

В. И. Старосельский

Физика полупроводниковых приборов микроэлектроники

УЧЕБНОЕ ПОСОБИЕ

*Рекомендовано учебно-методическим объединением
вузов Российской Федерации по образованию в
области радиотехники, электроники, биомедицинской
техники и автоматизации в качестве учебного пособия
для студентов высших учебных заведений, обучающихся
по направлению подготовки 210100 «Электроника
и микроэлектроника»*

МОСКВА • ЮРАЙТ • 2011



УДК 621.3
ББК 32.844/32.852я73
С77



Настоящая методическая разработка выполнена в рамках инновационной образовательной программы МИЭТ «Современное профессиональное образование для российской инновационной системы в области электроники»

Рецензенты:

Таиров Ю. М. — доктор технических наук, профессор, заведующий кафедрой микроэлектроники Санкт-Петербургского государственного электротехнического университета им. В. И. Ульянова (Ленина);

Перишков В. С. — доктор технических наук, профессор, заведующий кафедрой микроэлектроники Московского инженерно-физического института (государственного университета).

Физика полупроводниковых приборов микроэлектроники : учеб. пособие / В. И. Старосельский — М. : Издательство Юрайт ; ИД Юрайт, 2011. — 463 с. — (Основы наук).

ISBN 978-5-9916-0808-4 (Издательство Юрайт)

ISBN 978-5-9692-0962-6 (ИД Юрайт)

Рассмотрены базовые полупроводниковые приборы современной микроэлектроники и физические процессы, обеспечивающие их работу. Анализируются статические, частотные и импульсные характеристики приборов, рассматриваются методы схмотехнического моделирования приборов и приводятся их эквивалентные схемы. Рассмотрены предельные параметры современных приборов микроэлектроники. Для каждого прибора делается краткий обзор современных методов их структурной реализации в интегральных схемах.

Для студентов, обучающихся по направлению 210100 «Электроника и микроэлектроника» (210100.62 — бакалавр, 210100.68 — магистр) и по инженерным специальностям 210104.65 «Микроэлектроника и твердотельная электроника», 210108.65 «Микросистемная техника», 010803.65 «Микроэлектроника и полупроводниковые приборы», 210601.65 «Нанотехнологии в электронике». Материал книги может быть полезен также научным работникам, инженерам и аспирантам, стремящимся получить необходимые профессиональные знания.

УДК 621.3
ББК 32.844/32.852я73

ISBN 978-5-9916-0808-4
(Издательство Юрайт)
ISBN 978-5-9692-0962-6
(ИД Юрайт)

© Старосельский В. И., 2008
© ООО «ИД Юрайт», 2011



Оглавление

Предисловие редактора Ю. А. Парменова	11
Глава I. Основные сведения из физики полупроводников	13
I.1. Элементы зонной теории.....	13
I.2. Собственные и примесные полупроводники	18
I.3. Концентрации электронов и дырок	19
I.3.1. Равновесное состояние.....	19
I.3.2. Неравновесное состояние.....	27
I.3.3. Уровень инжекции.....	29
I.4. Кинетика носителей заряда	29
I.5. Рекомбинация и генерация носителей заряда.....	33
I.5.1. Механизмы рекомбинации—генерации	33
I.5.2. Прямая межзонная рекомбинация.....	34
I.5.3. Ловушечная рекомбинация	35
I.5.4. Ударная рекомбинация	36
I.5.5. Поверхностная рекомбинация.....	36
I.5.6. Превалирующие механизмы рекомбинации	37
I.6. Общий подход к анализу полупроводниковых приборов.....	38
I.7. Уравнения непрерывности.....	39
<i>Литература</i>	<i>44</i>
Глава II. p-n переходы и полупроводниковые диоды	45
II.1. Понятие о переходе	45
II.1.1. Определение и классификация p - n переходов.....	45
II.1.2. Структура p - n перехода.....	48
Основные выводы.....	50
II.2. Равновесное состояние p - n перехода.....	51
II.2.1. Энергетическая диаграмма p - n перехода.....	51
II.2.2. Токи через p - n переход в равновесном состоянии.....	53
II.2.3. Методика определения параметров p - n перехода	57
II.2.4. Расчет параметров ступенчатого p - n перехода.....	60
II.2.5. Переход с линейным распределением примеси	62
II.2.6. Диффузионные p - n переходы	65
Основные выводы.....	67
II.3. Неравновесное состояние p - n перехода	69
II.3.1. Прямое и обратное включение p - n перехода	69
II.3.2. Энергетические диаграммы неравновесного p - n перехода.....	69



П.3.3. Граничные условия и уровни инжекции	71
П.3.4. Ширина p - n перехода.....	74
Основные выводы.....	75
П.4. Анализ идеализированного диода	76
П.4.1. Модель идеализированного диода	76
П.4.2. Методика анализа ВАХ идеализированного диода	80
П.4.3. Распределения неосновных носителей заряда в квазинейтральных областях.....	81
П.4.4. ВАХ идеализированного диода	84
П.4.5. Тепловой ток	85
П.4.6. Температурная зависимость прямой ветви ВАХ диода	88
П.4.7. Характеристические сопротивления диода	89
П.4.8. Коэффициенты инжекции и эффективность эмиттера.....	90
Основные выводы.....	91
П.5. ВАХ реального диода	92
П.5.1. Особенности ВАХ реального диода.....	92
П.5.2. Термогенерация и рекомбинация носителей заряда в p - n переходе.....	92
П.5.3. Ток термогенерации (обратное смещение перехода) ..	94
П.5.4. Ток рекомбинации (прямое смещение перехода)	95
П.5.5. Сопротивление базы.....	97
П.5.6. Характеристики диода при высоком уровне инжекции	99
П.5.6.1. Особенности высокого уровня инжекции.....	99
П.5.6.2. Распределение носителей заряда в базе.....	100
П.5.6.3. Электрическое поле в базе.....	101
П.5.6.4. ВАХ p - n перехода.....	102
П.5.6.5. ВАХ диода.....	102
П.5.6.6. Эффективность эмиттера	103
П.5.6.7. Границы высокого и низкого уровней инжекции	104
П.5.7. ВАХ реального диода	104
Основные выводы.....	106
П.6. Пробой p - n перехода.....	107
П.6.1. Механизмы пробоя p - n перехода.....	107
П.6.2. Лавинный пробой.....	107
П.6.3. Туннельный (зенеровский) пробой	111
П.6.4. Особенности лавинного и туннельного механизмов пробоя.....	113
П.6.5. Тепловой пробой.....	114
Основные выводы.....	116



II.7. Частотные и импульсные свойства полупроводникового диода.....	117
II.7.1. Механизмы инерционности диода.....	117
II.7.2. Барьерная емкость p - n перехода	117
II.7.3. Диффузионные емкости диода.....	121
II.7.4. Частотные свойства диода на малом переменном сигнале	125
II.7.5. Импульсные свойства диода	128
Основные выводы.....	131
<i>Литература</i>	132
 Глава III. Транзисторы со структурой металл—диэлектрик— полупроводник	133
III.1. Эффект поля в структурах МДП.....	133
III.1.1. Структура МДП.....	133
III.1.2. Идеальная структура МДП в условиях термодинамического равновесия	134
III.1.3. Эффект поля в идеальной структуре МДП.....	136
III.1.4. Вольтфарадные характеристики идеальной структуры МДП	141
III.1.5. Особенности эффекта поля в реальной структуре МДП.....	144
III.1.6. Пороговое напряжение структуры МДП.....	145
Основные выводы.....	150
III.2. Устройство и принцип действия МДП-транзистора.....	152
III.2.1. Принцип действия МДПТ	152
III.2.2. Устройство и разновидности МДПТ	152
III.2.3. Качественный анализ характеристик МДПТ.....	155
III.2.4. Пороговое напряжение МДПТ и влияние потенциала подложки.....	159
Основные выводы.....	165
III.3. Анализ идеализированного транзистора.....	165
III.3.1. Допущения модели идеализированного МДПТ	165
III.3.2. Характеристики идеализированного транзистора... ..	167
III.3.3. Инерционные свойства идеализированного МДПТ.....	175
Основные выводы.....	178
III.4. Статические ВАХ МДП-транзисторов с длинным каналом	178
III.4.1. Влияние неоднородности ОПЗ под затвором.....	178
III.4.2. Подпороговый ток	182
III.4.3. Подвижность носителей заряда в канале	187
III.4.4. Влияние температуры.....	187
III.4.5. Умножение носителей в канале	189
Основные выводы.....	190



III.5. Эффекты короткого канала.....	191
III.5.1. Природа эффектов короткого канала.....	191
III.5.2. Пороговое напряжение	192
III.5.3. Эффект смыкания канала.....	195
III.5.4. Зарядка оксида	198
III.5.5. Ограничение дрейфовой скорости носителей в канале.....	198
III.5.6. Влияние ограничения дрейфовой скорости носителей в канале на основные свойства МДПТ	200
III.5.7. Модуляция длины канала	204
III.5.8. Пологая область ВАХ МДПТ и коэффициент усиления	208
III.5.9. Емкости затвор—исток и затвор—сток	212
III.5.10. Особенности ВАХ короткоканального МДПТ	213
III.5.11. Подпороговый ток.....	215
III.5.12. Критерий короткого канала.....	215
Основные выводы	215
Приложение III.1. Определение зависимости эффективной длины канала от напряжения сток-исток	216
III.6. Моделирование МДП-транзисторов	217
III.6.1. Классификация моделей полупроводниковых приборов.....	219
III.6.2. Особенности моделирования МДПТ	219
III.6.3. Нелинейная эквивалентная схема МДПТ для большого сигнала	222
III.6.4. Линейные эквивалентные схемы МДПТ для малого сигнала	224
III.6.5. Формальные линейные модели.....	231
Основные выводы	235
III.7. Масштабирование МДП-транзисторов	238
III.8. Структуры короткоканальных МДП-транзисторов.....	244
III.8.1. Требования к МДПТ в СБИС	244
III.8.2. Типовая структура короткоканального МДПТ.....	245
III.8.3. Структурные модификации короткоканальных МДПТ	249
<i>Литература</i>	252
Глава IV. Биполярные транзисторы	254
IV.1. Устройство и принцип действия биполярного транзистора	254
IV.1.1. Устройство биполярного транзистора.....	254
IV.1.2. Принцип действия и режимы работы транзистора	256
IV.1.3. Разновидности биполярных транзисторов	259
IV.1.4. Основные физические процессы в транзисторе (нормальный режим).....	263



IV.1.5. Эффективность эмиттера и коэффициент переноса.....	266
Основные выводы.....	267
IV.2. Модель идеализированного транзистора Эберса—Молла.....	267
IV.2.1. Модель Эберса—Молла.....	267
IV.2.2. Статические характеристики идеализированного транзистора в схеме ОБ.....	274
IV.2.3. Статические характеристики идеализированного транзистора в схеме ОЭ.....	276
IV.2.4. Тепловые токи коллектора и эмиттера.....	279
Основные выводы.....	281
IV.3. Параметры идеализированного транзистора.....	282
IV.3.1. Коэффициент переноса.....	283
IV.3.2. Тепловые токи эмиттерного диода и эффективность эмиттера.....	286
IV.3.3. Влияние коэффициента переноса и эффективности эмиттера на усилительные свойства транзистора.....	287
IV.3.4. Инверсные параметры.....	288
Основные выводы.....	291
IV.4. Частотные и импульсные свойства коэффициентов передачи тока.....	292
IV.4.1. Механизм внутренней инерционности биполярного транзистора и форма представления частотных и импульсных характеристик.....	292
IV.4.2. Частотная характеристика коэффициента переноса.....	294
IV.4.3. Частотная характеристика эффективности эмиттера.....	295
IV.4.4. Частотные и импульсные свойства коэффициента передачи эмиттерного тока.....	297
IV.4.5. Частотные и импульсные свойства коэффициента усиления тока базы.....	299
IV.4.6. Сравнительный анализ частотных и импульсных характеристик коэффициентов α и β	301
IV.4.7. Диффузионные емкости в транзисторе.....	303
IV.4.8. Накопление заряда в коллекторе.....	305
Основные выводы.....	307
IV.5. Действие факторов, не учтенных в идеализированной модели транзистора.....	308
IV.5.1. Факторы, не учтенные в идеализированной модели.....	308
IV.5.2. Эффект Эрли.....	309



IV.5.3. Сопротивления базы и коллектора.....	316
IV.5.4. Эффект оттеснения эмиттерного тока.....	317
IV.5.5. Влияние режима работы транзистора на коэффициенты передачи тока	319
IV.5.6. Ограничение скорости носителей заряда в базе и в коллекторном переходе.....	321
IV.5.7. Специфика пробоя в биполярных транзисторах.....	326
IV.5.8. Диодное включение транзисторов	328
Основные выводы.....	330
IV.6. Особенности дрейфовых планарных транзисторов	331
IV.6.1. Примесный профиль планарного дрейфового транзистора и встроенные электрические поля	332
IV.6.2. Распределение избыточных носителей заряда в базе	336
IV.6.3. Время пролета неосновных носителей через базу и тепловые токи.....	339
IV.6.4. Коэффициент передачи эмиттерного тока.....	342
IV.6.5. Тиристорный эффект.....	344
Основные выводы.....	348
IV.7. Моделирование биполярных транзисторов.....	349
IV.7.1. Особенности моделирования биполярных транзисторов	349
IV.7.2. Эквивалентные схемы для большого сигнала на основе модели Эберса—Молла	350
IV.7.3. Метод Гуммеля—Пуна.....	351
IV.7.4. Модель Гуммеля—Пуна	357
IV.7.5. Эквивалентные схемы для малого сигнала.....	359
IV.7.6. Формальные линейные модели.....	364
Основные выводы.....	366
IV.8. Структуры интегральных биполярных транзисторов.....	367
<i>Литература</i>	375
Глава V. Сравнительный анализ принципов действия и свойств биполярных и полевых транзисторов.....	377
V.1. Сходства и различия в устройстве и принципе действия МДП- и биполярных транзисторов	377
V.2. Основные электрические свойства МДП- и биполярных транзисторов в рамках классических моделей	379
V.3. Эффекты, ограничивающие характеристики реальных МДП- и биполярных транзисторов	381
V.4. Важнейшие характеристики МДП- и биполярных транзисторов	381
V.5. Анализ важнейших характеристик МДП- и биполярных транзисторов	384



V.6. Области применения МДП- и биполярных транзисторов	391
Глава VI. Контакты металл—полупроводник	393
VI.1. Контакты металл—полупроводник.....	393
VI.1.1. Типы и основные характеристики контактов металл—полупроводник	393
VI.1.2. Параметры барьерных контактов	394
VI.1.3. Эффект Шоттки	397
VI.2. Теория выпрямления на контакте металл—полупроводник	399
VI.2.1. Общий подход.....	399
VI.2.2. Теория термоэлектронной эмиссии (диодная теория выпрямления).....	400
VI.2.3. Диффузионная теория выпрямления.....	402
VI.2.4. Сравнение результатов ТТЭ и диффузионной теории выпрямления	405
VI.3. Омические контакты	405
VI.4. Диоды Шоттки	407
Основные выводы.....	410
Литература	411
Глава VII. Полевые транзисторы с управляющим барьерным переходом	412
VII.1. Полевые транзисторы с управляющим <i>p-n</i> переходом.....	412
VII.1.1. Устройство и принцип действия.....	412
VII.1.2. ВАХ идеализированного транзистора.....	414
VII.1.3. Эквивалентная схема и особенности ВАХ	417
VII.1.4. Сравнение полевого транзистора с управляющим <i>p-n</i> переходом и МДПТ.....	418
Основные выводы.....	420
VII.2. Полевые транзисторы с затвором Шоттки на основе GaAs.....	420
VII.2.1. Особенности GaAs как материала микроэлектроники.....	421
VII.2.2. Устройство и особенности ПТШ на GaAs.....	423
VII.2.3. Особенности характеристик и применения ПТШ	425
Основные выводы.....	427
Литература	428
Глава VIII. Гетеропереходные транзисторы.....	429
VIII.1. Гетеропереходы	429
VIII.1.1. Особенности гетеропереходов.....	430



VIII.1.2. Гетеропереходы на основе GaAs.....	432
VIII.1.3. Инверсионный слой в гетеропереходе и двухмерный электронный газ.....	434
Основные выводы.....	436
VIII.2. Гетеропереходные полевые транзисторы на основе GaAs.....	437
VIII.2.1. Устройство и принцип действия ГПТ на основе GaAs	437
VIII.2.2. Пороговое напряжение.....	439
VIII.2.3. ВАХ ГПТ	441
VIII.2.4. Разновидности ГПТ на основе соединений A^3B^5	442
VIII.2.5. ГПТ на основе структур Si/SiGe	444
Основные выводы.....	446
VIII.3. Гетеропереходные биполярные транзисторы.....	446
VIII.3.1. Возможности улучшения характеристик биполярных транзисторов за счет использования гетеропереходов.....	446
VIII.3.2. ГBT на основе GaAs	449
VIII.3.3. ГBT на основе гетероструктур Si-Ge/Si	454
Основные выводы.....	455
<i>Литература</i>	455
ПРИЛОЖЕНИЯ	458
Приложение 1. Физические константы	458
Приложение 2. Список общих обозначений	458
Приложение 3. свойства некоторых полупроводниковых материалов ($T = 300\text{ K}$)	460
Приложение 4. Энергии ионизации примесей в Si, Ge и GaAs	462
Приложение 5. Греческий алфавит	462
Литература к приложениям 3, 4	463



ПРЕДИСЛОВИЕ РЕДАКТОРА

Автор настоящей книги задумывал ее как учебное пособие для студентов факультета электроники и компьютерных технологий Московского государственного института электронной техники (технического университета), обучающихся по направлению 210100 «Электроника и микроэлектроника» (210100.62 — бакалавр, 210100.68 — магистр) и по инженерным специальностям 210104.65 «Микроэлектроника и твердотельная электроника», 210108.65 «Микросистемная техника», 010803.65 «Микроэлектроника и полупроводниковые приборы», 210601.65 «Нанотехнологии в электронике» и специализирующихся в области проектирования и технологии изготовления полупроводниковых интегральных схем.

В учебном пособии использованы материалы лекционных курсов «Элементы твердотельной электроники и микроэлектроники», «Физика полупроводниковых приборов» «Физика полупроводников и полупроводниковых приборов», а также факультативных курсов, прочитанных автором студентам МИЭТ (ТУ) на протяжении более 30 лет. Материалы дополнены новыми сведениями в области теории и проектирования полупроводниковых приборов с субмикронными размерами активных областей и приборов на новых полупроводниковых материалах микроэлектроники.

В то же время содержание учебного пособия по объему и глубине изложения выходит далеко за пределы учебных планов указанных направлений (специальностей). В книге нашли отражения также и результаты многочисленных научных исследований автора в области физики кремниевых и арсенидгаллиевых полупроводниковых приборов, в том числе МДП- и биполярных транзисторов, гетеропереходных полевых транзисторов, СВЧ приборов. Обладая разносторонними знаниями как в области физики, так и в области проектирования полупроводниковых приборов, автор при изложении материала обращает внимание на особенности использования изучаемого прибора в интегральных схемах, исследует предельные возможности приборов, большое внимание уделяет рассмотрению моделей приборов для схемотехнического применения.

К сожалению, автор не успел увидеть свою книгу изданной. Книга редактировалась уже после его скоропостижной кончины. Этим он лишил нас возможности проводить научные и методические дискуссии по вопросам, возникающим в процессе редактирования.

Отдавая дань научной и педагогической деятельности Виктора Игоревича Старосельского, редактор стремился сохранить рукопись в том



виде, в каком она была написана автором и не исправлял ее даже в том случае, если придерживался другой точки зрения. Коррективы подвергались только явные опечатки и неточности в формулах и тексте.

В полном виде учебное пособие предназначено для студентов, освоивших курсы физики твердого тела и физики полупроводников, хотя для удобства пользования в него включена глава, содержащая краткие сведения из физики полупроводников.

Ограниченный объем учебного пособия позволяет охватить только базовые элементы современной микроэлектроники: p - n переходы, полупроводниковые диоды, транзисторы со структурой металл—диэлектрик—полупроводник, биполярные транзисторы, контакты металл—полупроводник, полевые транзисторы с управляющим барьерным переходом и гетеропереходные полевые и биполярные транзисторы.

Свойства всех полупроводниковых приборов, включенных в материал учебного пособия, рассматриваются с единых методических позиций. Вначале описываются устройство и принцип действия приборов, анализируются их статические характеристики в рамках максимально упрощенной идеализированной модели, после чего рассматриваются особенности реальных приборов, не учтенные простой моделью. Затем анализируются механизмы внутренней инерционности прибора, определяющие его частотные и импульсные характеристики. Большое внимание уделяется анализу эффектов малых размеров, которые определяют предельные характеристики современных приборов микроэлектроники. В заключение рассматриваются методы схемотехнического моделирования приборов и приводятся их эквивалентные схемы. Для каждого прибора делается краткий обзор современных методов их структурной реализации в интегральных схемах. Отдельные разделы посвящены масштабированию МДП-транзисторов и сравнительному анализу принципов действия, свойств и областей применения наиболее важных активных приборов современной микроэлектроники — полевых и биполярных транзисторов.

Для удобства пользования в приложениях даны главные сведения о физических свойствах основных полупроводниковых материалов микроэлектроники.

В целом учебное пособие безусловно полезно не только студентам указанных выше направлений (специальностей), но также и в сокращенном виде для обучения студентов родственных направлений, изучающих элементы твердотельной электроники. Материал книги может быть полезен научным работникам инженерам и аспирантам, не получившим достаточно широкого образования в области физики полупроводниковых приборов и стремящимся получить необходимые профессиональные знания.

Ю. А. Парменов



Глава I

ОСНОВНЫЕ СВЕДЕНИЯ ИЗ ФИЗИКИ ПОЛУПРОВОДНИКОВ

Поскольку по физике полупроводников имеется обширная литература [1—3 и др.], в этом разделе с соответствующими ссылками конспективно изложены лишь те сведения, которые будут использованы при анализе свойств полупроводниковых приборов, а также даны необходимые определения. Основное внимание уделено полупроводниковым материалам, которые широко применяются для изготовления электронных приборов — моноатомным полупроводникам кремнию (Si) и германию (Ge), а также полупроводниковым соединениям A^3B^5 , из которых наиболее освоен арсенид галлия (GaAs). Математические формулы максимально упрощены с учетом специфики этих материалов, а также условий эксплуатации изделий (ограниченный температурный диапазон). Далее в книге приведенный материал будет использован для необходимых ссылок, чтобы не перегружать выкладками основной текст.

I.1. Элементы зонной теории

В твердых телах с упорядоченной кристаллической структурой электроны имеют дискретный энергетический спектр. Разрешенные энергетические состояния группируются в *разрешенные энергетические зоны*, между которыми находятся *запрещенные зоны* — энергетические интервалы, недоступные для электронов. С учетом спинового вырождения число разрешенных s -состояний в разрешенной зоне равно удвоенному количеству атомов в кристалле, поэтому в единице объема каждая энергетическая зона содержит $N_{st} = 2N_{at}$ разрешенных состояний, где $N_{at} \approx 10^{22} \text{ см}^{-3}$ — концентрация атомов в кристалле. Ширина разрешенных зон возрастает с увеличе-



нием энергии. Если две разрешенные зоны перекрываются, число разрешенных состояний в расширенной разрешенной зоне составляет $N_{st} = 4N_{at}$.

Вероятность заполнения электроном разрешенного уровня снижается с увеличением энергии уровня и возрастает с ростом температуры. Поэтому разрешенные зоны, соответствующие достаточно высоким энергиям, остаются пустыми вплоть до температуры плавления кристалла. Последняя разрешенная зона, содержащая электроны, называется *зоной проводимости*, так как находящиеся в ней электроны могут передвигаться в каком-либо выделенном направлении (участвовать в проводимости) и создавать электрический ток. Электроны, заполняющие разрешенную зону полностью, участвовать в проводимости не могут.

В металлах зона проводимости заполнена частично, а расположенная под ней разрешенная зона — полностью. Поэтому электропроводность металлов обусловлена только электронами проводимости и весьма велика.

В нелегированных (собственных) полупроводниках при нулевой абсолютной температуре зона проводимости пуста, а находящаяся под ней *валентная зона* заполнена полностью. При этом электропроводность полупроводника равна нулю. При конечной температуре некоторое количество электронов переходит из валентной зоны в зону проводимости.

В твердом теле электроны *обобществлены всем кристаллом*, т.е. не принадлежат какому-либо конкретному атому. Поведение электрона описывается волновой функцией ψ , которая является решением уравнения Шредингера. В идеальном кристалле на электрон действует самосогласованное электрическое поле, которое является суммой полей ионов решетки и всех электронов кристалла. Зонная теория твердого тела предполагает, что потенциальная энергия электрона в этом поле является периодической функцией с периодом решетки. При этом уравнение Шредингера имеет решение в виде функции Блоха

$$\Psi_{\vec{k}}(\vec{r}) = \psi_{\vec{k}}(\vec{r})e^{i\vec{k}\vec{r}},$$

где $\psi_{\vec{k}}$ — периодическая функция с периодом решетки кристалла; $\vec{k}$ — волновой вектор, $\vec{r}$ — радиус-вектор.

Таким образом, волновая функция $\Psi_{\vec{k}}(\vec{r})$ представляет собой плоскую волну $e^{i\vec{k}\vec{r}}$ с периодически модулированной амплитудой.



Для свободного электрона с потенциальной энергией E_0 модуляция амплитуды плоской волны отсутствует. В этом случае энергия электрона E и его импульс $\vec{p}$ составляют

$$E = E_0 + \hbar\omega = E_0 + m_0 v_n^2; \quad (I.1.1)$$

$$\vec{p} = \hbar\vec{k} = m_0 v_n, \quad (I.1.2)$$

где m_0 — масса электрона; v_n — групповая скорость.

Зависимость

$$E = E_0 + \hbar\omega = E_0 + p^2 / 2m_0 = E_0 + \hbar^2 k^2 / 2m_0 \quad (I.1.3)$$

представляет собой закон дисперсии свободного электрона.

Из (I.1.3) следует, что групповая скорость свободного электрона составляет

$$\vec{v}_n = d\omega / d\vec{k} = dE / d(\hbar\vec{k}) = dE / d\vec{p}. \quad (I.1.4)$$

Под действием внешней силы $\vec{F}$ импульс и волновой вектор свободного электрона изменяются по закону

$$\vec{F} = d\vec{p} / dt = d(\hbar\vec{k}) / dt. \quad (I.1.5)$$

В полупроводниковом кристалле закон дисперсии (I.1.3) искажен полем периодического потенциала. Он может быть представлен в виде многозначной функции $E(\vec{k})$, полностью определенной в зоне Бриллюэна — области $\vec{k}$ -пространства, размеры которой определяются периодом кристаллической решетки $\vec{a}$. Для одномерной решетки $-\pi/a \leq k \leq \pi/a$. Для трехмерной решетки границы зоны Бриллюэна находятся в определенном соответствии с границами кристаллической ячейки в координатном пространстве. Многозначность функции $E(\vec{k})$ соответствует разделению энергетического спектра на разрешенные зоны.

В зоне проводимости электроны заполняют разрешенные состояния вблизи минимальной энергии E_c (дно зоны проводимости). Энергии E_c соответствует значение волнового вектора $\vec{k}_c$. Если $\vec{k}_c = 0$ и $E(\vec{k}) = E(k)$ (т.е. энергетический минимум находится в центре зоны Бриллюэна и функция $E(\vec{k})$ не зависит от направления волнового вектора), то в окрестности точки $\vec{k} = \vec{k}_c = 0$ функция $E(\vec{k})$ может быть представлена в виде

$$E(\vec{k}) = E_c + \frac{1}{2} \cdot \left[\frac{\partial^2 E}{\partial k_x^2} k_x^2 + \frac{\partial^2 E}{\partial k_y^2} k_y^2 + \frac{\partial^2 E}{\partial k_z^2} k_z^2 \right], \quad (I.1.6a)$$



$$\text{или} \quad E = E_c + \hbar^2 k^2 / 2m_n, \quad (\text{I.1.66})$$

$$\text{где} \quad m_n = \hbar^2 / d^2 E / dk^2 \Big|_{k=0}. \quad (\text{I.1.7})$$

Сравнение соотношений (I.1.66) и (I.1.3) показывает, что вблизи дна зоны проводимости электроны могут рассматриваться как *квазичастицы — электроны проводимости*, свойства которых отличаются от свойств свободных электронов только значением *эффективной массы*, определенной в (I.1.7). Энергия E_c может считаться *потенциальной энергией электрона проводимости*, энергия $\hbar^2 k^2 / 2m_n = p_n^2 / 2m_n$ — его *кинетической энергией*, а величина

$$\vec{p}_n = m_n \vec{v}_n = \hbar \vec{k} \quad (\text{I.1.8})$$

— *квазиимпульсом электрона проводимости*. При этом величина $\vec{v}_n$ сохраняет значение групповой скорости, а соотношение (I.1.5) остается справедливым.

В общем случае $\vec{k}_c \neq 0$ и обратная величина эффективной массы является симметричным тензором, который в диагональном виде имеет компоненты m_{n1}^{-1} , m_{n2}^{-1} , m_{n3}^{-1} . При этом

$$E = E_c + \frac{\hbar^2}{2} \left[\frac{(k_x - k_{cx})^2}{m_{n1}} + \frac{(k_y - k_{cy})^2}{m_{n2}} + \frac{(k_z - k_{cz})^2}{m_{n3}} \right]; \quad (\text{I.1.9})$$

$$\vec{p}_n = \hbar(\vec{k} - \vec{k}_c). \quad (\text{I.1.10})$$

Для полупроводников типа A^3B^5 (например, GaAs) энергетический минимум $E = E_c$ находится в центре зоны Бриллюэна ($\vec{k}_c = 0$), а эффективная масса электрона может считаться скаляром ($m_{n1} = m_{n2} = m_{n3} = m_n$). Для монокристаллических полупроводников (Si, Ge) энергетический минимум $E = E_c$ находится вблизи края зоны Бриллюэна, $m_{n1} = m_{n2} = m_{nt}$ (поперечная эффективная масса) и $m_{n3} = m_{nl}$ (продольная эффективная масса).

В валентной зоне незаполненными электронами остаются только разрешенные состояния вблизи максимальной энергии E_v (потолок валентной зоны). При этом оказывается удобным рассматривать перемещения этих незаполненных состояний, введя понятие *квазичастиц — дырок* в валентной зоне. Заряд дырки *положителен* и равен заряду свободного электрона с противоположным знаком: $q_p = e$. Закон дисперсии дырок симметричен закону дисперсии электронов (I.1.66), так как энергии E_v обычно соответствует нулевое значение волнового



вектора $\vec{k}_v = 0$, и $E(\vec{k}) = E(k)$. Точка $\vec{k}_v = 0$ является точкой вырождения — в ней смыкаются две дисперсионные ветви

$$E_1 = E_v - \hbar^2 k^2 / 2m_{pl}; \quad (I.1.11a)$$

$$E_2 = E_v - \hbar^2 k^2 / 2m_{ph}, \quad (I.1.11б)$$

где $m_{pl} > 0$ и $m_{ph} > m_{pl}$ — эффективные массы легких и тяжелых дырок. Энергия E_v может считаться потенциальной энергией дырок, энергии $\hbar^2 k^2 / 2m_{pl,ph} = p_{pl,ph}^2 / 2m_{pl,ph}$ — их кинетической энергией, а величины $\vec{p}_{pl,ph} = m_{pl,ph} \vec{v}_{pl,ph} = \hbar \vec{k}$ — квазиимпульсами дырок. Величины $\vec{v}_{pl,ph}$ сохраняют значения групповых скоростей.

Вдали от границ разрешенных зон законы дисперсии квазичастиц и свободных электронов качественно различны.

Подробная информация о свойствах квазичастиц в полупроводниках содержится, например, в работах [4, 5].

Состояние электронов и дырок удобно представлять с помощью энергетических диаграмм, на которых изображаются только окрестности запрещенной зоны шириной

$$E_g = E_c - E_v.$$

В полупроводниках ширина запрещенной зоны составляет около 1 эВ. Этим они отличаются от диэлектриков, в которых ширина запрещенной зоны составляет несколько эВ.

При общей концентрации разрешенных состояний в зоне проводимости N_{sc} их концентрация в диапазоне энергии от E_c до $E > E_c$ составляет $N_{sc}(E)$. Плотностью состояний в зоне проводимости называется концентрация разрешенных энергетических уровней в единичном интервале энергии $g_c(E) = dN_{sc}(E) / dE$. Вблизи дна зоны проводимости

$$g_c(E) = \sqrt{2}(E - E_c)^{1/2} m_{nd}^{3/2} / \pi^2 \hbar^3, \quad (I.1.12a)$$

где $m_{nd} = (M_c^2 m_{n1} m_{n2} m_{n3})^{1/3}$ — эффективная масса плотности состояний для электронов; M_c — число эквивалентных минимумов энергии в зоне Бриллюэна. Аналогично, вблизи потолка валентной зоны

$$g_v(E) = \sqrt{2}(E_v - E)^{1/2} m_{pd}^{3/2} / \pi^2 \hbar^3, \quad (I.1.12б)$$

где $m_{pd} = (m_{pl}^{3/2} + m_{ph}^{3/2})^{2/3}$ — эффективная масса плотности состояний для дырок.



I.2. Собственные и примесные полупроводники

Собственными называются полупроводники, не содержащие примесных атомов. Концентрации электронов и дырок в собственных полупроводниках одинаковы и для определенного полупроводника зависят только от температуры:

$$n_i = n_i(T).$$

Примесными называются полупроводники, легированные определенными примесными атомами с целью регулировки концентраций электронов n и дырок p . Для этого используются примесные атомы — *доноры* и *акцепторы*, создающие в запрещенной зоне *мелкие примесные уровни*. Доноры с концентрацией N_d создают примесные уровни E_d , расположенные вблизи дна зоны проводимости, и характеризуются энергией активации

$$\Delta E_d = E_c - E_d.$$

Акцепторы с концентрацией N_a создают примесные уровни E_a , расположенные вблизи потолка валентной зоны и имеющие энергию активации

$$\Delta E_a = E_a - E_v.$$

Энергии активации обычно не превышают 0,1 эВ. Примесные уровни и энергии активации показаны на рис. I.2.1.

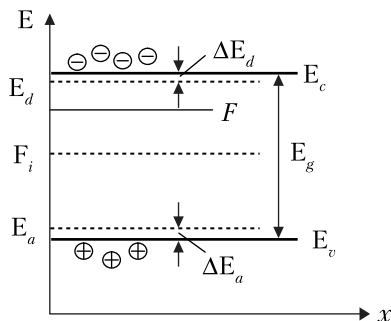


Рис. I.2.1. Энергетическая диаграмма полупроводника:
F — уровень Ферми в полупроводнике n -типа

Валентная оболочка доноров содержит на один электрон больше, а акцепторов — меньше, чем у атомов полупроводника. Поэтому доноры и акцепторы могут ионизироваться,

отдавая один электрон в зону проводимости или поглощая один электрон из валентной зоны. При этом нейтральные атомы доноров и акцепторов превращаются в *примесные ионы* с концентрациями, соответственно, N_d^+ и N_a^- .

Если электронейтральный полупроводник содержит только донорную примесь, концентрация электронов превышает концентрацию дырок

$$n = p + N_d^+ > p.$$

В случае легирования акцепторами концентрация дырок больше, чем электронов:

$$p = n + N_a^- > n.$$

В первом случае говорят о полупроводнике *n*-типа, электроны являются *основными носителями*, а дырки — *неосновными*. Во втором случае полупроводник имеет *p*-тип проводимости, основными носителями являются дырки, а неосновными — электроны.

Если концентрации донорных и акцепторных ионов одинаковы, то избыточные электроны доноров поглощаются акцепторами. В этом случае полупроводник называется *компенсированным*, а концентрации электронов и дырок равны собственным концентрациям n_i .

В электронных приборах полупроводник обычно содержит оба типа примесей. При этом тип его проводимости зависит от того, каких примесных ионов больше: при $N_d^+ > N_a^-$ реализуется *n*-тип, а при $N_a^- > N_d^+$ — *p*-тип проводимости.

1.3. Концентрации электронов и дырок

1.3.1. Равновесное состояние

В состоянии термодинамического равновесия функции распределения электронов и дырок по энергиям соответствуют статистике Ферми—Дирака:

$$f_{n0}(E) = \frac{1}{1 + e^{(E-F)/kT}}; \quad (1.3.1a)$$

$$f_{p0}(E) = 1 - f_{n0}(E) = \frac{1}{1 + e^{(F-E)/kT}}. \quad (1.3.1б)$$



Функции распределения имеют два параметра: *энергию Ферми* F и *температуру* T . Функции распределения симметричны относительно точки $E = F$: $f_0(E - F) + f_0(F - E) = 1$, причем вероятность заполнения энергетического уровня $E = F$ равна $1/2$. Температура характеризует степень «размытости» функций распределения относительно оси $E = F$ (рис. I.3.1). Полупроводник называется *невыврожденным*, если уровень Ферми расположен внутри запрещенной зоны на достаточном расстоянии от ее границ: $E_c - F > \approx 3kT$, $F - E_v > \approx 3kT$. В этом случае $f_{n0} \ll 1$, $f_{p0} \ll 1$, и функции распределения соответствуют статистике Максвелла–Больцмана

$$f_{n0}(E) = e^{(F-E)/kT}; \quad (I.3.2a)$$

$$f_{p0}(E) = e^{(E-F)/kT}. \quad (I.3.2б)$$

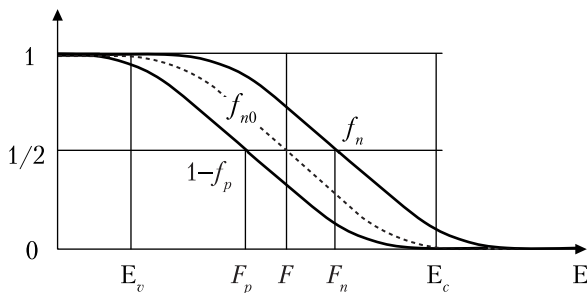


Рис. I.3.1. Функции распределения:

f_{n0} — (I.3.1a), f_n — (I.3.25a) и $1-f_p$ — (I.3.25б)

С учетом (I.1.12a, б) функции распределения позволяют связать с энергией Ферми равновесные концентрации электронов и дырок

$$n_0 = \int_{E_c}^{\infty} g_c(E) f_{n0}(E) dE; \quad (I.3.3a)$$

$$p_0 = \int_{-\infty}^{E_v} g_v(E) f_{p0}(E) dE. \quad (I.3.3б)$$

В (I.3.3a, б) учтено, что электроны занимают только энергетические уровни вблизи дна зоны проводимости, а дырки — вблизи потолка валентной зоны.

Собственный полупроводник может быть вырожден только при очень высокой температуре, находящейся за пределами температурного диапазона эксплуатации полупроводниковых приборов (самый жесткий «военный» диапазон: $-60 - +125^\circ\text{C}$). Поэтому энергию Ферми F_i , а также собственные концентрации электронов и дырок можно найти с использованием функций распределения (1.3.2а, б) и соотношений (1.3.3а, б) при условии $n_0 = p_0 = n_i$

$$F_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \frac{m_{pd}}{m_{nd}}; \quad (1.3.4)$$

$$n_i = (N_c N_v)^{1/2} e^{-E_g/2kT}, \quad (1.3.5)$$

где $N_{c,v} = 2(2\pi m_{nd,pd} kT / h^2)^{3/2}$ — эффективные плотности квантовых состояний в зоне проводимости и в валентной зоне, имеющие значения $\sim 10^{19} \text{ см}^{-3}$.

Энергия Ферми F_i в собственном полупроводнике называется *электростатической энергией*. Из (1.3.4) следует, что в диапазоне эксплуатации полупроводниковых приборов уровень электростатической энергии находится вблизи середины запрещенной зоны (см. рис. 1.2.1).

В примесном полупроводнике для нахождения энергии Ферми, а также концентраций электронов и дырок следует использовать уравнения (1.3.3а, б), функции распределения (1.3.1а, б) и условие электронейтральности

$$n_0 + N_a^- = p_0 + N_d^+, \quad (1.3.6)$$

а также уравнения для коэффициентов ионизации примесей N_d^+ / N_d и N_a^- / N_a . Последние определяются вероятностями заполнения примесных уровней E_d и E_a соответственно электронами и дырками:

$$N_d^+ / N_d = 1 - f_{nd0}(E_d); \quad (1.3.7а)$$

$$N_a^- / N_a = 1 - f_{pa0}(E_a). \quad (1.3.7б)$$

Функции $f_{nd0}(E_d)$ и $f_{pa0}(E_a)$ аналогичны функциям распределения (1.3.1а, б) с поправкой на фактор вырождения [1, 3].

В случае когда концентрация основных носителей много меньше соответствующей эффективной плотности квантовых состояний ($p_0 < n_0 \ll N_c$ или $n_0 < p_0 \ll N_v$), полупроводник является *невыврожденным* и допустимо использование функ-

ций распределения (I.3.2а, б). В этом случае из (I.3.3а, б) и (I.3.2а, б) следует, что

$$n_0 = N_c e^{(F - E_c)/kT} = n_i e^{(F - F_i)/kT}; \quad (I.3.8a)$$

$$p_0 = N_v e^{(E_v - F)/kT} = n_i e^{(F_i - F)/kT}; \quad (I.3.8б)$$

$$n_0 p_0 = n_i^2. \quad (I.3.9)$$

Соотношение (I.3.9) представляет собой закон действующих масс для невырожденного полупроводника.

Для наших целей нет необходимости решать полную систему уравнений, определяющих равновесные концентрации электронов и дырок. Рассмотрим лишь случаи, практически важные для полупроводниковых приборов.

1. Однородный невырожденный полупроводник с одним типом примеси

Положим для определенности, что полупроводник легирован только донорами. В этом случае уровень Ферми расположен выше уровня F_i , однако ниже примесного уровня E_d (см. рис. I.2.1). Поэтому *практически все донорные атомы ионизованы*: $N_d^+ = N_d$.

Если концентрация основных носителей (в данном случае электронов) в какой-либо области прибора не сказывается заметным образом на его характеристиках, определять ее нет необходимости. В противном случае она не должна зависеть от температуры. Это достигается достаточно сильным легированием полупроводникового слоя, когда концентрация доноров значительно превышает собственную концентрацию носителей при максимальной рабочей температуре: $N_d \gg n_{i\max}$ (в кремниевых ИС концентрация примеси в наименее легированном слое — подложке — составляет не менее 10^{15} см^{-3}). При этом равновесная концентрация основных носителей (электронов) практически равна концентрации доноров, а равновесная концентрация неосновных (дырок) определяется законом действующих масс (I.3.9)

$$n_0 = N_d; \quad p_0 = n_i^2 / n_0 = n_i^2 / N_d.$$

Аналогично для невырожденного полупроводника p -типа, легированного только акцепторами, справедливы соотношения

$$p_0 = N_a; \quad n_0 = n_i^2 / p_0 = n_i^2 / N_a.$$



2. Однородный невырожденный полупроводник с двумя типами примеси

Положим для определенности, что мы имеем дело с полупроводником n -типа: $N_d > N_a$. Положение уровня Ферми соответствует предыдущему случаю (см. рис. 1.2.1). При этом ионизованы как акцепторные, так и донорные уровни. Акцепторные уровни E_a заполнены электронами с донорных уровней, а оставшиеся электроны с концентрацией $N_d - N_a$ переходят с донорных уровней E_d в зону проводимости. Независимость концентрации основных носителей (электронов) от температуры достигается выполнением условия

$$N \gg n_{i\max},$$

где $N = N_d - N_a$ — результирующая концентрация примеси; $n_{i\max}$ — собственная концентрация при максимальной рабочей температуре.

При этом равновесная концентрация основных носителей (электронов) практически равна результирующей концентрации примеси, а равновесная концентрация неосновных (дырок) определяется законом действующих масс (1.3.9)

$$n_0 = N; \quad p_0 = n_i^2 / n_0 = n_i^2 / N.$$

Аналогично для невырожденного полупроводника p -типа при $N_a > N_d$ справедливы соотношения

$$p_0 = N; \quad n_0 = n_i^2 / p_0 = n_i^2 / N,$$

где $N = N_a - N_d$ — результирующая концентрация примеси.

Таким образом, для *однородного невырожденного полупроводника* можно сделать следующие общие выводы.

Полупроводник является невырожденным, если *результурующая концентрация примеси* $N = |N_d - N_a|$ много меньше соответствующей эффективной плотности квантовых состояний ($N \ll N_c$ для n -типа, $N \ll N_v$ для p -типа):

$$N = |N_d - N_a| \ll \begin{cases} N_c, & n\text{-тип,} \\ N_v, & p\text{-тип.} \end{cases} \quad (1.3.10)$$

Равновесная концентрация основных носителей может считаться равной результирующей концентрации примеси, а равновесная концентрация неосновных носителей определяется законом действующих масс (1.3.9).



Для полупроводника n -типа

$$\begin{cases} n_0 = N \gg n_i, \\ p_0 = n_i^2 / n_0 = n_i^2 / N \ll n_i. \end{cases} \quad (\text{I.3.11a})$$

Для полупроводника p -типа

$$\begin{cases} p_0 = N \gg n_i, \\ n_0 = n_i^2 / p_0 = n_i^2 / N \ll n_i. \end{cases} \quad (\text{I.3.11б})$$

Положение уровня Ферми может быть установлено из соотношений (I.3.11) и (I.3.8).

3. Однородный вырожденный полупроводник

Рассмотрим для определенности сильно вырожденный полупроводник n^+ -типа с концентрацией доноров $N_d \gg N_c$. Особенностью такого полупроводника является расщепление примесного донорного уровня E_d в примесную зону, которая перекрывается с зоной проводимости, а также расположение уровня Ферми F внутри результирующей разрешенной зоны. Эта ситуация показана на рис. I.3.2, *а*, где изображена зонная диаграмма полупроводника n^+ -типа и зависимость от энергии плотности разрешенных состояний в зоне проводимости $g_c(E)$, примесной зоне $g_d(E)$ и результирующей разрешенной зоне $g(E) = g_c(E) + g_d(E)$. Для

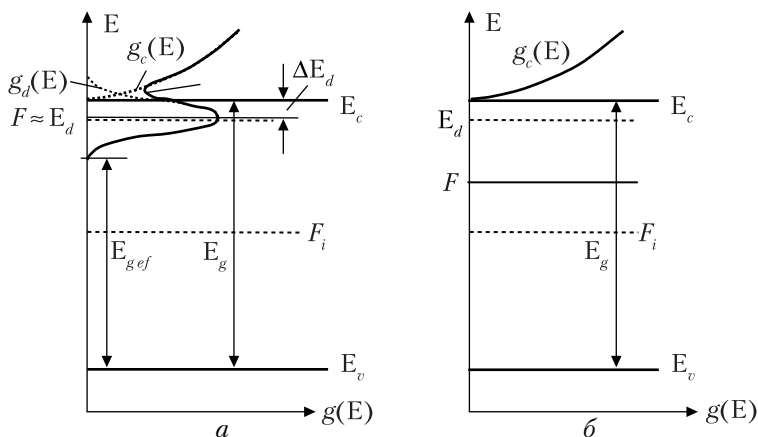


Рис. I.3.2. Зонная диаграмма полупроводника n^+ -типа при $N_d \gg N_c$ и зависимость от энергии плотности разрешенных состояний в зоне проводимости $g_c(E)$, примесной зоне $g_d(E)$ и результирующей разрешенной зоне $g(E) = g_c(E) + g_d(E)$ (*а*); зонная диаграмма невырожденного полупроводника n -типа (*б*)

сравнения на рис 1.3.2, б представлена та же информация для слаболегированного невырожденного полупроводника n -типа.

Хотя часть примесных уровней в примесной зоне лежит ниже уровня Ферми, в проводимости могут участвовать все электроны введенных доноров, поскольку их волновые функции перекрываются. Поэтому доноры не могут считаться неионизированными, и концентрация основных носителей — электронов — равна концентрации доноров

$$n_0 = N_d. \quad (1.3.12)$$

Концентрация неосновных носителей — дырок — может быть найдена с использованием уравнений (1.3.3а, б), функций распределения (1.3.1а, б) и условия электронейтральности

$$n_0 = p_0 + N_d,$$

если известна зависимость от энергии плотности разрешенных состояний $g(E) = g_c(E) + g_d(E)$ в результирующей разрешенной зоне.

Для кремния с примесью фосфора эта задача решена численно с использованием аппроксимации $g_d(E)$ в виде гауссовой функции [6]. Простой метод, основанный на допущении о симметричности функции $g_d(E)$ относительно оси $E = E_d$, предложен в работе [7]. Хотя закон действующих масс в форме (1.3.9) для вырожденного полупроводника неприменим, формально концентрацию дырок можно записать в виде

$$p_0(N_d) = n_i^2 / N_{def}(N_d). \quad (1.3.13)$$

При этом задача нахождения концентрации неосновных носителей — дырок — состоит в отыскании функции $N_{def}(N_d)$. Поскольку дырочный газ невырожден, остается справедливым соотношение (1.3.8б), которое совместно с (1.3.13) позволяет связать неизвестную концентрацию $N_{def}(N_d)$ с также неизвестным положением уровня Ферми:

$$N_{def} = (n_i^2 / N_v) e^{(F - E_v)/kT}. \quad (1.3.14)$$

Отсюда с учетом (1.3.5) получим

$$N_{def} = N_c e^{(F - E_c)/kT}. \quad (1.3.15)$$

При $N_d \gg N_c$ вырожденная примесная зона содержит в единице объема $2N_d$ уровней, на которых находятся N_d электронов. Если, как предполагается, функция $g_d(E)$ симметрична относительно оси $E = E_d$, то при $T = 0$ заполнены все примесные уровни с энергиями $E \leq E_d$. Это означает, что энергия Ферми в пределе $N_d \gg N_c$ близка к E_d (см. рис. 1.3.2, а):



$$F \approx E_d. \quad (I.3.16)$$

Справедливость (I.3.16) подтверждается численными расчетами в работе [6] не только при $T = 0$, но и при $T = 300$ К. Учитывая, что $E_c - E_v = E_g$, в соответствии с (I.3.14) получаем

$$N_{def} \xrightarrow{N_d \gg N_c} N_{dm}, \quad (I.3.17)$$

где
$$N_{dm} = N_c e^{-\Delta E_d / kT}. \quad (I.3.18)$$

С другой стороны, при низкой степени легирования полупроводник невырожден и

$$N_{def} \xrightarrow{N_d \ll N_c} N_d. \quad (I.3.19)$$

Соотношения (I.3.17) и (I.3.19) определяют асимптотическое поведение искомой функции $N_{def}(N_d)$ (рис. I.3.3). В качестве асимптотически верного приближения можно, например, положить

$$N_{def} \approx N_d / (1 + N_d / N_{dm}), \quad (I.3.20a)$$

или

$$N_{def} \approx N_{dm} \cdot \text{th}(1 + N_d / N_{dm}). \quad (I.3.20б)$$

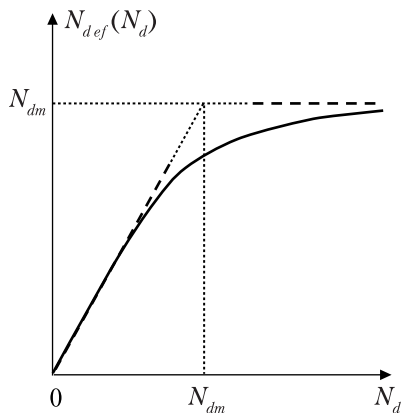


Рис. I.3.3. Характер функции $N_{def}(N_d)$ (сплошная линия) и ее асимптоты (I.3.19) и (I.3.20) (штриховые линии)

По аналогии с (I.3.12), (I.3.13), (I.3.18) и (I.3.20a) для вырожденного полупроводника p^+ -типа получаем

$$p_0 = N_a; \quad (I.3.21)$$

$$n_0(N_a) = n_i^2 / N_{aef}(N_a); \quad (1.3.22)$$

$$N_{am} = N_v e^{-\Delta E_a / kT}, \quad (1.3.23)$$

где
$$N_{aef} \approx N_a / (1 + \frac{N_a}{N_{am}}). \quad (1.3.24)$$

Описанный метод определения концентраций неосновных носителей в вырожденном полупроводнике косвенно подтвержден экспериментом [7].

Значения параметров, определяющих концентрацию дырок в n^+ -Si, для наиболее распространенных донорных примесей приведены в табл. 1.3.1.

Таблица 1.3.1

Значения параметров, определяющих концентрацию дырок в n^+ -Si ($T = 300$ К)

$N_d, \text{см}^{-3}$	Донорная примесь	Энергия активации $\Delta E_d, \text{эВ}$	$N_{dm}, \text{см}^{-3}$
$2,2 \cdot 10^{19}$	Фосфор (P)	0,043	$4,3 \cdot 10^{18}$
	Мышьяк (As)	0,049	$3,4 \cdot 10^{18}$

Отметим, что в вырожденном полупроводнике вследствие перекрытия примесной зоны с зоной проводимости (для полупроводника n -типа) или валентной зоной (для полупроводника p -типа) эффективная ширина запрещенной зоны E_{gef} уменьшается (рис. 1.3.2).

1.3.2. Неравновесное состояние

Под влиянием внешних воздействий функции распределения электронов и дырок отклоняются от равновесных. Эти отклонения выражаются в изменении средней энергии носителей заряда, а также в ненулевом среднем значении импульса (точнее, квазиимпульса). После снятия внешнего воздействия вид функций распределения возвращается к равновесному вследствие столкновений частиц между собой, с дефектами кристаллической решетки, а также межзонных переходов.

Для электронов проводимости и дырок в валентной зоне равновесные значения средней энергии и импульса восстанавливаются соответственно за *время релаксации импульса* τ_p и *время релаксации энергии* τ_E . Это характерное время имеет весьма малые значения $\sim (10^{-13} - 10^{-11})$ с.

Межзонные переходы восстанавливают равновесные значения *концентрации* электронов и дырок за *время жизни* неравно-



весных носителей τ . Время жизни неравновесных носителей в полупроводниках составляет $(10^{-9} - 10^{-5})$ с, т.е. значительно превышает время релаксации импульса и энергии.

Вследствие этого при не слишком сильных внешних воздействиях неравновесные функции распределения похожи по виду на равновесные (I.3.1а, б) или (I.3.2а, б), но сдвинуты по энергиям, что соответствует замене энергетического уровня Ферми F квазиуровнями Ферми для электронов F_n и дырок F_p (см. рис. I.3.1)

$$f_n(E) = \frac{1}{1 + e^{(E - F_n)/kT}}; \quad (I.3.25a)$$

$$f_p(E) = \frac{1}{1 + e^{(F_p - E)/kT}}. \quad (I.3.25б)$$

Квазиуровень F_n характеризует заполнение электронами зоны проводимости, а квазиуровень F_p — заполнение дырками валентной зоны. Отсутствие вырождения электронного и дырочного газов соответствует теперь условиям: $E_c - F_n > 3kT$, $F_p - E_v > 3kT$. При этом функции распределения определяются статистикой Максвелла—Больцмана. По аналогии с равновесными функциями (I.3.2а, б)

$$f_n(E) = e^{(F_n - E)/kT}; \quad (I.3.26a)$$

$$f_{p0}(E) = e^{(E - F)/kT}. \quad (I.3.26б)$$

Положения квазиуровней Ферми определяют значения неравновесных концентраций электронов и дырок. В отсутствие вырождения ($f_n \ll 1$, $f_p \ll 1$) справедливы соотношения, соответствующие соотношениям (I.3.8а, б) и (I.3.9) для равновесного состояния

$$n = n_i e^{(F_n - F_i)/kT} = N_c e^{(F_n - E_c)/kT}; \quad (I.3.27a)$$

$$p = n_i e^{(F_i - F_p)/kT} = N_v e^{(E_v - F_p)/kT}. \quad (I.3.27б)$$

$$np = n_i^2 e^{(F_n - F_p)/kT}. \quad (I.3.28)$$

Избыточными концентрациями электронов и дырок называются их превышения над равновесными

$$\Delta n = n - n_0; \quad (I.3.29a)$$

$$\Delta p = p - p_0. \quad (I.3.29б)$$

Локальная электрическая нейтральность полупроводника соответствует выполнению равенства $p + N_d^+ = n + N_a^-$. Разу-



меется, однородно легированный полупроводник в состоянии равновесия электрически нейтрален: $p_0 + N_d^+ = n_0 + N_a^-$. Вычитая второе равенство из первого, получаем условие локальной нейтральности в виде

$$\Delta n = \Delta p. \quad (1.3.30)$$

1.3.3. Уровень инжекции

Отношение δ концентрации неравновесных неосновных носителей к равновесной концентрации основных называется *уровнем инжекции*

$$\delta = \begin{cases} p / n_0, & n_0 > p_0 \quad (n\text{-тип}), \\ n / p_0, & p_0 > n_0 \quad (p\text{-тип}). \end{cases} \quad (1.3.31)$$

В условиях *электронейтральности* при низком уровне инжекции ($\delta \ll 1$) в однородном полупроводнике концентрация основных носителей близка к равновесной:

$$n \approx n_0 \quad \text{— для полупроводника } n\text{-типа}; \quad (1.3.32a)$$

$$p \approx p_0 \quad \text{— для полупроводника } p\text{-типа}. \quad (1.3.32b)$$

При *высоком уровне инжекции* ($\delta \gg 1$) практически одинаковы концентрации основных, неосновных и избыточных носителей

$$n \approx p \approx \Delta n = \Delta p. \quad (1.3.33)$$

1.4. Кинетика носителей заряда

Направленный перенос электронов и дырок (электрический ток) обусловлен градиентом скалярных величин, влияющих на их поведение, — концентраций носителей, электрического потенциала, температуры. Градиенты концентраций электронов и дырок являются причинами *диффузионных токов*

$$\vec{j}_{n \text{ dif}} = e D_n \nabla n; \quad (1.4.1a)$$

$$\vec{j}_{p \text{ dif}} = -e D_p \nabla p, \quad (1.4.16)$$

градиент потенциала $\nabla \phi = -E$ — причиной *дрейфовых токов*

$$\vec{j}_{n \text{ dr}} = -en\mu_n \nabla \phi = en\mu_n \vec{E}; \quad (1.4.2a)$$

$$\vec{j}_{p \text{ dr}} = -ep\mu_p \nabla \phi = ep\mu_p \vec{E}, \quad (1.4.26)$$



градиент температуры — причиной *термотоков* (в дальнейшем они не учитываются).

Таким образом, плотность тока является суммой плотностей электронного и дырочного токов

$$\vec{j} = \vec{j}_n + \vec{j}_p, \quad (I.4.3)$$

каждая из которых есть сумма диффузионной и дрейфовой составляющих

$$\vec{j}_{n,p} = \vec{j}_{n,p\,dr} + \vec{j}_{n,p\,dif}. \quad (I.4.4)$$

Независимо от механизма тока, плотности токов определяют-ся концентрациями и средними скоростями движения частиц

$$\vec{j}_n = en\vec{v}_n; \quad (I.4.5a)$$

$$\vec{j}_p = ep\vec{v}_p. \quad (I.4.5b)$$

Уравнения для токов (I.4.1), (I.4.2) содержат *кинетические коэффициенты* — *коэффициенты диффузии* $D_{n,p}$ и *подвижности* $\mu_{n,p}$ электронов и дырок. Сами уравнения для токов получены путем интегрирования *кинетических уравнений* Больцмана [1], которые отражают непрерывность потоков функций распределения электронов и дырок.

При малых внешних возмущениях подвижности определяются соотношениями [1]

$$\mu_{n,p} = e\tau_{p\,n,p} / m_{n,p}, \quad (I.4.6)$$

где $\tau_{p\,n,p}$ — времена релаксации импульсов электронов и дырок; $m_{n,p}$ — их эффективные массы.

При $\vec{j}_{n,p\,dif} = 0$ из уравнений (I.4.2) и (I.4.5) следует, что подвижности суть коэффициенты пропорциональности между дрейфовой скоростью и напряженностью электрического поля

$$v_{n,p} = \mu_{n,p} E. \quad (I.4.7)$$

Основным механизмом релаксации импульса носителей, определяющим значения $\tau_{p\,n,p}$ и $\mu_{n,p}$ в (I.4.6), является рассеяние на тепловых колебаниях решетки (акустических фононах) и на ионах примеси. Для данного полупроводника подвижность μ_L , обусловленная рассеянием на фононах, зависит только от температуры

$$\mu_L(T) = \mu_L(T_0)(T / T_0)^{-3/2}, \quad (I.4.8a)$$

где $T_0 = 300$ К.



Подвижность μ_I , обусловленная рассеянием на ионах примеси, зависит от концентрации примесных ионов (а при их полной ионизации — от концентрации примеси N) и температуры. Теоретические формулы, определяющие эту зависимость, приведены, например, в работе [8]. Температурная зависимость подвижности μ_I описывается соотношением

$$\mu_I(T) = \mu_I(T_0)(T/T_0)^{3/2}. \quad (1.4.86)$$

Для каждой фиксированной энергии частиц E результирующая подвижность определяется сложением обратных величин подвижностей для каждого механизма рассеяния

$$\mu(E) = 1 / [\mu_L^{-1}(E) + \mu_I^{-1}(E)],$$

однако для интегральной подвижности (усредненной по всем энергиям) это правило не выполняется. Практически используются экспериментальные данные [8] либо их аппроксимации. Для кремния при $T = 300$ К хорошие результаты дает аппроксимация

$$\mu(N_\Sigma) = \begin{cases} \mu_1, & N_\Sigma \leq N_1, \\ \mu_{n1}[1 - \lambda \ln(N_\Sigma / N_1)], & N_1 < N_\Sigma < N_2, \\ \mu_{n2}, & N_\Sigma \geq N_2, \end{cases} \quad (1.4.9)$$

где N_Σ — суммарная концентрация примеси, а значения параметров для основных носителей приведены в табл. 1.4.1. Для неосновных носителей эффективные значения концентраций примеси $N_{1,2}$ несколько выше, так как неосновные носители рассеиваются также на основных, движущихся навстречу.

Таблица 1.4.1

**Значения параметров аппроксимации (1.4.9)
для основных носителей заряда в кремнии ($T = 300$ К)**

Носители	$N_1, \text{см}^{-3}$	$N_2, \text{см}^{-3}$	$\mu_1, \text{см}^2/\text{В}\cdot\text{с}$	$\mu_2, \text{см}^2/\text{В}\cdot\text{с}$	λ
Электроны	$3 \cdot 10^{15}$	$1 \cdot 10^{19}$	1300	85	0,115
Дырки	$2 \cdot 10^{15}$	$1 \cdot 10^{19}$	500	50	0,145

Линейная зависимость $v_{n,p}(E)$ (1.4.7) нарушается в сильных полях, когда на длине свободного пробега носители успевают набрать энергию, достаточную для возбуждения оптических фононов. В однородных образцах дрейфовая скорость носителей насыщается с ростом поля (рис. 1.4.1), стремясь в большинстве полупроводниковых материалов к величине $v_s \approx 10^7$ см/с. Наиболее простые аналитические



результаты получаются при аппроксимации зависимости $v(E)$ кусочно-линейной функцией:

$$v(E) = \begin{cases} \mu E, & E \leq E_s, \\ v_s = \mu E_s, & E > E_s \end{cases} \quad (\text{I.4.10})$$

(штриховая линия на рис. I.4.1), где μ — подвижность носителей в слабом поле.

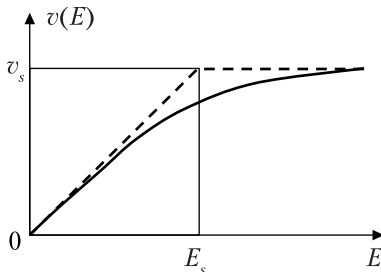


Рис. I.4.1. Зависимость дрейфовой скорости носителей заряда в полупроводнике от напряженности электрического поля и ее аппроксимация (I.4.10)

Предельные скорости дрейфа электронов и дырок в кремнии составляют

$$v_{sn} \approx 1,1 \cdot 10^7 (T_0 / T)^{0,87} \text{ см/с}; \quad v_{sp} \approx 0,95 \cdot 10^7 (T_0 / T)^{0,52} \text{ см/с},$$

где $T_0 = 300 \text{ К}$.

Эти значения практически не зависят от концентрации примеси, так как скорость ограничивается при достижении носителями энергии оптического фонона. Напряженность поля E_s , соответствующая достижению предельной скорости, в зависимости от степени легирования составляет

$$E_s = \frac{v_s}{\mu} \approx \begin{cases} 15\text{--}25 \text{ кВ/см} & \text{— для } n\text{-Si}; \\ 30\text{--}50 \text{ кВ/см} & \text{— для } p\text{-Si}. \end{cases}$$

Для невырожденных носителей заряда коэффициент диффузии связан с подвижностью соотношением Эйнштейна

$$D_{n,p} = \Phi_T \mu_{n,p}, \quad (\text{I.4.11a})$$

где

$$\Phi_T = kT / e \quad (\text{I.4.11б})$$

— температурный потенциал.

Для рассматриваемых нами приборов нет необходимости учитывать диффузионные токи вырожденных носителей.

Понятие квазиуровней Ферми позволяет в невырожденном полупроводнике выразить полную плотность тока через градиент единой скалярной величины — квазиуровня Ферми. Согласно (1.3.27а)

$$\nabla n = \nabla [n_i e^{(F_n - F_i)/kT}] = n_i e^{(F_n - F_i)/kT} (\nabla F_n - \nabla F_i) / kT = n (\nabla F_n - \nabla F_i) / kT.$$

Таким образом,

$$\vec{j}_n = n\mu_n \nabla F_i + eD_n n (\nabla F_n - \nabla F_i) / kT.$$

Аналогичное выражение может быть получено для дырочного тока

$$\vec{j}_p = p\mu_p \nabla F_i + eD_p p (\nabla F_p - \nabla F_i) / kT.$$

С учетом (1.4.11а)

$$\begin{cases} \vec{j}_n = n\mu_n \nabla F_n, \\ \vec{j}_p = p\mu_p \nabla F_p. \end{cases} \quad \begin{matrix} (1.4.12a) \\ (1.4.12б) \end{matrix}$$

Энергетические диаграммы полупроводника для случаев чистой диффузии электронов ($\vec{E} = \nabla F_i = 0$) и чистого дрейфа ($\nabla n \sim \nabla F_n - \nabla F_i = 0$) показаны на рис. 1.4.2.

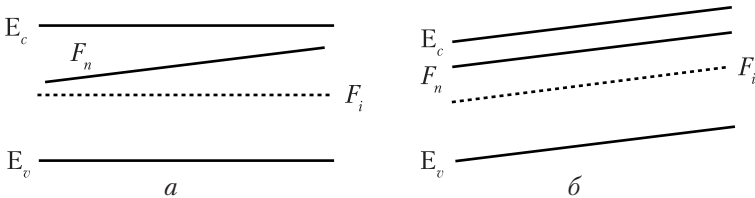


Рис. 1.4.2. Энергетические диаграммы полупроводника для чистой диффузии электронов (а) и чистого дрейфа (б)

1.5. Рекомбинация и генерация носителей заряда

1.5.1. Механизмы рекомбинации—генерации

Основными механизмами рекомбинации носителей заряда в полупроводниках являются: 1) прямая межзонная; 2) ловушечная; 3) ударная (рекомбинация Оже); 4) поверхностная. Каждому механизму рекомбинации соответствует противоположный механизм генерации носителей. Рекомбинация и генерация носителей происходят парами: при каждом акте исчезает или возникает электронно-дырочная

пара, что на языке реальных (а не квази-) частиц можно интерпретировать как переход электрона из валентной зоны в зону проводимости или наоборот. Единственное исключение составляет захват носителей на мелкие ловушечные центры в запрещенной зоне и обратный переход в ту же зону. Эти процессы в дальнейшем не рассматриваются.

Процессы рекомбинации и генерации электронно-дырочных пар характеризуются *скоростью рекомбинации* r и *скоростью генерации* g , которые определяются как число электронно-дырочных пар, рекомбинирующих и генерированных в единице объема в единицу времени. Скорость рекомбинации носителей зависит от их концентрации ($r = 0$, если $n = 0$ или $p = 0$). Скорость генерации зависит от средней энергии носителей, т.е. от температуры. От концентрации носителей она может зависеть только в случае, когда электронный или дырочный газ вырожден и заметная часть разрешенных состояний в валентной зоне или в зоне проводимости заполнена. В дальнейшем будем считать, что *скорость генерации зависит только от температуры (скорость термогенерации)*. При этом в состоянии равновесия

$$r = r_0 = g. \quad (I.5.1)$$

Избыточной скоростью рекомбинации называется разность $r - g$. Другой характеристикой процессов рекомбинации и генерации может служить *время жизни* электронов τ_n и дырок τ_p , определяемые соотношением

$$r - g = \Delta n / \tau_n = \Delta p / \tau_p. \quad (I.5.2)$$

Понятие времени жизни целесообразно использовать в тех случаях, когда оно не зависит от избыточной концентрации носителей.

I.5.2. Прямая межзонная рекомбинация

Прямая межзонная рекомбинация называется *излучательной*, поскольку энергия, выделяемая при рекомбинации каждой пары, излучается в виде фотона. Электроны и дырки имеют энергии, соответственно близкие к E_c и E_v , поэтому энергия излучаемых фотонов определяется шириной запрещенной зоны ($h\nu \approx E_c - E_v = E_g$), а длина волны $\lambda = hc / E_g$, где h — постоянная Планка; c — скорость света. Скорость излучательной рекомбинации пропорциональна произведению концентраций электронов и дырок:

$$r = \gamma np.$$



В состоянии равновесия с учетом (1.5.1) и (1.3.9)

$$r_0 = \gamma n_0 p_0 = g; \quad r - g = \gamma(np - n_0 p_0) = \gamma(np - n_i^2).$$

Сравнение этого результата с (1.5.2) показывает, что в однородном электронейтральном полупроводнике для излучательной рекомбинации время жизни не зависит от избыточной концентрации носителей при низком уровне инжекции

$$\tau = 1 / \gamma(p_0 + n_0) = 1 / \gamma N. \quad (1.5.3)$$

1.5.3. Ловушечная рекомбинация

Ловушечная рекомбинация (*рекомбинация Шокли—Рида—Холла*) осуществляется путем последовательного захвата электронов и дырок глубокими ловушечными уровнями, которые создаются в глубине запрещенной зоны фоновыми примесями. Механизм ловушечной рекомбинации исследован Шокли, Ридом и Холлом в предположении, что в запрещенной зоне имеется единственный ловушечный уровень с энергией E_t и получено соотношение, определяющее избыточную скорость рекомбинации:

$$r - g = \frac{C_n C_p N_t (np - n_i p_1)}{C_n (n + n_1) + C_p (p + p_1)} = \frac{np - n_i^2}{\tau_{n0} (n + n_1) + \tau_{p0} (p + p_1)}, \quad (1.5.4)$$

где N_t — концентрация ловушек; $C_{n,p}$ — вероятности захвата ловушками электронов и дырок; $\tau_{n0,p0} = 1 / N_t C_{p,n}$, $n_1 = n_i e^{(E_t - E_i)/kT}$, $p_1 = n_i e^{(E_i - E_t)/kT}$.

Сравнение соотношений (1.5.4) и (1.5.2) с учетом (1.3.8а, б) показывает, что в однородном электронейтральном полупроводнике время жизни *неосновных* носителей при ловушечной рекомбинации не зависит от избыточной концентрации носителей, если

$$n_0 + p_0 \gg n_1 + p_1 \quad (1.5.5)$$

и составляет

$$\tau = \tau_{n0} \quad \text{— для полупроводника } n\text{-типа}; \quad (1.5.6a)$$

$$\tau = \tau_{p0} \quad \text{— для полупроводника } p\text{-типа} \quad (1.5.6b)$$

Таким образом, при условии (1.5.5) время жизни *неосновных* носителей не зависит явно от степени легирования полупроводника. Практически повышение степени легирования ведет к увеличению концентрации ловушек N_t , в результате чего время жизни уменьшается.



1.5.4. Ударная рекомбинация

При ударной рекомбинации (*рекомбинации Оже*) энергия, выделяющаяся при рекомбинации электронно-дырочной пары, передается третьему носителю — электрону или дырке, который после этого передает ее кристаллической решетке (фононам) за малое время релаксации энергии. Скорость ударной рекомбинации пропорциональна концентрации неосновных носителей и квадрату концентрации основных носителей. Используя соотношения (1.5.1) и (1.5.2), нетрудно показать, что для Оже-рекомбинации время жизни в однородном электронейтральном полупроводнике не зависит от избыточной концентрации носителей при низком уровне инжекции

$$\tau = 1 / A(n_0^2 + p_0^2) = 1 / AN^2, \quad (1.5.7)$$

где A — константа Оже-рекомбинации.

1.5.5. Поверхностная рекомбинация

На поверхности полупроводника обрывается его кристаллическая решетка. Кроме того, имеются различные *поверхностные дефекты* физической, химической или механической природы. Все это приводит к тому, что в поверхностном слое полупроводника запрещенная зона содержит множество разрешенных состояний, способствующих повышению интенсивности рекомбинационно-генерационных процессов. Рекомбинация избыточных носителей заряда на поверхности вызывает приток к ней электронов и дырок. Таким образом, в направлении, нормальном к поверхности, протекают равные по абсолютной величине электронный и дырочный токи. В первом приближении плотность этих токов пропорциональна избыточной концентрации носителей заряда на поверхности Δn_s и Δp_s

$$j_{sn} = -e\Delta n_s s; \quad (1.5.8a)$$

$$j_{sp} = e\Delta p_s s, \quad (1.5.8б)$$

где s — *скорость поверхностной рекомбинации*, характеризующая интенсивность механизма поверхностной рекомбинации.

Для идеальной поверхности, эквивалентной любой воображаемой поверхности в объеме полупроводника, $s = 0$. Для поверхности идеального металлического контакта $s = \infty$. Бесконечное значение скорости поверхностной рекомби-

нации означает, что на поверхности полупроводника всегда $\Delta n_s = \Delta p_s = 0$, т.е. поверхностные концентрации электронов и дырок всегда остаются равновесными ($n_s = n_0$, $p_s = p_0$). Такие идеальные контакты называются *омическими*. На поверхности реальных металлических контактов к полупроводнику скорость поверхностной рекомбинации остается конечной, но весьма высокой. В моделях электронных приборов ее обычно полагают бесконечной.

В полупроводниковых приборах и ИС поверхность полупроводника никогда не оставляют открытой. Чтобы избежать деградации характеристик прибора, используют пассивирующие покрытия. В кремниевых приборах и ИС хорошим пассивирующим покрытием является диоксид кремния SiO_2 . В зависимости от его качества скорость поверхностной рекомбинации на пассивированной поверхности может составлять от 10 до 10^3 см/с и выше.

1.5.6. Превалирующие механизмы рекомбинации

Все рассмотренные механизмы рекомбинации действуют одновременно, причем результирующие скорости рекомбинации и генерации суммируются

$$r_\Sigma = \Sigma r_i; \quad g_\Sigma = \Sigma g_i,$$

а результирующее время жизни составляет $\tau_\Sigma = 1 / \Sigma \tau_i^{-1}$.

Степень проявления различных механизмов рекомбинации зависит от особенностей зонной структуры полупроводника и степени его легирования.

В моноатомных полупроводниках (Si, Ge) излучательный механизм рекомбинации проявляется слабо, так как в них дно зоны проводимости и потолок валентной зоны соответствуют различным значениям квазиимпульсов электронов и дырок. Поэтому прямые (без изменения квазиимпульса) межзонные переходы возможны только для частиц, имеющих значительную кинетическую энергию, и маловероятны.

В невырожденных моноатомных полупроводниках ($N < 10^{19}$ см⁻³) превалирующим является ловушечный механизм рекомбинации Шокли—Рида—Холла. Его интенсивность зависит от чистоты материала (концентрации глубоких ловушечных уровней N_t). При степени легирования $N = (10^{15} - 10^{17})$ см⁻³ время жизни в кремнии составляет обычно от 0,1 до 10 мкс, а при $N \sim 10^{14}$ см⁻³ может достигать 1 мс. Для некоторых областей применения желательно снижение време-



ни жизни. Это достигается путем специального легирования полупроводника примесями, создающими глубокие ловушечные уровни. Для кремния такой примесью является Au.

В вырожденном полупроводнике ($N > 10^{19} \text{ см}^{-3}$) наиболее вероятен механизм Оже-рекомбинации, для которого время жизни $\tau \sim 1/N^2$. При степени легирования $N > 10^{20} \text{ см}^{-3}$ время жизни составляет от 1 до 10 нс.

Прямые межзонные переходы характерны для полярных полупроводников типа A^3B^5 (GaAs, AlAs, InP и др) и их соединений. В этих полупроводниках энергетический минимум в зоне проводимости ($E = E_c$) и максимум в валентной зоне ($E = E_v$) расположены в центре зоны Бриллюэна ($\vec{p}_{n0} = \vec{p}_{p0} = 0$), и основным механизмом является излучательная рекомбинация ($\tau \sim 1/N$), сопровождающаяся излучением фотонов с энергией E_g . Такие полупроводники применяются для создания источников монохроматического излучения (в том числе лазеров) ближнего инфракрасного и оптического диапазона.

Поверхностная рекомбинация проявляется во всех полупроводниковых приборах биполярного типа, основанных на использовании как основных, так и неосновных носителей.

I.6. Общий подход к анализу полупроводниковых приборов

Полупроводниковые приборы являются весьма сложными объектами. Их анализ включает исследование статических характеристик прибора, его работу на высокочастотном гармоническом сигнале или в импульсном режиме, а также синтез моделей, позволяющих описывать поведение прибора в составе электронных устройств обработки информации.

Для определения токов в выводах прибора необходимо найти потоки электронов и дырок через некоторые поверхности, отделяющие электроды от внутренних областей прибора. Плотности токов определяются пространственными распределениями концентраций подвижных носителей и напряженности электрического поля (см. разд. I.4). Уравнения для токов содержат кинетические коэффициенты (подвижности и коэффициенты диффузии частиц), которые являются локальными функциями средней энергии носителей. Таким образом, для определения плотностей тока необходимо найти пространственные распределения концентраций носителей заряда и их средних энергий, а также напряженности электрического поля.



При известном распределении концентраций носителей (т.е. плотности объемного заряда) напряженность электрического поля может быть найдена из уравнения Пуассона

$$\nabla \vec{E} = \rho / \epsilon \epsilon_0 = e(\Delta p - \Delta n) / \epsilon \epsilon_0. \quad (I.6.1)$$

Распределения концентраций носителей и их средних энергий могут быть найдены из *уравнений непрерывности* потоков частиц и потоков их энергии. Для наших задач с достаточной точностью средние энергии носителей могут считаться локальными функциями напряженности электрического поля. Таким образом, задача сводится к отысканию пространственных распределений концентраций частиц $n(\vec{r})$, $p(\vec{r})$ и напряженности электрического поля $\vec{E}(\vec{r})$ путем совместного решения уравнений непрерывности потоков электронов и дырок, а также уравнения Пуассона (I.6.1) при заданных граничных условиях на контактах и поверхностях прибора.

Уравнения непрерывности будут рассмотрены в разд. I.7, однако следует сразу сказать, что решение самосогласованной задачи отыскания функций $n(\vec{r})$, $p(\vec{r})$ и $\vec{E}(\vec{r})$, как правило, может быть получено только численными методами. Для получения достаточно общих результатов, способствующих физическому пониманию свойств приборов и развитию инженерной интуиции, необходимы дополнительные упрощения.

Таким упрощением является разбиение прибора на области пространственного заряда (ОПЗ) и квазинейтральные области. В ОПЗ используется приближение полного обеднения области подвижными носителями заряда, что позволяет свести задачу отыскания поля к электростатической. В квазинейтральных областях существенного упрощения можно достигнуть, полагая избыточные концентрации электронов и дырок равными, что позволяет заменить уравнения непрерывности потоков электронов и дырок единым и более простым *биполярным уравнением непрерывности*.

I.7. Уравнения непрерывности

В физике полупроводников и полупроводниковых приборов используется ряд *уравнений непрерывности*, которые констатируют непрерывность потоков различных физических величин, характеризующих ансамбли электронов и дырок.



В рамках квазиклассического приближения исходными являются *уравнения непрерывности потоков функций распределения* $f_{n,p}(\vec{r}, \vec{p}_{n,p}, t)$ — кинетические уравнения Больцмана [1]. Аргументами функций распределения в общем случае являются координата $\vec{r}$, импульсы (или квазиимпульсы) частиц $\vec{p}_{n,p}$ и время t , а в условиях равновесия ($f_{n,p} = f_{n0,p0}$, (I.3.1)) — энергия частиц $E_{n,p}$.

Интегрирование кинетических уравнений Больцмана по всем импульсам дает *уравнения непрерывности потоков частиц* — электронов и дырок, которые часто называют просто уравнениями непрерывности. Они и являются предметом изучения данной главы.

Интегрирование по всем импульсам кинетических уравнений Больцмана, умноженных на импульс, дает *уравнения непрерывности потоков среднего импульса* частиц, которые в полупроводниках с небольшим упрощением соответствуют уравнениям для токов.

Интегрирование по всем импульсам кинетических уравнений Больцмана, умноженных на квадрат импульса, дает *уравнения непрерывности потоков средней энергии* частиц. Эти уравнения необходимы при наличии сильных внешних электрических полей, когда средние энергии частиц $\bar{E}_{n,p}$ (или характеризующие их температуры $T_{n,p}$) заметно отличаются от равновесного значения (или температуры решетки T).

Если внешние поля настолько деформируют функции распределения, что им не могут быть приписаны определенные значения температуры, ряд уравнений непрерывности может быть продолжен.

Для наших основных целей достаточно использовать уравнения непрерывности потоков частиц. Эти уравнения имеют вид

$$\frac{\partial p}{\partial t} = g_p - r_p - \frac{1}{e} \nabla \vec{j}_p; \quad (I.7.1a)$$

$$\frac{\partial n}{\partial t} = g_n - r_n + \frac{1}{e} \nabla \vec{j}_n. \quad (I.7.1б)$$

Левые части уравнений (I.7.1a, б) имеют смысл суммарной скорости увеличения концентраций частиц, первые слагаемые правых частей — скорости увеличения концентраций частиц за счет процессов их генерации и рекомбинации, а последние слагаемые — скорости увеличения концентраций частиц за счет дивергенции их потоков (преимущественного втекания частиц в выбранный элементарный объем).



Подстановка в уравнения (1.7.1а, б) уравнений для токов (1.4.1)–(1.4.4) и уравнений генерации-рекомбинации (1.5.2) дает

$$\frac{\partial p}{\partial t} = -\frac{\Delta p}{\tau_p} - \mu_p (\vec{E}, \nabla p) - \mu_p p \nabla \vec{E} + D_p \nabla^2 p; \quad (1.7.2a)$$

$$\frac{\partial n}{\partial t} = -\frac{\Delta n}{\tau_n} + \mu_n (\vec{E}, \nabla n) + \mu_n n \nabla \vec{E} + D_n \nabla^2 n. \quad (1.7.2б)$$

В дальнейшем уравнения непрерывности будут применяться лишь для анализа *квазинейтральных* областей приборов, где условие электрической нейтральности (1.3.30) выполнено приближенно:

$$\Delta p = p - p_0 \approx \Delta n = n - n_0. \quad (1.7.3)$$

Квазинейтральность обеспечивается кулоновским притяжением избыточных электронов и избыточных дырок. При ее нарушении возникает электрическое поле, напряженность которого $\vec{E}$ определяется уравнением Пуассона (1.6.1):

$$\nabla \vec{E} = \rho / \epsilon \epsilon_0 = e(\Delta p - \Delta n) / \epsilon \epsilon_0.$$

Это поле направлено таким образом, чтобы восстановить локальную нейтральность полупроводника. В полупроводниковых приборах условие квазинейтральности (1.7.3) выполняется с высокой точностью, если *внешнее* электрическое поле, действующее на электроны и дырки в противоположных направлениях, не слишком велико. Электронейтральность может нарушаться и в равновесном состоянии при резком изменении концентрации примеси (см. разд. II.1). Эти области мы из рассмотрения исключим.

Использование уравнения квазинейтральности (1.7.3) позволяет существенно упростить задачу отыскания функций $n(x, t)$ и $p(x, t)$, так как вместо двух уравнений в частных производных (1.7.2а, б) достаточно иметь одно, содержащее в качестве неизвестного Δp , либо примерно равное ему Δn .

Разумеется, нельзя просто положить в (1.7.2а, б) $\Delta p = \Delta n$, так как при этом мы получим два *различных* уравнения с двумя *разными* решениями для одинаковых неизвестных. Оба эти уравнения некорректны, так как при $\Delta p = \Delta n$ $\nabla \vec{E} = 0$ и исчезает фактор, поддерживающий квазинейтральность, — члены $-\mu_p p \nabla \vec{E}$ в (1.7.2а) и $\mu_n n \nabla \vec{E}$ в (1.7.2б), содержащие $\nabla \vec{E}$.

Обойти эту трудность можно, исключая из уравнений (1.7.2а) и (1.7.2б) слагаемые, содержащие $\nabla \vec{E}$ (путем сложе-



ния этих уравнений, умноженных соответственно на $\mu_n n$ и $\mu_p p$, а затем деления результата на $\mu_n n + \mu_p p$). В остальных слагаемых с высокой точностью можно полагать

$$\Delta n = \Delta p, \quad \partial p / \partial t = \partial n / \partial t, \quad \nabla p = \nabla n \quad \text{и} \quad \nabla^2 p = \nabla^2 n.$$

Полученное в результате этой процедуры *биполярное уравнение непрерывности* можно записать в виде

$$\frac{\partial p}{\partial t} = -\frac{\Delta p}{\tau} - \mu E \nabla p + D \nabla^2 p \quad (I.7.4a)$$

либо в виде

$$\frac{\partial n}{\partial t} = -\frac{\Delta n}{\tau} - \mu E \nabla n + D \nabla^2 n, \quad (I.7.4б)$$

где

$$D = \frac{p + n}{p / D_n + n / D_p} \quad (I.7.5)$$

— *биполярный коэффициент диффузии*;

$$\mu = \frac{n - p}{p / \mu_n + n / \mu_p} \quad (I.7.6)$$

— *биполярная подвижность*;

$$\tau = \frac{p + n}{p / \tau_n + n / \tau_p} \quad (I.7.7)$$

— *биполярное время жизни* носителей заряда.

Единственное биполярное уравнение непрерывности в форме (I.7.4a) или (I.7.4б) заменяет два уравнения непрерывности (I.7.2a) и (I.7.2б), а также уравнение Пуассона, что является несомненным упрощением. Теперь трудность состоит в том, что биполярное уравнение непрерывности содержит биполярные коэффициенты D , μ и τ , зависящие от концентраций электронов и дырок. К счастью, в практически важных случаях низкого и высокого уровней инжекции биполярные коэффициенты оказываются постоянными.

При *низком* уровне инжекции в n -области $p \ll n$, и в соответствии с уравнениями (I.7.5)–(I.7.7) биполярные кинетические коэффициенты, а также биполярное время жизни совпадают с соответствующими коэффициентами для неосновных носителей

$$D = D_p; \quad \mu = \mu_p; \quad \tau = \tau_p.$$



При этом уравнение (1.7.4а) имеет вид

$$\frac{\partial p_n}{\partial t} = -\frac{\Delta p_n}{\tau_p} - \mu_p E \nabla p_n + D_p \nabla^2 p_n. \quad (1.7.8а)$$

Для низкого уровня инжекции в p -области $n \ll p$. При этом

$$D = D_n; \quad \mu = -\mu_n; \quad \tau = \tau_n$$

и уравнение (1.7.4б) имеет вид

$$\frac{\partial n_p}{\partial t} = -\frac{\Delta n_p}{\tau_n} + \mu_n E \nabla n_p + D_n \nabla^2 n_p. \quad (1.7.8б)$$

Сравнивая уравнения (1.7.8) и (1.7.2), приходим к следующему заключению: *при низком уровне инжекции биполярное уравнение непрерывности совпадает с уравнением непрерывности для неосновных носителей без члена с $\nabla \vec{E}$.*

Физическая причина определяющей роли неосновных носителей состоит в следующем. Равновесная концентрация неосновных носителей (например, дырок с концентрацией p_{n0} в полупроводнике n -типа) весьма мала (в пределе $p_{n0} \rightarrow 0$). Поэтому под действием внешних факторов (например, электрического поля) перемещаются именно те избыточные неосновные носители, которые имеются в наличии. Равное количество избыточных основных носителей (в нашем примере электроны с концентрацией $\Delta n_n(x) = \Delta p_n(x)$) перемещается под действием кулоновского поля, создаваемого избытком неосновных носителей, в том же направлении. Их перемещение, однако, осуществляется в виде волны на фоне значительного количества равновесных основных носителей n_{n0} . Такое перемещение осуществить значительно легче, чем физическое перемещение определенных частиц. Поэтому избыточные носители перемещаются в направлении внешней силы, действующей на *неосновные* носители. В полупроводнике p -типа перемещение избыточных носителей в тех же условиях осуществляется в противоположном направлении. Рассмотренная ситуация поясняется рис. 1.7.1.

При *высоком* уровне инжекции из уравнений (1.7.5)–(1.7.7) получаем

$$n \approx p; \quad D = 2D_n D_p / (D_n + D_p); \quad \mu = 0; \quad \tau = 2\tau_n \tau_p / (\tau_n + \tau_p),$$

и биполярные уравнения непрерывности, записанные в виде (1.7.4а) и (1.7.4б), совпадают:

$$\frac{\partial p}{\partial t} = -\frac{\Delta p}{\tau} + D \nabla^2 p, \quad (1.7.9а)$$



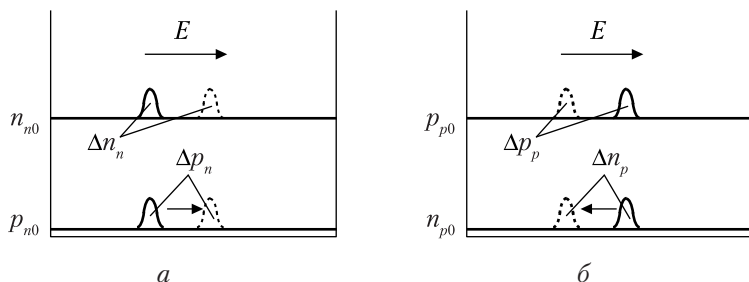


Рис. I.7.1. Перемещение избыточных носителей заряда под действием внешнего электрического поля при низком уровне инжекции в полупроводниках *n*-типа (а) и *p*-типа (б)

или
$$\frac{\partial n}{\partial t} = -\frac{\Delta n}{\tau} + D \nabla^2 n. \quad (I.7.96)$$

Сравнение уравнений (I.7.8) и (I.7.9) позволяет сделать следующий вывод: *при высоком уровне инжекции биполярное уравнение непрерывности совпадает с любым из уравнений непрерывности для электронов и дырок, в которых изъят член с ∇E , а коэффициенты определяются такими соотношениями, как $D = 2D_n D_p / (D_n + D_p)$; $\mu = 0$; $\tau = 2\tau_n \tau_p / (\tau_n + \tau_p)$.*

Литература

1. Бонч-Бруевич, В. Л. Физика полупроводников / В. Л. Бонч-Бруевич, С. Г. Калашников. — М.: Наука, 1977.
2. Киреев, П. С. Физика полупроводников / П. С. Киреев. — М.: Высш. школа, 1969.
3. Шалимова, К. В. Физика полупроводников / К. В. Шалимова. — М.: Энергия, 1985.
4. Киттель, Ч. Введение в физику твердого тела: пер. с англ. / Ч. Киттель. — М.: Физматгиз, 1963.
5. Парменов, Ю. А. Физика полупроводников: учеб. пособие / Ю. А. Парменов. — М.: МИЭТ, 2002.
6. Kleppenger, D. D. Impurity concentration dependent density of states and resulting Fermi level for silicon / D. D. Kleppenger, F. A. Lindholm // Solid State Electron., 1971. — Vol. 14. — № 5. — P. 407.
7. Лапшинов, О. Н. Коэффициент усиления интегрального транзистора: сб. науч. трудов по проблемам микроэлектроники; сер. Полупроводниковые ИС памяти — схемотехника и технология / О. Н. Лапшинов, В. И. Старосельский — М.: МИЭТ, 1979. — С. 123–140.
8. Sze, S. M. Physics of Semiconductor Devices (3-rd edition) / S. M. Sze, K. Ng. Kwok. — Wiley-Interscience, 2007.

Глава II

***p-n* ПЕРЕХОДЫ И ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ**

Переходные области между полупроводниками *p*- и *n*-типа играют важную роль в современной электронике как основа самостоятельного полупроводникового прибора (диода) и для понимания работы других полупроводниковых приборов. Основы теории *p-n* переходов заложены Шокли [1, 2] и развиты Са, Нойсом, Шокли [3], а также Моллом [4].

II.1. Понятие о *p-n* переходе

II.1.1. Определение и классификация *p-n* переходов

Одним из основных понятий полупроводниковой электроники является понятие о *p-n переходе* — области пространственного заряда (ОПЗ), возникающей при резком пространственном изменении типа проводимости полупроводника.

На рис. II.1.1 представлены равновесные распределения зарядов в полупроводнике при плавном и резком изменении типа проводимости. При плавном изменении типа проводимости (рис. II.1.1, *a*) градиент концентрации результирующей примеси $N = N_d - N_a$ мал, соответственно малы и диффузионные токи электронов и дырок. Эти токи компенсируются дрейфовыми токами, которые вызваны электрическим полем, связанным с нарушением условия электрической нейтральности,

$$N_d - N_a + p_0 - n_0 = 0. \quad (\text{II.1.1})$$

Для компенсации диффузионных токов достаточно незначительного нарушения нейтральности, и условие (II.1.1) можно считать приближенно выполненным:

$$p_0 + n_0 \gg N_d - N_a + p_0 - n_0 \approx 0. \quad (\text{II.1.2})$$



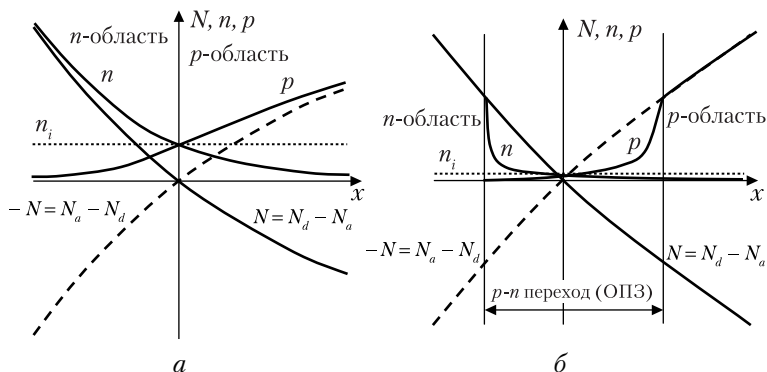


Рис. II.1.1. Распределение примеси и носителей заряда в полупроводнике при плавном (а) и резком (б) изменении типа проводимости

При резком изменении типа проводимости (см. рис. II.1.1, б) диффузионные токи велики, и для их компенсации необходимо существенное нарушение электронейтральности.

Известно, что экранирование малого электрического заряда в полупроводнике подвижными носителями осуществляется в области с размерами порядка дебаевской длины экранирования

$$L_D = \sqrt{\frac{\epsilon \epsilon_0 \Phi_T}{e(n_0 + p_0)}}, \quad (\text{II.1.3})$$

где $\Phi_T = kT/e$ — температурный потенциал. При этом электрическая нейтральность существенно нарушается, если на дебаевской длине экранирования изменение результирующей концентрации примеси велико. Таким образом, нейтральность существенно нарушается при условии

$$\left| \frac{dN}{dx} \right| \gg \frac{n_0 + p_0}{L_D}. \quad (\text{II.1.4})$$

В состоянии термодинамического равновесия при отсутствии вырождения справедлив закон действующих масс

$$n_0 p_0 = n_i^2. \quad (\text{II.1.5})$$

При условии (II.1.5) правая часть (II.1.4) достигает минимума при $n_0 = p_0 = n_i$. Поэтому условие существования p - n перехода (условие существенного нарушения нейтральности) имеет вид

$$\left| \frac{dN}{dx} \right| \gg \frac{2n_i}{L_{Di}}, \quad (\text{II.1.6})$$

где $L_{Di} = \sqrt{\epsilon\epsilon_0\Phi_T / e2n_i}$ — дебаевская длина экранирования в собственном полупроводнике.

Переходы, в которых изменение концентрации примеси на границе слоев p - и n -типа может считаться скачкообразным ($|dN/dx| \rightarrow \infty$), называются *ступенчатыми*. В *плавных* p - n переходах градиент концентрации примеси конечен, но удовлетворяет неравенству (II.1.6).

Практически ступенчатыми могут считаться p - n переходы, в которых изменение концентрации примеси происходит на длине, много меньшей ширины перехода. Такие переходы могут быть получены путем сплавления или эпитаксии.

По отношению к концентрации основных носителей в слоях p - и n -типа переходы делятся на *симметричные* и *несимметричные*.

В симметричных переходах функция $N(x)$ обладает центральной симметрией относительно точки $x=0$, где $N=0$. В несимметричных p - n переходах более легированная (и, соответственно, более электропроводная) область называется *эмиттером*, а менее легированная (высокоомная) — *базой*. Отметим сразу, что эмиттер может оказаться *вырожденным*; база практически никогда вырожденной не бывает за исключением специальных случаев, которые будут оговорены.

Особую группу составляют *точечные* p - n переходы, в которых размер эмиттера настолько мал, что переход имеет форму сферического слоя. Точечные переходы могут быть получены на основе контакта металлической иглы с полупроводником при последующем вжигании. Точечные переходы в микроэлектронике практически не используются и в дальнейшем рассматриваться не будут. В противоположность им *плоскостные* переходы ограничены со стороны p - и n -областей плоскостями. Обычно размеры плоскостных p - n переходов в поперечном направлении намного больше ширины перехода, что позволяет использовать при их анализе одномерное приближение (краевые эффекты не учитываются).



В заключение отметим, что формула (II.1.1) предполагает полную ионизацию избыточных примесей (плотность заряда примесных атомов составляет $e(N_d - N_a)$), а формула (II.1.5) — отсутствие вырождения. Случаи, когда эти допущения не выполнены, будут рассматриваться особо.

II.1.2. Структура *p-n* перехода

Наиболее просто поддаются анализу ступенчатые переходы. Структура ступенчатого перехода представлена на рис. II.1.2. Практически всегда концентрации примеси в *p*- и *n*-областях намного превышают собственную концентрацию носителей заряда: $N_{E,B} \gg n_i$.

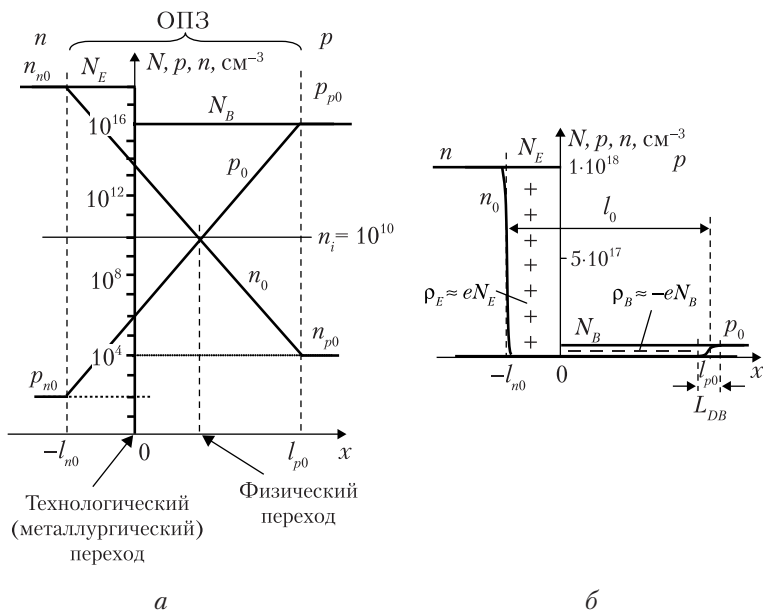


Рис. II.1.2. Распределение примеси и носителей заряда в ступенчатом *p-n* переходе:

a — полулогарифмический масштаб, *б* — линейный масштаб

Для определенности будем полагать, что эмиттером является *n*-область, а базой — *p*-область. В большинстве практических случаев выполняется неравенство

$$N_E \gg N_B, \quad (\text{II.1.7})$$

где N_E, N_B — результирующие концентрации примеси в эмиттере и базе.

На рис. II.1.2 показан кремниевый переход ($n_i \approx 10^{10} \text{ см}^{-3}$ при комнатной температуре $T = 290 \text{ K}$) с концентрациями примеси $N_E = 10^{18} \text{ см}^{-3}$, $N_B = 10^{16} \text{ см}^{-3}$.

В глубине эмиттера и базы концентрации основных носителей заряда практически совпадают с результирующими концентрациями примеси

$$\begin{cases} n_{n0} = N_E, \end{cases} \quad (\text{II.1.8a})$$

$$\begin{cases} p_{p0} = N_B, \end{cases} \quad (\text{II.1.8б})$$

а концентрации неосновных носителей определяются законом действующих масс

$$\begin{cases} p_{n0} = n_i^2 / N_{Eef}, \end{cases} \quad (\text{II.1.9a})$$

$$\begin{cases} n_{p0} = n_i^2 / p_{p0} = n_i^2 / N_B. \end{cases} \quad (\text{II.1.9б})$$

Здесь и ниже индексы « p » и « n » соответствуют p - и n -областям, а индекс «0» — состоянию термодинамического равновесия. В соотношении (II.1.9a) возможное вырождение эмиттера учтено использованием эффективного значения концентрации примеси в эмиттере (см. разд. I.3), которое определяется соотношениями (I.3.20) или (I.3.24). Заметим, что концентрация неосновных носителей в базе больше, чем в эмиттере (а при $N_{Eef} \gg N_B$ — много больше).

На рис. II.1.2, а распределение примеси и носителей заряда представлено в полулогарифмическом масштабе, так как в линейном масштабе изобразить концентрации неосновных носителей невозможно. Переход занимает область $-l_{n0} < x < l_{p0}$. Границы перехода $x = -l_{n0}$ и $x = l_{p0}$ определены в некоторой степени условно, так как концентрации основных носителей изменяются плавно. Однако из рисунка очевидно, что уже на небольшом расстоянии от границ внутри перехода выполняются неравенства

$$\begin{cases} n_0 \ll N_E, \end{cases} \quad (\text{II.1.10a})$$

$$\begin{cases} p_0 \ll N_B. \end{cases} \quad (\text{II.1.10б})$$

Неравенства (II.1.10) соответствуют *приближению полного обеднения* ОПЗ носителями заряда, согласно которому внутри ОПЗ

$$n_0 + p_0 \ll N_B \leq N_E. \quad (\text{II.1.11})$$



На рис. II.1.2, б распределения концентрации примесей и носителей заряда изображены в линейном масштабе. Из рисунка очевидно, что в эмиттерной области перехода $-l_{n0} < x < 0$ концентрация подвижных носителей пренебрежимо мала по сравнению с концентрацией примеси. Эта область имеет положительный объемный заряд, плотность которого с учетом (II.1.10а) не зависит от координаты. В базовой области перехода ($0 < x < l_{p0}$) плотность объемного заряда отрицательна:

$$\left\{ \begin{aligned} \rho_E &= eN_E, \\ \rho_B &= -eN_B. \end{aligned} \right. \quad \begin{aligned} & \text{(II.1.12a)} \\ & \text{(II.1.12б)} \end{aligned}$$

Физическая причина образования объемного заряда заключается в том, что под действием градиента концентрации дырки и электроны диффундируют из эмиттера в базу и из базы в эмиттер, соответственно, оставляя в *p-n* переходе оголенные примесные ионы. Процесс диффузии продолжается до тех пор, пока не уравнивается возникающим электрическим полем.

Из рис. II.1.2 очевидно, что в некоторой плоскости $x = x_{phys}$ концентрации электронов и дырок одинаковы

$$n_0(x_{phys}) = p_0(x_{phys}) = n_i. \quad \text{(II.1.13)}$$

Эта плоскость называется плоскостью *физического* перехода в отличие от плоскости *металлургического* (или *технологического*) перехода $x = 0$, где результирующая концентрация примеси равна нулю. В симметричных переходах плоскости физического и металлургического переходов совпадают. Заметим, что физический и металлургический переходы являются плоскостями, в то время как собственно *p-n* переход — пространственная область конечного объема.

Основные выводы

1. При резком изменении типа проводимости в полупроводнике возникает область объемного заряда, называемая *p-n* переходом.
2. Условием образования *p-n* перехода является существенное изменение результирующей концентрации примеси на дебаевской длине экранирования (II.1.6).
3. В области *p-n* перехода выполняется приближение полного обеднения ОПЗ подвижными носителями заряда, согласно которому концентрации подвижных носителей пренебрежимо малы по сравнению с концентрациями примесных ионов (II.1.11).



II.2. Равновесное состояние p - n перехода

II.2.1. Энергетическая диаграмма p - n перехода

Для анализа физических процессов в полупроводниковых приборах удобно использовать метод энергетических диаграмм. Энергетическая диаграмма p - n перехода в состоянии термодинамического равновесия представлена на рис. II.2.1. По оси ординат отложена энергия электрона E . Энергия дырок на диаграмме возрастает в направлении $-E$. Так как частицы стремятся занять состояния с минимальной энергией, электроны на диаграмме имеют тенденцию «утонуть», а дырки — «всплыть». При отсутствии вырождения общий для всей системы уровень Ферми с энергией F расположен внутри запрещенной зоны, ширина которой $E_g = eV_g$ не зависит от координаты.

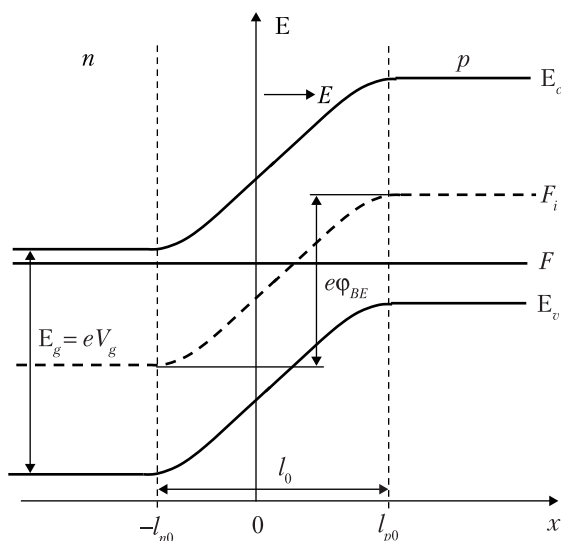


Рис. II.2.1. Энергетическая диаграмма равновесного p - n перехода

Уровень электростатической энергии F_i , показанный на рис. II.2.1 штриховой линией, соответствует положению уровня Ферми в собственном полупроводнике и расположен вблизи середины запрещенной зоны:

$$F_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \frac{m_{pd}}{m_{nd}} \approx \frac{E_c + E_v}{2},$$

где m_{pd} и m_{nd} — эффективные массы плотности состояния электронов и дырок.

Очевидно, что расстояния между уровнями E_c , E_v и F_i не зависят от координаты.

Поскольку групповые скорости электронов с энергией E_c и дырок с энергией E_v равны нулю, границы запрещенной зоны могут рассматриваться как потенциальные энергии электронов E_c и дырок E_v . При этом электрический потенциал, соответствующий точке с координатой $\vec{r}$, составляет

$$\varphi(\vec{r}) = C_1 - E_c(\vec{r})/e = C_2 - E_v(\vec{r})/e = C_3 - F_i(\vec{r})/e, \quad (\text{II.2.1})$$

где C_1, C_2, C_3 — константы, не зависящие от $\vec{r}$.

Вместо энергетических диаграмм иногда используются потенциальные диаграммы, в которых по оси ординат отложен электрический потенциал. В соответствии с (II.2.1) потенциальные диаграммы отличаются от энергетических направлением вертикальной оси и масштабным коэффициентом, равным заряду электрона.

Наклон уровней E_c , E_v и F_i на энергетической диаграмме пропорционален напряженности электрического поля E . Действительно,

$$E = -\nabla\varphi = \nabla E_c / e = \nabla E_v / e = \nabla F_i / e. \quad (\text{II.2.2})$$

В электрически нейтральных областях эмиттера ($x < -l_{n0}$) и базы ($x > l_{p0}$) поле равно нулю, и уровни E_c , E_v и F_i расположены горизонтально. В области *p-n* перехода ($-l_{n0} < x < l_{p0}$) электрическое поле направлено слева направо (вдоль градиента F_i), что находится в соответствии с рис. II.1.2, б (от «плюса» к «минусу»).

Газ *неосновных* носителей заряда всегда невырожден, поэтому их равновесные концентрации определяются взаимным расположением уровней F и $E_{c,v}$ либо уровней F и F_i соотношениями

$$n_{p0} = N_c e^{(F - E_c)/kT} = n_i e^{(F - F_i)/kT}, \quad (\text{II.2.3})$$

$$p_{n0} = N_v e^{(E_v - F)/kT} = n_i e^{(F_i - F)/kT}. \quad (\text{II.2.4})$$

Для *основных* носителей в вырожденном эмиттере эти соотношения несправедливы.



В эмиттере n -типа ($x < -l_{n0}$) фермиевская энергия больше электростатической энергии: $F > F_i$, $n_{n0} > p_{n0}$; в базе p -типа ($x > l_{p0}$): $F < F_i$, $p_{p0} > n_{p0}$.

В плоскости физического перехода $x = x_{phys}$ выполняется условие

$$F_i(x_{phys}) = F. \quad (\text{II.2.5})$$

Ввиду искривления запрещенной зоны в области перехода между эмиттером и базой существует энергетический барьер, высота которого равна разности электростатических энергий в n - и p -областях (см. рис. II.2.1). Соответственно потенциалы базы и эмиттера отличаются на величину

$$\varphi_{BE} = (F_{ip} - F_{in}) / e, \quad (\text{II.2.6})$$

которая называется *контактной разностью потенциалов*.

Энергетический барьер препятствует диффузионным потокам электронов из эмиттера в базу и дырок из базы в эмиттер. Величина барьера автоматически устанавливается такой, чтобы точно скомпенсировать диффузионные потоки.

II.2.2. Токи через p - n переход в равновесном состоянии

Сквозные токи через p - n переход в состоянии равновесия связаны с наличием градиентов концентрации подвижных носителей заряда, а также градиента электрического потенциала в p - n переходе. Кроме того, через переход протекают токи, связанные с рекомбинацией и генерацией носителей заряда внутри p - n перехода. Механизмы протекания этих токов схематически представлены на рис. II.2.2, где дырки изображены кружками со знаком «+», а электроны — со знаком «-».

Концентрация электронов в зоне проводимости n -области уменьшается по мере увеличения их энергии от уровня E_c . Под действием хаотического теплового движения электроны могут попадать из n -области в p - n переход. Наименее энергичные электроны (с энергией, близкой к E_c) отражаются барьером и возвращаются в n -область (процесс 1 на рис. II.2.2, *a*). Более энергичные электроны дальше проникают в область p - n перехода, однако, если их кинетическая энергия меньше высоты барьера $e\varphi_{BE}$, они также возвращаются в n -область, не вызывая тока через переход (процесс 2). Наконец, энергичные электроны с кинетической энергией, большей $e\varphi_{BE}$, могут преодолеть барьер (процесс 3). С этими электронами связан ток инжекции $j_{n\text{inj}0}$, направленный от p -области к



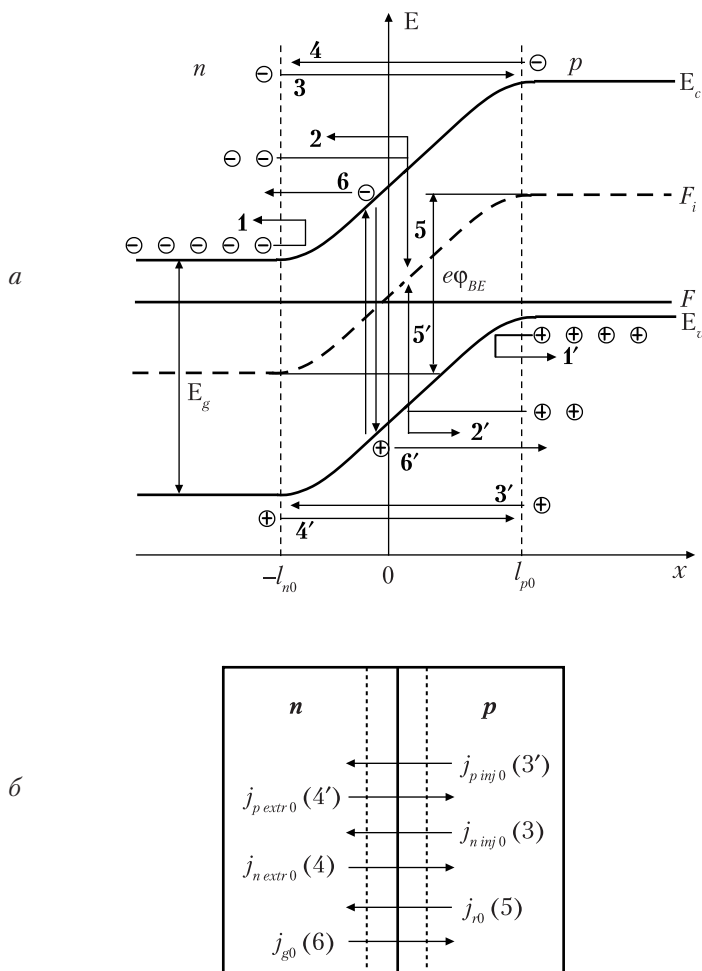


Рис. II.2.2. Механизмы переноса носителей заряда через *p-n* переход (а) и связанные с ними токи инжекции, экстракции, рекомбинации и генерации (б)

n-области (рис. II.2.2, б), т.е. против оси *x* (направление электронного тока инжекции противоположно направлению потока электронов).

Этот ток полностью компенсируется встречным потоком электронов из *p*-области. В *p*-области электроны являются

неосновными носителями и содержатся в весьма малом количестве. Если под действием теплового движения электроны попадают из *p*-области в *p-n* переход, они подхватываются электрическим полем перехода и переходят в *n*-область (процесс 4), вызывая протекание через переход *тока экстракции* $j_{n\text{extr}0}$, направленного вдоль оси *x* (вдоль поля). В состоянии термодинамического равновесия электронные токи инжекции и экстракции в точности компенсируют друг друга. Аналогичным образом взаимно компенсируются токи инжекции и экстракции дырок (процессы 3'—4'):

$$\begin{cases} j_{n\text{inj}0} + j_{n\text{extr}0} = 0. \end{cases} \quad (\text{II.2.7a})$$

$$\begin{cases} j_{p\text{inj}0} + j_{p\text{extr}0} = 0. \end{cases} \quad (\text{II.2.7б})$$

Помимо рассмотренных механизмов протекания тока существуют токи, связанные с процессами термогенерации и рекомбинации электронно-дырочных пар в области перехода. Дырки и электроны, проникающие в переход со стороны *p*- и *n*-областей, соответственно, имеют конечную вероятность рекомбинировать в переходе (процесс 5—5'). Рекомбинирующие носители попадают в место рекомбинации с *разных* сторон перехода (электроны — из *n*-, а дырки — из *p*-области). Поэтому каждый акт рекомбинации связан с переносом заряда *e* через переход от *p*- к *n*-области. Таким образом, рекомбинация носителей заряда в ОПЗ приводит к протеканию *тока рекомбинации* j_{r0} в направлении оси *x*. При термогенерации электронно-дырочных пар в переходе образовавшиеся носители заряда подхватываются электрическим полем, причем электроны переносятся в *n*-область, а дырки — в *p*-область (процесс 6—6'). Возникающий при этом *ток термогенерации* j_{g0} направлен против оси *x* (вдоль поля) и в точности компенсирует ток рекомбинации

$$j_{r0} + j_{g0} = 0. \quad (\text{II.2.7в})$$

Разумеется, суммарная плотность тока через переход в состоянии равновесия равна нулю

$$j_0 = j_{n\text{inj}0} + j_{n\text{extr}0} + j_{p\text{inj}0} + j_{p\text{extr}0} + j_{r0} + j_{g0} = 0. \quad (\text{II.2.8})$$

Система равенств (II.2.7а, б, в) сильнее равенства (II.2.8), так как свидетельствует о *детальном* равновесии встречных потоков электронов и дырок. Это равновесие является частным случаем проявления общего принципа *детального* равновесия.



Заметим также, что каждый из рассмотренных токов должен иметь малую величину. Токи экстракции малы ввиду того, что они создаются неосновными носителями в *p*- и *n*-областях, концентрации которых чрезвычайно малы. Токи инжекции малы ввиду того, что они создаются только наиболее энергичными носителями с кинетической энергией, большей высоты барьера $e\phi_{BE}$, число которых также мало. Наконец, токи рекомбинации-генерации малы ввиду малых размеров *p-n* перехода (число генерированных пар мало) и малого времени пребывания носителей в переходе (число рекомбинирующих пар мало).

Интересно отметить, что несмотря на равенство нулю суммарного тока, рассмотренные механизмы токов могут быть обнаружены экспериментально. Каждый из механизмов имеет случайный характер и поэтому проявляется в виде шумового тока, который может быть рассчитан и измерен. Разумеется, средняя по времени величина каждого из шумовых токов равна нулю, но соответствие рассчитанных и измеренных значений их спектральных плотностей служит прямым подтверждением теоретических представлений.

Выявленные механизмы токов в равновесном состоянии позволяют предсказать качественный вид ВАХ *p-n* перехода. Если приложить к переходу напряжение положительной полярности $V > 0$ («+» к *p*-области, «-» к *n*-области), высота потенциального барьера $e\phi_{BE}$ между *p*- и *n*-областями снижается, что должно привести к резкому возрастанию токов инжекции. При обратной полярности внешнего напряжения высота потенциального барьера возрастает, и токи инжекции практически полностью перекрываются (кривая 1 на рис. II.2.3, *a*). Токи экстракции не связаны с необходимостью преодоления потенциального барьера, поэтому их величина не зависит от приложенного к переходу напряжения (прямая 2 на рис. II.2.3, *a*). В результате ВАХ токов инжекции-экстракции имеет ярко выраженный «диодный» характер, обеспечивающий односторонний характер проводимости (кривая 3 на рис. II.2.3, *a*).

Ток рекомбинации в переходе также должен быстро возрастать с ростом напряжения положительной полярности, так как снижение барьера приводит к увеличению расстояния, на которое основные носители заряда могут проникать в переход. Таким образом, с ростом прямого напряжения повышается количество электронов и дырок в переходе, и, следовательно, интенсивность их рекомбинации. При обратном напряжении



достаточно большой величины носители заряда практически не попадают в ОПЗ, и ток рекомбинации снижается до нуля (кривая 4 на рис. II.2.3, б). Ток термогенерации не зависит от концентрации электронов и дырок и на первый взгляд не должен зависеть от напряжения. Однако, как будет показано ниже, с ростом обратного напряжения увеличивается ширина ОПЗ, что приводит к возрастанию тока термогенерации (кривая 5 на рис. II.2.3, б). В результате ВАХ токов генерации-рекомбинации также имеет «диодный» характер, хотя и не столь ярко выраженный (кривая 6 на рис. II.2.3, б).

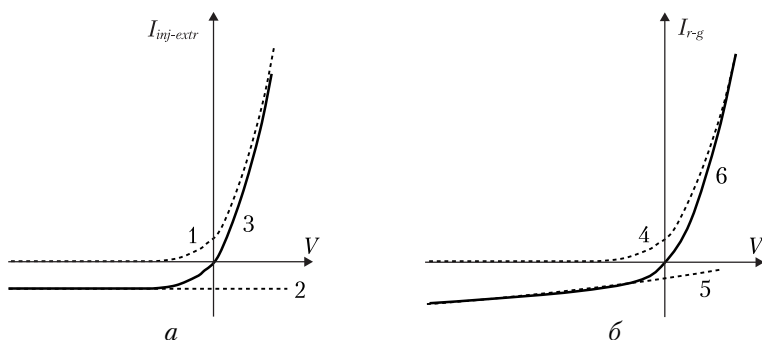


Рис. II.2.3. Вид ВАХ токов инжекции-экстракции (а) и рекомбинации-генерации (б):

1 — $I_{inj}(V)$; 2 — $I_{extr}(V)$; 3 — $I_{inj}(V) + I_{extr}(V)$; 4 — $I_r(V)$;
5 — $I_g(V)$; 6 — $I_r(V) + I_g(V)$

II.2.3. Методика определения параметров p - n перехода

Основными параметрами p - n перехода являются контактная разность потенциалов Φ_{BE} , протяженности перехода в p - и n -областях l_{p0} и l_{n0} , а также максимальная напряженность электрического поля E_{m0} . Кроме того, необходимо знать распределение напряженности электрического поля в ОПЗ $E(x)$. Исходными данными являются результирующие примесные профили в базе $N_B(x)$ и эмиттере $N_E(x)$, а также электрофизические параметры полупроводникового материала.

Контактная разность потенциалов может быть определена с помощью соотношений (II.2.3)—(II.2.4) и (II.2.6). Учитывая, что концентрации носителей заряда на границах перехода (в плоскостях $x = -l_{n0}$ и $x = l_{p0}$ на рис. II.2.1) соответствуют равновесным значениям (см. рис. II.1.2, а), получаем

$$p_{n0}(-l_{n0}) = n_i e^{[F_i(-l_{n0}) - F]/kT} = n_i^2 / N_{Eef}(-l_{n0}), \quad (\text{II.2.9a})$$

$$n_{p0}(l_{p0}) = n_i e^{[F - F_i(l_{p0})]/kT} = n_i^2 / N_B(l_{p0}). \quad (\text{II.2.9б})$$

Перемножив равенства (II.2.9a, б), находим

$$\varphi_{BE} = \varphi_T \ln \frac{N_{Eef}(-l_{n0}) N_B(l_{p0})}{n_i^2}. \quad (\text{II.2.10})$$

Из рис. II.2.1 следует, что при отсутствии вырождения (когда уровень Ферми лежит в запрещенной зоне) высота потенциального барьера не может превышать ширины запрещенной зоны E_g . При этом

$$\varphi_{BE} < V_g = E_g / e. \quad (\text{II.2.11})$$

Очевидно, что контактная разность потенциалов увеличивается с увеличением степени легирования эмиттера и базы. Переходы, изготовленные на основе полупроводника с большей шириной запрещенной зоны (и, следовательно, меньшей собственной концентрацией носителей заряда n_i), имеют большую контактную разность потенциалов.

Основными допущениями при анализе перехода являются пренебрежение концентрациями подвижных носителей заряда в ОПЗ по сравнению с концентрациями примесей — приближение полного обеднения (II.1.11), а также электронейтральность полупроводника вне ОПЗ. При этом распределение плотности объемного заряда имеет вид

$$\rho(x) = \begin{cases} 0, & x \leq -l_{n0}, \\ eN(x), & -l_{n0} < x < l_{p0}, \\ 0, & x \geq l_{p0}. \end{cases} \quad (\text{II.2.12})$$

Электрическое поле может быть найдено из уравнения Пуассона

$$\frac{dE(x)}{dx} = \frac{\rho(x)}{\epsilon \epsilon_0}. \quad (\text{II.2.13})$$

На границах перехода (в плоскостях $x = -l_{n0}$ и $x = l_{p0}$) поле равно нулю; поэтому согласно теореме Гаусса в целом переход электронейтрален



$$\int_{-l_{n0}}^{l_{p0}} N(x) dx = 0, \quad (\text{II.2.14})$$

а распределение поля в ОПЗ определяется соотношением

$$E(x) = \frac{e}{\epsilon \epsilon_0} \int_{-l_{n0}}^x N(x) dx. \quad (\text{II.2.15})$$

Поскольку функция $\rho(x)$ меняет знак в плоскости $x = 0$, максимальная напряженность электрического поля составляет

$$E_{m0} = \frac{e}{\epsilon \epsilon_0} \int_{-l_{n0}}^0 N(x) dx = -\frac{e}{\epsilon \epsilon_0} \int_0^{l_{p0}} N(x) dx. \quad (\text{II.2.16})$$

Интегрирование напряженности электрического поля по всей ОПЗ дает второе уравнение для контактной разности потенциалов

$$\Phi_{BE} = - \int_{-l_{n0}}^{l_{p0}} E(x) dx = -\frac{e}{\epsilon \epsilon_0} \int_{-l_{n0}}^{l_{p0}} dx \int_{-l_{n0}}^x N(x) dx.$$

Интегрируя по частям, с учетом (II.2.14) получаем

$$\Phi_{BE} = \frac{e}{\epsilon \epsilon_0} \int_{-l_{n0}}^{l_{p0}} x N(x) dx. \quad (\text{II.2.17})$$

Уравнения (II.2.10), (II.2.14), (II.2.16) и (II.2.17) с учетом очевидного равенства

$$l_0 = l_{p0} + l_{n0} \quad (\text{II.2.18})$$

могут быть решены относительно неизвестных Φ_{BE} , l_{p0} , l_{n0} и E_{m0} , после чего из (II.2.15) определяется функция $E(x)$.

Отметим, что в набор уравнений, определяющих параметры p - n перехода, входят два уравнения для контактной разности потенциалов. Уравнение (II.2.10) определяет значение Φ_{BE} из соображений статистики носителей заряда, а уравнение (II.2.17) является уравнением электростатики.



II.2.4. Расчет параметров ступенчатого *p-n* перехода

Наиболее просто определяются параметры ступенчатого *p-n* перехода, так как в этом случае функция $N(x)$ имеет вид

$$N(x) = \begin{cases} N_E, & x \leq 0, \\ N_B, & x \geq 0, \end{cases} \quad (\text{II.2.19})$$

и значения граничных концентраций примеси $N(-l_{n0})$ и $N(l_{p0})$ известны

$$\begin{cases} N(-l_{n0}) = N_E, & (\text{II.2.20a}) \end{cases}$$

$$\begin{cases} N(l_{p0}) = N_B, & (\text{II.2.20б}) \end{cases}$$

На рис. II.2.4 представлены эпюры распределения примеси и подвижных носителей заряда, плотности заряда, электрического поля и потенциала в ступенчатом *p-n* переходе.

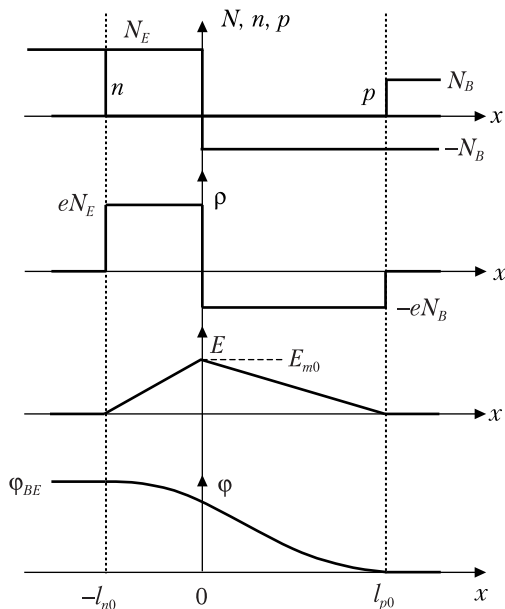


Рис. II.2.4. Эпюры распределения примеси и подвижных носителей заряда, плотности заряда, электрического поля и потенциала в ступенчатом *p-n* переходе

Контактная разность потенциалов определяется из уравнения (II.2.10)

$$\Phi_{BE} = \Phi_T \ln \frac{N_{Eef} N_B}{n_i^2}. \quad (\text{II.2.21})$$

Подставляя (II.2.19) в (II.2.14) и (II.2.17), с учетом (II.2.18) получаем

$$l_0 = \sqrt{2\epsilon\epsilon_0\Phi_{BE}(N_B^{-1} + N_E^{-1}) / e}; \quad (\text{II.2.22})$$

$$l_{n0} = l_0 N_B / (N_E + N_B); \quad (\text{II.2.23a})$$

$$l_{p0} = l_0 N_E / (N_E + N_B). \quad (\text{II.2.23б})$$

Максимальная напряженность электрического поля определяется из (II.2.16)

$$E_{m0} = \sqrt{\frac{2e\Phi_{BE}}{\epsilon\epsilon_0(N_E^{-1} + N_B^{-1})}}. \quad (\text{II.2.24})$$

Из (II.2.22) следует, что при условии $N_E \gg N_B$ практически весь переход сосредоточен в области базы ($l_{n0} \ll l_{p0} \approx l_0$). Величина Φ_{BE} слабо (логарифмически) зависит от концентрации примеси в эмиттере, поэтому при $N_E \gg N_B$ параметры перехода определяются практически только свойствами базы:

$$l_0 \approx \sqrt{2\epsilon\epsilon_0\Phi_{BE} / eN_B}; \quad (\text{II.2.25})$$

$$E_{m0} = \frac{2\Phi_{BE}}{l_0} \approx \sqrt{\frac{2e\Phi_{BE}N_B}{\epsilon\epsilon_0}}. \quad (\text{II.2.26})$$

В качестве примера приведем значения параметров кремниевых и германиевых ступенчатых *p-n* переходов для случая $N_E = 1 \cdot 10^{18} \text{ см}^{-3}$, $N_B = 1 \cdot 10^{17} \text{ см}^{-3}$. При $T = 300 \text{ К}$ для кремния: $E_g = 1,12 \text{ эВ}$, $\epsilon = 12$, $n_i = 10^{10} \text{ см}^{-3}$; для германия: $E_g = 0,72 \text{ эВ}$, $\epsilon = 16$, $n_i = 10^{13} \text{ см}^{-3}$. Из (II.2.10), (II.2.25) и (II.2.26) получаем

Параметр	Si	Ge
Φ_{BE} , В	0,79	0,54
l_0 , мкм	0,10	0,10
E_{m0} , В/см	$1,52 \cdot 10^5$	$1,04 \cdot 10^5$

В обоих случаях контактная разность потенциалов составляет примерно $\frac{2}{3}$ ширины запрещенной зоны в вольтах, а ширина перехода весьма мала (0,1 мкм). Ввиду этого электрические поля достигают более чем 100 кВ/см, причем для кремниевого перехода эта величина больше из-за более широкой запрещенной зоны. С уменьшением степени легирования базы ширина перехода возрастает, электрическое поле в переходе уменьшается, величина Φ_{BE} тоже уменьшается, но весьма слабо.

Типичные значения параметров *p-n* перехода

$$\Phi_{BE} \approx 2E_g / e3; \quad l_0 = (0,01-1) \text{ мкм}; \quad E_{m0} = (30-300) \text{ кВ/см}.$$

Принятые допущения предполагают, что распределения дырок и электронов в ступенчатом переходе также имеют характер ступенчатых функций (см. рис. II.2.4). Реально функции $p(x)$ и $n(x)$ не имеют скачков. Длины «хвостов» этих функций примерно равны дебаевским длинам экранирования соответственно в эмиттере и базе. Наибольшую длину имеет длина «хвоста» функции $n(x)$ в более слабо легированной базе, где дебаевская длина экранирования составляет

$$L_{DB} = \sqrt{\frac{\epsilon \epsilon_0 \Phi_T}{e N_B}}.$$

Сравнение этого соотношения с (II.2.25) показывает, что модель полного обеднения может считаться справедливой при условии

$$2\Phi_{BE} \gg \Phi_T. \quad (\text{II.2.27})$$

В этом случае ширина ОПЗ значительно превышает длину «хвоста» функции $n(x)$ ($l_0 \gg L_{DB}$). Условие (II.2.27) выполнено практически всегда.

II.2.5. Переход с линейным распределением примеси

В реальных *p-n* переходах распределение результирующей примеси $N(x) = N_d(x) - N_a(x)$ описывается непрерывной функцией. Если кривизна функции $N(x)$ невелика, а переход достаточно узкий, то в пределах перехода ($-l_{n0} \leq x \leq l_{p0}$) функция $N(x)$ может считаться линейной:

$$N(x) \approx -N' \cdot x, \quad (\text{II.2.28})$$

где $N' = |dN / dx|$ — абсолютная величина градиента концентрации примеси. Эпюры распределения примеси, носителей



заряда, плотности заряда, поля и потенциала в линейном p - n переходе представлены на рис. II.2.5.

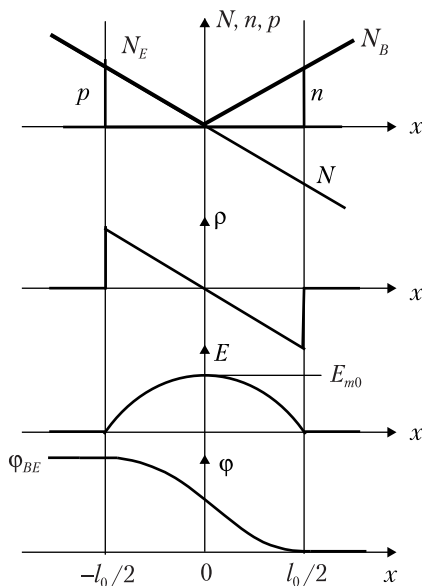


Рис. II.2.5. Эпюры распределения примеси и подвижных носителей заряда, плотности заряда, электрического поля и потенциала в линейном p - n переходе

Разумеется, линейный переход является симметричным

$$l_{p0} = l_{n0} = l_0 / 2.$$

На границах линейного перехода полупроводник практически никогда не бывает вырожден и $N_{Eef}(-l_{n0}) = N_E(-l_{n0}) = N_B(l_{p0}) = N' l_0 / 2$. При этом из (II.2.10) получаем

$$\phi_{BE} = 2\phi_T \ln \frac{N' l_0}{2n_i}. \quad (\text{II.2.29})$$

В отличие от ступенчатого, в линейном переходе величина контактной разности потенциалов не может быть определена независимо от ширины перехода l_0 . Подставляя (II.2.28) в (II.2.17) и интегрируя, получаем

$$\phi_{BE} = \frac{e N' l_0^3}{12 \epsilon \epsilon_0}. \quad (\text{II.2.30})$$

Отсюда

$$l_0 = \sqrt[3]{\frac{12\epsilon\epsilon_0\phi_{BE}}{eN'}}. \quad (\text{II.2.31})$$

Параметры ϕ_{BE} и l_0 находятся совместным решением уравнений (II.2.29) и (II.2.31). Для их определения удобно использовать метод последовательных приближений, основанный на том, что функция $\phi_{BE}(l_0)$, описываемая уравнением (II.2.31), изменяется довольно быстро ($\phi_{BE} \sim l_0^3$), а уравнением (II.2.29) — чрезвычайно медленно (логарифмически). Рисунок II.2.6 иллюстрирует порядок решения уравнений. Ширина перехода в i -ом приближении $l_0^{(i)}$ определяется уравнением (II.2.31) при $\phi_{BE} = \phi_{BE}^{(i-1)}$

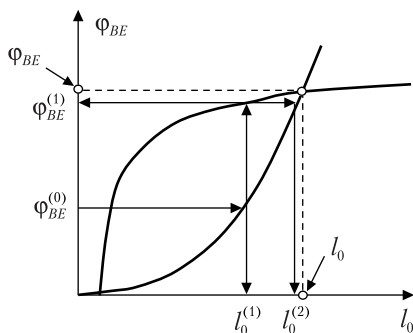


Рис. II.2.6. Определение параметров ϕ_{BE} и l_0 в линейном *p-n* переходе методом итераций

$$l_0^{(i)} = \sqrt[3]{\frac{12\epsilon\epsilon_0\phi_{BE}^{(i-1)}}{eN'}}. \quad (\text{II.2.32})$$

Контактная разность потенциалов в i -ом приближении определяется уравнением (II.2.29) при $l_0 = l_0^{(i)}$:

$$\phi_{BE}^{(i)} = 2\phi_T \ln \frac{N' l_0^{(i)}}{2n_i}. \quad (\text{II.2.33})$$

Нулевое приближение $\phi_{BE}^{(0)}$ выбирается в значительной степени произвольно, например, $\phi_{BE}^{(0)} = 2E_g / 3e$. В большинстве практических случаев необходимая точность обеспечивается вторым приближением.

II.2.6. Диффузионные *p-n* переходы

В *p-n* переходах, полученных методом диффузии, распределение результирующей примеси имеет вид одной из следующих функций:

$$N(x) = N_d(x) - N_a(x) = N_{ds} \exp \frac{-(x-x_s)^2}{4Dt} - N_a(x); \quad (\text{II.2.34a})$$

$$N(x) = N_d(x) - N_a(x) = N_{ds} \operatorname{erfc} \frac{x-x_s}{2\sqrt{Dt}} - N_a(x), \quad (\text{II.2.34б})$$

где x_s — координата поверхности кристалла; N_{ds} — поверхностная концентрация доноров; D — коэффициент диффузии донорной примеси; t — время диффузии (рис. II.2.7). Конкретный вид функции $N(x)$ зависит от условий диффузии [5]. Как правило, концентрация исходной примеси (в данном случае акцепторов) в ОПЗ изменяется незначительно и поэтому может считаться постоянной:

$$N_a(x) \approx N_a(0) = N_0 \quad \text{при} \quad -l_{n0} < x < l_{p0}. \quad (\text{II.2.35})$$

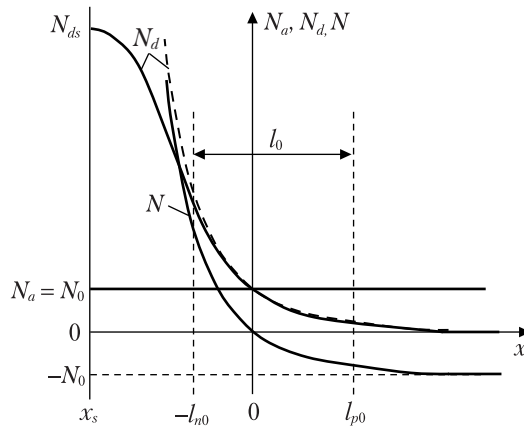


Рис. II.2.7. Распределение доноров N_d , акцепторов N_a и результирующей примеси $N = N_d - N_a$ в диффузионном *p-n* переходе

С достаточной для практических целей точностью распределение концентрации акцепторов в ОПЗ можно аппроксимировать экспонентой (штриховая линия на рис. II.2.7):

$$N_d(x) \approx N_0 e^{-x/L} \quad \text{при} \quad -l_{n0} < x < l_{p0}. \quad (\text{II.2.36})$$

Параметр L выбирается таким образом, чтобы в плоскости металлургического перехода ($x = 0$) градиент концентрации примеси, полученный из (II.2.36), соответствовал градиенту действительного распределения (II.2.34а, б):

$$L = \frac{N_0}{|dN_d / dx|}. \quad (\text{II.2.37})$$

Таким образом, результирующее распределение примеси в переходе (см. рис. II.2.7) описывается функцией

$$N(x) = N_0(e^{-x/L} - 1). \quad (\text{II.2.38})$$

Подстановка (II.2.38) в (II.2.10), (II.2.14) и (II.2.17) с учетом соотношения $l_0 = l_{n0} + l_{p0}$ дает

$$\left\{ \begin{aligned} e^{l_{n0}/L} &= \frac{l_0 / L}{1 - e^{-l_0/L}}, \end{aligned} \right. \quad (\text{II.2.39a})$$

$$\left\{ \begin{aligned} e^{-l_{p0}/L} &= \frac{l_0 / L}{e^{l_0/L} - 1}, \end{aligned} \right. \quad (\text{II.2.39б})$$

$$\left\{ \begin{aligned} \frac{\Phi_{BE}}{\Phi_T} &= 2 \ln \frac{N_0}{n_i} + \ln \left[\left(\frac{l_0 / L}{1 - e^{-l_0/L}} - 1 \right) \left(1 - \frac{l_0 / L}{e^{l_0/L} - 1} \right) \right], \end{aligned} \right. \quad (\text{II.2.40a})$$

или
$$\frac{\Phi_{BE}}{\Phi_T} = \frac{L^2}{L_D^2} \cdot \frac{l_0}{L_D} \left(1 + \frac{l_0}{2L} \operatorname{cth} \frac{l_0}{2L} \right). \quad (\text{II.2.40б})$$

В соотношении (II.2.40б)

$$L_D = \sqrt{\frac{\epsilon \epsilon_0 \Phi_T}{e N_0}} \quad (\text{II.2.41})$$

— дебаевская длина экранирования в исходном полупроводнике с концентрацией примеси N_0 .

Анализ соотношений (II.2.39)–(II.2.41) показывает, что при выполнении неравенства

$$l_0 / L < 1 \quad (\text{II.2.42})$$

свойства диффузионного *p-n* перехода очень близки к свойствам перехода с линейным распределением примеси ($l_{n0} \approx l_{p0}$, $l_0 \sim \Phi_{BE}^{1/3}$), и параметры перехода можно рассчитывать по методике, изложенной в параграфе II.2.5. При выполнении неравенства

$$l_0 / L > 10 \quad (\text{II.2.43})$$



в соотношениях (II.2.39)–(II.2.41) можно считать

$$e^{l_0/L} \gg l_0 / L \gg 1; \quad \text{cth}(l_0 / L) \approx 1,$$

и форма их записи упрощается

$$\frac{\Phi_{BE}}{\Phi_T} \approx \ln \frac{N_0^2 l_0}{n_i^2 L}; \quad (\text{II.2.44})$$

$$\frac{\Phi_{BE}}{\Phi_T} \approx \frac{l_0^2}{2L_D^2}; \quad (\text{II.2.45})$$

$$l_{n0} \ll l_{p0} \approx l_0. \quad (\text{II.2.46})$$

В этом случае свойства диффузионного перехода очень близки к свойствам резко асимметричного ступенчатого перехода ($l_{n0} \ll l_{p0} \approx l_0 \sim \Phi_{BE}^{1/2}$). Для нахождения параметров l_0 и Φ_{BE} из уравнений (II.2.44) и (II.2.45) можно использовать метод последовательных приближений, полагая $\Phi_{BE}^{(0)} = 2E_g / 3e$:

$$l_0^{(i)} = L_D \sqrt{2\Phi_{BE}^{(i-1)} / \Phi_T}; \quad (\text{II.2.47})$$

$$\Phi_{BE}^{(i)} = \Phi_T \ln \frac{N_0^2 l_0^{(i)}}{n_i^2 L}. \quad (\text{II.2.48})$$

Условие (II.2.42) выполняется при выполнении неравенства

$$L / L_D \gg \ln^{1/2}(N_0 / n_i), \quad (\text{II.2.49})$$

а условие (II.2.43) — при выполнении неравенства

$$L / L_D \ll \ln^{1/2}(N_0 / n_i). \quad (\text{II.2.50})$$

В промежуточном случае для определения параметров диффузионных переходов следует решать систему уравнений (II.2.39)–(II.2.41) численными методами либо использовать номограмму, приведенную в работе [5].

Основные выводы

1. В состоянии термодинамического равновесия между p - и n -областями образуется контактная разность потенциалов. Через p - n переход протекают токи инжекции и экстракции, а также токи, связанные с термогенерацией и рекомбинацией носителей заряда в ОПЗ.



2. Токи инжекции обусловлены диффузией носителей заряда из области, где они являются основными, в ту область, где они являются неосновными. Эти токи создаются энергичными носителями заряда, которые имеют кинетическую энергию, большую высоты потенциального барьера. Токи инжекции направлены от p -области к n -области.

3. Токи экстракции обусловлены попаданием в ОПЗ неосновных носителей заряда из эмиттера и базы и перебросом этих носителей через переход электрическим полем. Эти токи направлены от n -области к p -области.

4. Ток рекомбинации связан с рекомбинацией электронов и дырок, попадающих в ОПЗ из тех областей, где они являются основными носителями заряда. Этот ток направлен от p -области к n -области.

5. Ток термогенерации связан с тепловым возбуждением электронно-дырочных пар в p - n переходе. Генерируемые носители увлекаются электрическим полем в противоположных направлениях, создавая ток, направленный от n -области к p -области.

6. Высота потенциального барьера в состоянии равновесия устанавливается такой, что электронные токи, дырочные токи и токи генерации-рекомбинации в области перехода взаимно компенсируются.

7. Основными параметрами перехода являются: контактная разность потенциалов, ширина перехода и максимальная напряженность электрического поля. Протяженности p - n перехода в p - и n -областях, а также контактная разность потенциалов определяются совместным решением системы уравнений (II.2.10), (II.2.14), (II.2.17) и (II.2.18). Максимальное поле в переходе определяется уравнением (II.2.16).

8. Параметры ступенчатого перехода определяются соотношениями (II.2.21)–(II.2.24). Ширина перехода пропорциональна квадратному корню из контактной разности потенциалов и уменьшается с увеличением степени легирования базы и эмиттера. В случае, когда эмиттер легирован значительно сильнее базы, практически весь переход расположен в базовой области, а его параметры определяются свойствами базы.

9. Переход с линейным распределением примеси является симметричным. Его параметры определяются уравнениями (II.2.29) и (II.2.30), которые удобно решать методом последовательных приближений. Ширина перехода пропорциональна кубическому корню из контактной разности потенциалов и уменьшается с увеличением градиента концентрации примеси в переходе.

10. Параметры диффузионных переходов удобно находить, аппроксимируя распределение примеси выражением (II.2.38). При выполнении неравенства (II.2.42) распределение примеси в диффузионном переходе может считаться линейным. При выполнении



неравенства (II.2.43) свойства диффузионного перехода близки к свойствам резко асимметричного ступенчатого перехода, а его параметры удобно находить путем решения уравнений (II.2.44) и (II.2.45) методом последовательных приближений.

II.3. Неравновесное состояние p - n перехода

II.3.1. Прямое и обратное включение p - n перехода

Подключение внешнего источника напряжения между p - и n -областями перехода изменяет высоту потенциального барьера. Если внешнее напряжение подключено плюсом к p -слою, высота барьера уменьшается. Такое включение называется прямым. Напряжение обратной полярности увеличивает высоту барьера. Такое включение называется обратным.

Так как область p - n перехода почти не содержит подвижных носителей заряда и обладает повышенным сопротивлением, большая часть внешнего напряжения V приложено к p - n переходу. Изменение высоты потенциального барьера приводит к трем главным следствиям:

1) изменяется ширина p - n перехода, значение которой зависит от напряжения:

$$l_0 \rightarrow l(V);$$

2) изменяются концентрации подвижных носителей заряда на границах p - n перехода:

$$p(-l_p, l_n) \neq p_0(-l_{p0}, l_{n0}), \quad n(-l_p, l_n) \neq n_0(-l_{p0}, l_{n0});$$

3) изменяются токи инжекции подвижных носителей заряда в те области, где они являются неосновными, а также токи рекомбинации—генерации носителей заряда в ОПЗ.

Прежде чем анализировать неравновесное состояние p - n перехода, полезно рассмотреть его энергетические диаграммы при прямом и обратном включениях.

II.3.2. Энергетические диаграммы неравновесного p - n перехода

Энергетические диаграммы p - n перехода при прямом и обратном включениях показаны на рис. II.3.1. Как мы видели (см. параграф II.2.2), в состоянии равновесия небольшая доля электронов из n -области и дырок из p -области имеет достаточ-



ную кинетическую энергию для преодоления потенциального барьера. Эти носители создают токи инжекции электронов из *n*- в *p*-область и дырок — в обратном направлении. Токи инжекции уравниваются токами экстракции неосновных носителей (электронов из *p*-области и дырок из *n*-области), которые случайно приближаются к барьеру и подхватываются его электрическим полем.

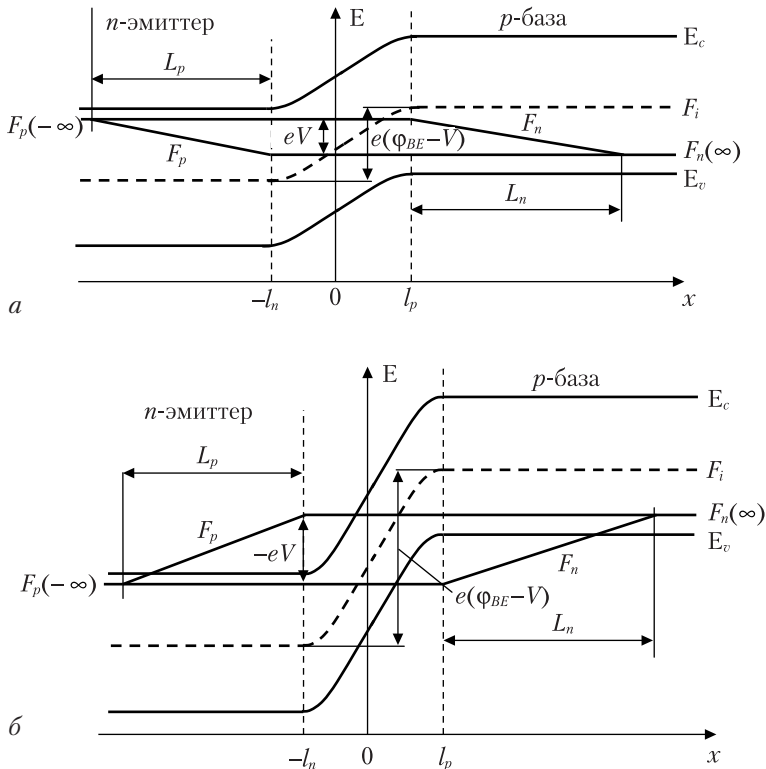


Рис. II.3.1. Энергетические диаграммы *p-n* перехода при прямом (а) и обратном (б) включениих

При *прямом* смещении (рис. II.3.1, а) вследствие уменьшения высоты барьера токи инжекции увеличиваются, а токи экстракции остаются неизменными. В результате из *p*- в *n*-область течет ток, величина которого тем больше, чем ниже барьер, т.е. чем больше прямое смещение. Высота барьера составляет $e(\phi_{BE} - V)$.

Электроны и дырки, инжектированные в *p*- и *n*-области, соответственно, нарушают равновесное распределение носителей вблизи перехода. В окрестности перехода квазиуровень Ферми для электронов лежит выше, чем квазиуровень Ферми для дырок ($F_n > F_p$). Вдали от перехода квазиуровни Ферми совпадают, так как инжектированные носители рекомбинируют, не успевая достигнуть этой области. При $V \rightarrow \phi_{BE}$ *p-n* переход заливается электронно-дырочной плазмой, его сопротивление падает и приложенное напряжение распределяется по всему образцу в соответствии с его проводимостью.

При обратном смещении (рис. II.3.1, б) вследствие увеличения высоты барьера токи инжекции практически прекращаются, а токи экстракции остаются неизменными. В результате через переход течет малый обратный ток, созданный неосновными носителями, случайно приблизившимися к *p-n* переходу. Этот ток не зависит от обратного напряжения, если оно достаточно велико. В окрестности *p-n* перехода квазиуровень Ферми для электронов расположен ниже, чем для дырок ($F_n < F_p$).

II.3.3. Граничные условия и уровни инжекции

Важной особенностью энергетических диаграмм на рис. II.3.1, а, б является *практически горизонтальное расположение квазиуровней Ферми в ОПЗ*

$$F_{n,p}(x) = \text{const.}$$

При этом квазиуровни Ферми для электронов и дырок параллельны друг другу, а расстояние между ними соответствует приложенному напряжению

$$F_n - F_p = eV. \quad (\text{II.3.1})$$

Строго говоря, утверждение о горизонтальном расположении квазиуровней Ферми в ОПЗ не может быть выполнено точно, так как согласно (I.4.12а, б) в этом случае должны быть равны нулю токи электронов и дырок. Оно справедливо в том смысле, что наклоны квазиуровней Ферми в ОПЗ много меньше, чем в квазинейтральных областях базы и эмиттера.

Качественное объяснение постоянства квазиуровней Ферми в ОПЗ основано на допущении о малости токов, связанных с рекомбинацией и генерацией носителей заряда в ОПЗ. При этом каждая из плотностей тока j_n и j_p постоянна в любой точке перехода и согласно (I.4.12а, б) ($\vec{j}_n = n\mu_n \nabla F_n$, $\vec{j}_p = p\mu_p \nabla F_p$):

$$\begin{cases} p(x)dF_p(x)/dx = \text{const}; \\ n(x)dF_n(x)/dx = \text{const}. \end{cases}$$

Очевидно, что минимальное значение концентрации электронов в ОПЗ достигается на ее границе с *p*-областью в плоскости $x = l_p$. В этой плоскости значение $dF_n(x)/dx$ максимально. По мере продвижения в ОПЗ (против оси x на рис. II.3.1) концентрация электронов резко возрастает (см., например, рис. II.1.2), следовательно, значение $dF_n(x)/dx$ резко снижается. Таким образом, практически во всей ОПЗ $dF_n(x)/dx$ остается много меньшим граничного значения в плоскости $x = l_p$, что и позволяет считать значение $F_n(x)$ постоянным. Аналогичные рассуждения применимы и в отношении функции $F_p(x)$, что приводит к равенству (II.3.1). Из соотношений (II.3.1) и (I.3.28) ($np = n_i^2 e^{(F_n - F_p)/kT}$) следует, что в ОПЗ

$$n(x)p(x) = n_i^2 e^{V/\Phi_T}. \quad (\text{II.3.2})$$

Таким образом, *произведение концентраций электронов и дырок в ОПЗ определяется напряжением на переходе и не зависит от координаты.*

В дальнейшем мы будем находить распределения концентраций носителей заряда $p(x)$ и $n(x)$ в областях базы и эмиттера путем решения биполярных уравнений непрерывности (I.7). Для этого необходимо использовать некоторые *граничные условия*. В качестве одного из граничных условий будут приняты значения *граничных концентраций* носителей заряда на границах областей базы и эмиттера с ОПЗ:

$$p_{nj} = p_n(-l_n) \text{ и } n_{pj} = n_p(l_p).$$

В параграфе I.3.2 были введены понятия *избыточных концентраций* электронов и дырок как превышения их концентраций над равновесными значениями (соотношения (I.3.29а, б)):

$$\Delta n = n - n_0; \quad \Delta p = p - p_0.$$

Области базы и эмиттера за пределами ОПЗ являются *квазинейтральными*, и согласно (I.3.30), (I.3.33)

$$\Delta n \approx \Delta p, \quad (\text{II.3.3а})$$

или

$$n - n_0 \approx p - p_0. \quad (\text{II.3.3б})$$



В дальнейшем будем полагать, что условия (II.3.3а, б) выполнены точно.

Для нахождения граничных концентраций носителей заряда можно использовать совместно условия (II.3.2) и (II.3.3а, б), полагая, что условие (II.3.2), справедливое для ОПЗ, и (II.3.3), справедливое для квазинейтральных областей базы и эмиттера, выполняются и на границах ОПЗ с квазинейтральными областями

$$\begin{cases} n_j p_j = n_i^2 e^{V/\Phi_T}, & \text{(II.3.4а)} \end{cases}$$

$$\begin{cases} n_j - n_0 \approx p_j - p_0. & \text{(II.3.4б)} \end{cases}$$

Из (II.3.4а, б) для границ ОПЗ с базой p -типа ($p_0 = N_B$, $n_0 = n_i^2 / N_B$) и эмиттером n -типа ($n_0 = N_E$, $p_0 = n_i^2 / N_{Eef}$) получим

$$\begin{cases} \Delta n_{pj} = \Delta p_{pj} = \frac{N_B}{2} \left(\sqrt{1 + \left(\frac{2n_i}{N_B} \right)^2 (e^{V/\Phi_T} - 1)} - 1 \right), & \text{(II.3.5а)} \end{cases}$$

$$\begin{cases} \Delta p_{nj} = \Delta n_{nj} = \frac{N_{Eef}}{2} \left(\sqrt{1 + \left(\frac{2n_i}{N_{Eef}} \right)^2 (e^{V/\Phi_T} - 1)} - 1 \right). & \text{(II.3.5б)} \end{cases}$$

Степень нарушения равновесия на границах оценивается уровнем инжекции δ , который определен соотношениями (I.3.31). Для базы p -типа

$$\delta_B = n_{pj} / p_{p0} = n_{pj} / N_B;$$

для эмиттера n -типа

$$\delta_E = p_{nj} / n_{n0} = p_{nj} / N_E.$$

При $\delta \ll 1$ уровень инжекции считается *низким*, при $\delta \gg 1$ — *высоким*. Отметим, что при обычно выполненном условии $N_E \gg N_B$ высокого уровня инжекции в эмиттере быть не может, так как концентрация неосновных носителей не может существенно превышать равновесную концентрацию основных носителей, равную N_E .

При низком уровне инжекции (НУИ) инжектированные через переход носители не изменяют существенно проводимость нейтральных областей.

Для низкого уровня инжекции из (II.3.5) получаем



$$\begin{cases} \Delta n_{pj} = n_{p0}(e^{V/\Phi_T} - 1), \\ \Delta p_{nj} = p_{n0}(e^{V/\Phi_T} - 1), \end{cases} \quad (\text{II.3.6a})$$

$$\begin{cases} \Delta n_{pj} = n_{p0}(e^{V/\Phi_T} - 1), \\ \Delta p_{nj} = p_{n0}(e^{V/\Phi_T} - 1), \end{cases} \quad (\text{II.3.6б})$$

или

$$\begin{cases} n_{pj} = n_{p0}e^{V/\Phi_T}, \\ p_{nj} = p_{n0}e^{V/\Phi_T}. \end{cases} \quad (\text{II.3.7a})$$

$$\begin{cases} n_{pj} = n_{p0}e^{V/\Phi_T}, \\ p_{nj} = p_{n0}e^{V/\Phi_T}. \end{cases} \quad (\text{II.3.7б})$$

Уравнения (II.3.6) и (II.3.7) получили название *граничных условий Шокли*. Заметим, что граничные условия Шокли могут быть получены непосредственно из соотношения (II.3.4а), учитывая, что при НУИ граничная концентрация основных носителей остается равновесной.

При НУИ избыточные граничные концентрации Δn_{pj} и Δp_{nj} пропорциональны равновесным концентрациям n_{p0} и p_{n0} . Если *n*-эмиттер легирован значительно сильнее *p*-базы, то $n_{p0} \gg p_{n0}$, и $\Delta n_{pj} \gg \Delta p_{nj}$. В этом случае *инжекция носит односторонний характер — из эмиттера в базу*.

Положительным напряжениям соответствует обогащение прилежащих к переходу областей неосновными носителями ($n_p > n_{p0}$, $p_n > p_{n0}$), что связано с преобладанием токов инжекции над токами экстракции. При отрицательных напряжениях преобладает процесс экстракция (отсасывания) неосновных носителей *p-n* переходом ($n_p < n_{p0}$, $p_n < p_{n0}$). При $V < -3\Phi_T$ практически можно считать, что на границе с переходом концентрация неосновных носителей равна нулю.

Как отмечалось выше, в эмиттере высокого уровня инжекции (ВУИ) быть не может. Для высокого уровня инжекции в *p*-базе из (II.3.5) получим

$$n_{pj} = n_i e^{V/2\Phi_T}. \quad (\text{II.3.8})$$

Граничное условие (II.3.8) также может быть получено непосредственно из соотношения (II.3.4а), учитывая что при ВУИ граничные концентрации основных и неосновных носителей одинаковы ($\Delta n_{pj} = \Delta p_{pj} \approx n_{pj} \approx p_{pj} \gg p_{p0}$).

II.3.4. Ширина *p-n* перехода

Формулы из параграфов II.2.4—II.2.6 могут быть использованы для расчета неравновесной ширины *p-n* перехода, если учесть изменение высоты контактного барьера под действием приложенного напряжения. Заменяя в (II.2.25) и (II.2.31) Φ_{BE} на $\Phi_{BE} - V$, получаем

$$l(V) = l_0 \sqrt[k]{(\varphi_{BE} - V) / \varphi_{BE}}, \quad (\text{II.3.9})$$

где $k = 2$ для ступенчатого перехода и диффузионного перехода в случае выполнения неравенства (II.2.43), $k = 3$ для линейного перехода и диффузионного перехода в случае выполнения неравенства (II.2.42). Таким образом, в ступенчатых p - n переходах зависимость ширины перехода от напряжения является более резкой, чем в линейных. Если ни одно из неравенств (II.2.42) и (II.2.43) в диффузионном переходе не выполнено, показатель степени корня в (II.3.9) принимает значения

$$2 < k < 3. \quad (\text{II.3.10})$$

Соотношение (II.3.9) показывает, что при обратном напряжении ($V < 0$) переход расширяется, а при прямом ($V > 0$) — сужается. Однако в последнем случае полученное выражение является скорее качественным, так как при прямом смещении сильно возрастает погрешность, связанная с приближением полного обеднения ОПЗ. По аналогии с (II.2.27) следует считать, что использование соотношения (II.3.9) допустимо при $2(\varphi_{BE} - V) \gg \varphi_T$. Практически его полагают справедливым при $V \lesssim \varphi_{BE} / 2$. При дальнейшем увеличении прямого напряжения ширина перехода стремится к дебаевской длине экранирования в базе.

Основные выводы

1. Дрейфовые и диффузионные токи пропорциональны концентрациям носителей заряда и градиентам квазиуровней Ферми.
2. Квазиуровни Ферми в ОПЗ постоянны. Расстояние между ними равно приложенному напряжению, умноженному на заряд электрона.
3. Приложение к p - n переходу внешнего напряжения изменяет высоту потенциального барьера, ширину перехода, а также граничные концентрации подвижных носителей заряда.
4. При положительном напряжении происходит инжекция неосновных носителей заряда через переход, при отрицательном напряжении — экстракция неосновных носителей заряда переходом.
5. При низком уровне инжекции концентрации неосновных носителей заряда на границах перехода описываются граничными условиями Шокли (II.3.6) и (II.3.7). При высоком уровне инжекции в базе справедливо граничное условие (II.3.8).
6. Зависимость ширины p - n перехода от напряжения при $V \lesssim \varphi_{BE} / 2$ описывается соотношением $l(V) = l_0 \sqrt[k]{(\varphi_{BE} - V) / \varphi_{BE}}$,



где $2 \leq k \leq 3$. Отрицательное напряжение приводит к расширению перехода, положительное — к сужению. В ступенчатых переходах зависимость ширины перехода от напряжения является более резкой, чем в линейных.

II.4. Анализ идеализированного диода

II.4.1. Модель идеализированного диода

Процессы, протекающие в полупроводниковом диоде при подаче на него внешнего напряжения, весьма сложны. Поэтому для количественного анализа вольтамперной характеристики (ВАХ) диода необходимо сделать некоторые упрощающие предположения. Наиболее простой является модель идеализированного диода, предложенная Шокли. Основные допущения, принимаемые в этой модели, поясняются рис. II.4.1.

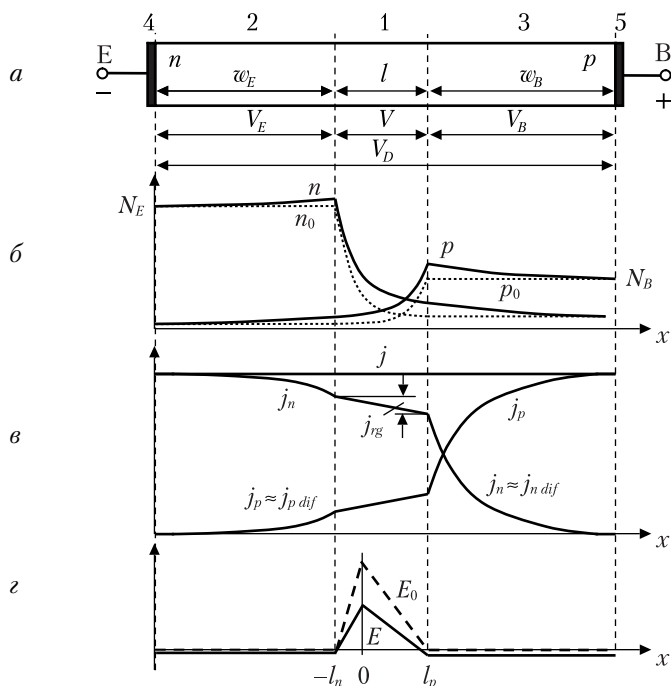


Рис. II.4.1. Структура полупроводникового диода (а) и распределения носителей заряда (б), плотностей токов (в) и напряженности электрического поля (г) при $V > 0$ (НУИ)

Структура диода схематически представлена на рис. II.4.1, *a*. В ней можно выделить области p - n перехода (1), эмиттера (2), базы (3), а также эмиттерного (4) и базового (5) контактов. На рис. II.4.1, *б* показаны распределения концентраций электронов и дырок в состоянии равновесия (n_0 и p_0), а также электронов и дырок при положительном напряжении на диоде (n и p). Как отмечалось выше, положительное напряжение вызывает инжекцию электронов из эмиттера в базу и дырок из базы в эмиттер, в результате чего концентрации неосновных носителей заряда в базе и эмиттере вблизи перехода увеличиваются. При этом нейтральность базы и эмиттера нарушается, и возникающее электрическое поле втягивает из контактов в базу дырки, а в эмиттер — электроны. В результате электрическая нейтральность базы и эмиттера восстанавливается практически полностью (имеет место «квазинейтральность» базы и эмиттера). Условия квазинейтральности базы и эмиттера имеют вид

$$\begin{cases} \Delta n_p(x) = n_p(x) - n_{p0} \approx \Delta p_p(x) = p_p(x) - p_{p0}, & (\text{II.4.1a}) \end{cases}$$

$$\begin{cases} \Delta p_n(x) = p_n(x) - p_{n0} \approx \Delta n_n(x) = n_n(x) - n_{n0}. & (\text{II.4.1б}) \end{cases}$$

На рис. II.4.1, *в*, *г* показана структура тока в диоде и изображено распределение электрического поля.

Приступая к составлению модели идеализированного диода, выскажем предварительно следующие соображения.

1. Если площадь поперечного сечения диода постоянна в направлении оси x , то полная плотность тока, равная сумме электронной и дырочной составляющих, не зависит от x (см. рис. II.4.1, *в*):

$$j = j_p(x) + j_n(x) = \text{const.} \quad (\text{II.4.2})$$

2. Механизмами переноса электронов и дырок являются диффузия и дрейф

$$\begin{cases} j_p(x) = -eD_p dp / dx + ep\mu_p E, & (\text{II.4.3a}) \end{cases}$$

$$\begin{cases} j_n(x) = eD_n dn / dx + en\mu_n E. & (\text{II.4.3б}) \end{cases}$$

Если эмиттер и база легированы однородно ($N_E(x) = \text{const}$, $N_B(x) = \text{const}$), то в состоянии равновесия напряженность поля E_0 вне перехода равна нулю. При положительном напряжении на диоде поле вне перехода связано с омическими сопротивлениями областей эмиттера и базы и направлено против оси x (см. рис. II.4.1, *г*). Поскольку удельное сопротивление квазинейтральных областей эмиттера и базы весьма

малы по сравнению с удельным сопротивлением перехода, поля в эмиттере и базе малы.

3. Если приложенное к диоду напряжение не слишком велико, концентрации неосновных носителей в базе и эмиттере во много раз меньше, чем основных (низкий уровень инжекции). Дрейфовые токи неосновных носителей при этом много меньше, чем дрейфовые токи основных носителей и, следовательно, много меньше полного тока через диод. В то же время ввиду условий квазинейтральности (II.4.1) градиенты концентраций основных и неосновных носителей в каждой точке эмиттера и базы одинаковы, и диффузионные токи имеют одинаковый порядок величины. Поэтому токи неосновных носителей заряда в квазинейтральных областях могут считаться чисто диффузионными (см. рис. II.4.1, *в*):

$$\begin{cases} j_p \approx j_{p\text{dif}} = -eD_p dp / dx, & -l_n < x < 0, \end{cases} \quad (\text{II.4.4a})$$

$$\begin{cases} j_n \approx j_{n\text{dif}} = eD_n dn / dx, & l_p > x > 0. \end{cases} \quad (\text{II.4.4б})$$

4. Приложенное к диоду напряжение V_D складывается из падений напряжения на *p-n* переходе V и на квазинейтральных областях эмиттера и базы (см. рис. II.4.1, *а*):

$$V_D = V + V_E + V_B. \quad (\text{II.4.5})$$

Если ток через диод не слишком велик, основная часть V полного напряжения V_D падает на переходе, так как области базы и эмиттера имеют значительно меньшее удельное сопротивление.

5. Изменения дырочного и электронного токов в области перехода могут быть связаны только с нарушением равновесия процессов термогенерации и рекомбинации носителей заряда в переходе. Поскольку электроны и дырки рекомбинируют и генерируются парами, должны выполняться равенства

$$\begin{cases} j_n(-l_n) - j_n(l_p) = j_{gr}, \end{cases} \quad (\text{II.4.6a})$$

$$\begin{cases} j_p(l_p) - j_p(-l_n) = j_{gr}, \end{cases} \quad (\text{II.4.6б})$$

где j_{gr} — плотность тока, связанного с генерацией и рекомбинацией электронно-дырочных пар в переходе (см. рис. II.4.1, *в*). Если ширина перехода l достаточно мала, ток j_{gr} много меньше полного тока через диод.

С учетом изложенных соображений модель *идеализированного диода* основана на следующих допущениях.

1. Поперечные размеры диода много больше продольных (в направлении оси x), в результате чего задача может считаться одномерной (величины p, n, j_p, j_n зависят только от координаты x).

2. Эмиттер и база легированы однородно.

3. Падением напряжения на квазинейтральных областях базы и эмиттера можно пренебречь: $V_D = V$.

4. Приложенное к переходу напряжение не слишком велико, так что уровень инжекции в базе (и, тем более, в эмиттере) можно считать низким

$$\begin{cases} n_p \ll p_p \approx p_{p0} = N_B, \\ p_n \ll n_n \approx n_{n0} = N_E. \end{cases} \quad (\text{II.4.7a})$$

$$\begin{cases} n_p \ll p_p \approx p_{p0} = N_B, \\ p_n \ll n_n \approx n_{n0} = N_E. \end{cases} \quad (\text{II.4.7б})$$

5. При низком уровне инжекции дрейфовые токи неосновных носителей малы; при этом выполняются приближенные равенства (II.4.4), а напряженности электрического поля в квазинейтральных областях базы и эмиттера можно считать равными нулю.

6. Токами генерации и рекомбинации носителей заряда в области перехода можно пренебречь.

Распределения носителей заряда и плотностей токов при $V > 0$, принятые в идеализированной модели диода, представлены на рис. II.4.2. Стационарные характеристики диода будут вначале проанализированы в рамках идеализированной модели, а затем мы рассмотрим, к чему приводят отклонения от принятых допущений.

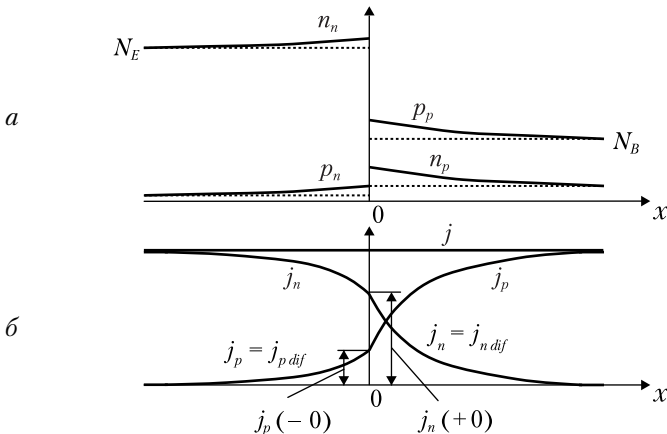


Рис. II.4.2. Распределения носителей заряда (а) и плотностей токов (б) в идеализированной модели диода

II.4.2. Методика анализа ВАХ идеализированного диода

Поскольку полный ток диода в любом сечении имеет одну и ту же величину (II.4.2), для нахождения ВАХ диода необходимо вычислить в соответствии с уравнениями (II.4.3) токи j_p и j_n при любом значении x в зависимости от приложенного напряжения V . Для этого в свою очередь следует знать распределения концентраций носителей заряда $n(x)$, $p(x)$ и электрического поля.

Функции $n(x)$, $p(x)$ при низком уровне инжекции в базе и эмиттере (допущение 5) определяются решением биполярных уравнений непрерывности (I.7.8а, б) в квазинейтральных областях базы и эмиттера, где ввиду допущения 5 члены, содержащие электрическое поле, могут быть опущены:

$$\begin{cases} \frac{\partial n_p}{\partial t} = -\frac{\Delta n_p}{\tau_n} + D_n \frac{\partial^2 n_p}{\partial x^2}, & \text{(II.4.8a)} \end{cases}$$

$$\begin{cases} \frac{\partial p_n}{\partial t} = -\frac{\Delta p_n}{\tau_p} + D_p \frac{\partial^2 p_n}{\partial x^2}. & \text{(II.4.8б)} \end{cases}$$

Допущения, принятые в идеализированной модели диода, позволяют применить для анализа его характеристик следующую методику.

1. Полный ток определяется в сечении $x=0$ как сумма тока электронов в p -области на границе с переходом $j_n(+0)$ и тока дырок в n -области на границе с переходом $j_p(-0)$ (см. рис. II.4.2, б):

$$j = j_n(+0) + j_p(-0). \quad \text{(II.4.9)}$$

Справедливость равенства (II.4.9) является следствием допущения 6 параграфа II.4.1 о малости токов генерации-рекомбинации в переходе.

2. Поскольку оба тока $j_n(+0)$ и $j_p(-0)$ переносятся неосновными носителями, они являются чисто диффузионными (допущение 5) и определяются соотношениями:

$$\begin{cases} j_n(+0) = -eD_n \frac{\partial n_p}{\partial x} (+0), & \text{(II.4.10a)} \end{cases}$$

$$\begin{cases} j_p(-0) = eD_p \frac{\partial p_n}{\partial x} (-0). & \text{(II.4.10б)} \end{cases}$$

3. Градиенты концентрации неосновных носителей заряда на границах перехода определяются решением уравнений (II.4.8).



4. При решении уравнений (II.4.8) используются граничные условия Шокли (II.3.6), (II.3.7), справедливые при низком уровне инжекции (допущение 4). Вследствие допущения 3 напряжение на переходе совпадает с полным напряжением на диоде

$$V_D = V.$$

Главное преимущество изложенной методики состоит в том, что отпадает необходимость в определении напряженности электрического поля и, следовательно, в использовании уравнения Пуассона. Заметим, что электрическое поле играет существенную роль в переносе основных носителей (например, вдали от перехода ток переносится основными носителями (см. рис. II.4.1, в) и является чисто дрейфовым). Для наших целей необходимо лишь определить плотности токов неосновных носителей по обе стороны ОПЗ (плоскости $x = \pm 0$). Это весьма полезное упрощение возможно благодаря допущению 6 о малости токов генерации и рекомбинации носителей заряда в области перехода.

II.4.3. Распределения неосновных носителей заряда в квазинейтральных областях

Распределение неосновных носителей (электронов) в p -базе описывается уравнением (II.4.8a), общее решение которого в стационарном случае ($\partial n_p / \partial t = 0$) имеет вид

$$\Delta n_p = A_1 e^{x/L_B} + A_2 e^{-x/L_B}, \quad (\text{II.4.11})$$

где
$$L_B = \sqrt{D_B \tau_B} \quad (\text{II.4.12})$$

— *диффузионная длина* неосновных носителей (электронов) в базе (среднее расстояние, которое электроны преодолевают под действием диффузии за время жизни), $D_B = D_n$ и $\tau_B = \tau_n$ — коэффициент диффузии и время жизни неосновных носителей (электронов) в базе. Индексы « B » вместо « p » использованы для того, чтобы конечный результат не зависел от типа проводимости базы. Следует помнить, что индексы « B » (и введенные ниже индексы « E ») относятся к *неосновным* носителям в базе (эмиттере).

Константы интегрирования A_1 и A_2 определяются граничными условиями в плоскостях $x = 0$ (граница базы с p - n переходом) и $x = w_B$ (граница базы с контактом).



В плоскости $x = 0$ для низкого уровня инжекции справедливы граничные условия Шокли (II.3.6)

$$\Delta n_p(0) = \Delta n_{pj} = n_{p0}(e^{V/\Phi_T} - 1); \quad (\text{II.4.13a})$$

$$\Delta p_n(0) = \Delta p_{nj} = p_{n0}(e^{V/\Phi_T} - 1). \quad (\text{II.4.13б})$$

На границе с омическим контактом концентрации неосновных носителей соответствуют равновесным значениям (см. параграф I.5.3)

$$\begin{cases} \Delta n_p(w_B) = 0; \\ \Delta p_n(w_E) = 0. \end{cases} \quad (\text{II.4.14a})$$

$$\begin{cases} \Delta n_p(w_B) = 0; \\ \Delta p_n(w_E) = 0. \end{cases} \quad (\text{II.4.14б})$$

Соотношения (II.4.14) соответствуют *определению омического контакта*. В реальных приборах условия (II.4.14) не являются вполне точными, однако это не влияет существенно на результат.

С учетом (II.4.13) и (II.4.14) из (II.4.11) получаем

$$\begin{cases} \Delta n_p(x) = \Delta n_p(0) \cdot \frac{\text{sh}[(w_B - x) / L_B]}{\text{sh}(w_B / L_B)}, \end{cases} \quad (\text{II.4.15a})$$

$$\begin{cases} n_p(x) = n_{p0} + \Delta n_p = n_{p0} + \Delta n_p(0) \cdot \frac{\text{sh}[(w_B - x) / L_B]}{\text{sh}(w_B / L_B)}. \end{cases} \quad (\text{II.4.15б})$$

Распределение электронов в базе для различных напряжений на переходе представлено на рис. II.4.3.

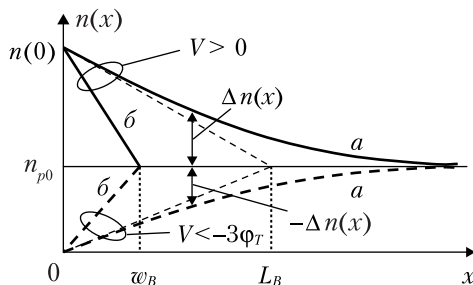


Рис. II.4.3. Распределение неосновных носителей заряда в базе диода:
 a — длинная база ($w_B \gg L_B$); $б$ — короткая база ($w_B \ll L_B$)

Для диода с длинной базой ($w_B \gg L_B$, рис. II.4.3, a) из (II.4.15a) следует

$$\Delta n_p(x) = n_p(0)e^{-x/L_B}. \quad (\text{II.4.16a})$$

Таким образом, на диффузионной длине избыточная концентрация носителей уменьшается вследствие рекомбинации в $e = 2,72$ раз. Заметим, что касательная, проведенная к кривой $n_p(x)$ в точке $x = 0$, пересекает линию n_{p0} в точке $x = L_B$ (см. рис. II.4.3).

В диоде с *короткой базой* ($w_B \ll L_B$, см. рис. II.4.3, б), из (II.4.15a) следует

$$\Delta n_p(x) = n_p(0) \cdot (w_B - x) / w_B. \quad (\text{II.4.16б})$$

Таким образом, в диоде с короткой базой распределение концентрации электронов есть линейная функция x . Это объясняется тем, что электроны не успевают рекомбинировать в базе за время, в течение которого они диффундируют от p - n перехода к омическому контакту. Поскольку потерь на рекомбинацию нет, электронный ток не зависит от x . Так как он является диффузионным (допущение 5), то от x не зависит также и градиент концентрации электронов, что и соответствует их линейному распределению.

Аналогичным образом находится распределение избыточных дырок в эмиттере

$$\Delta p_n(x) = p_n(0) \cdot \frac{\text{sh}[(w_E - x') / L_E]}{\text{sh}(w_E / L_E)}, \quad (\text{II.4.17})$$

где $\Delta p_n(x) = p_{n0}(e^{V/\varphi_T} - 1)$,

$x' = -x$, $L_E = \sqrt{D_E \tau_E} = \sqrt{D_p \tau_p}$ — диффузионная длина неосновных носителей (дырок) в n -эмиттере.

Графическое представление распределения неосновных носителей в квазинейтральных областях дает наглядное представление о многих электрических свойствах прибора. Отметим, что концентрация неосновных носителей на границе с ОПЗ ($x = 0$ на рис. II.4.3) соответствует напряжению на p - n переходе (граничное условие (II.4.13)), а тангенс угла наклона касательной к функции $n_p(x)$ (производная dn_p/dx) пропорционален электронному току, который имеет диффузионный механизм (соотношение (II.4.4)). Так, например, рис. II.4.3 показывает, что кривые a и b для случая $V > 0$ построены для одинаковых напряжений. При этом в случае короткой базы электронный ток больше, чем в случае длинной базы (кривая $n_p(x)$ имеет больший наклон).

Также отметим, что граничные концентрации $n_p(0)$ и $p_n(0)$ практически равны нулю уже при малых по абсолютной величине отрицательных напряжениях (при $V < -3\phi_T \approx 80$ мВ $n_p(0)/n_{p0} = p_n(0)/p_{n0} < 1/20$). Поэтому при $V < -3\phi_T$ вид функций $n_p(x)$, $p_n(x)$ и, следовательно, величина тока не зависят от напряжения.

II.4.4. ВАХ идеализированного диода

Дифференцируя соотношения (II.4.15), (II.4.17) и подставляя результат в (II.4.10) с учетом (II.4.13), получаем

$$\begin{cases} j_n(+0) = j_{Sn}(e^{V/\phi_T} - 1), & \text{(II.4.18a)} \\ j_p(-0) = j_{Sp}(e^{V/\phi_T} - 1), & \text{(II.4.18б)} \end{cases}$$

$$\text{где } j_{Sn} = \frac{eD_B n_{p0}}{L_B \text{th}(w_B/L_B)} = \begin{cases} eD_B n_{p0}/L_B, & w_B \gg L_B, \\ eD_B n_{p0}/w_B, & w_B \ll L_B; \end{cases} \quad \text{(II.4.19a)}$$

$$j_{Sp} = \frac{eD_E p_{n0}}{L_E \text{th}(w_E/L_E)} = \begin{cases} eD_E p_{n0}/L_E, & w_E \gg L_E, \\ eD_E p_{n0}/w_E, & w_E \ll L_E. \end{cases} \quad \text{(II.4.19б)}$$

Величины j_{Sn} и j_{Sp} называются плотностями *тепловых токов* электронов и дырок, соответственно.

С учетом (II.4.9) ВАХ идеализированного диода имеет вид

$$I = I_S(e^{V/\phi_T} - 1), \quad \text{(II.4.20)}$$

$$\text{где} \quad I_S = I_{Sn} + I_{Sp} \quad \text{(II.4.21)}$$

— *тепловой ток диода*;

$$I_{Sn} = S j_{Sn} \quad \text{(II.4.22a)}$$

— *тепловой ток электронов*;

$$I_{Sp} = S j_{Sp} \quad \text{(II.4.22б)}$$

— *тепловой ток дырок*; S — площадь перехода.

Тепловой ток является единственным параметром, полностью определяющим ВАХ идеализированного диода.

Поскольку $n_{p0} = n_i^2/p_{p0} = n_i^2/N_B$ и $p_{n0} = n_i^2/n_{n0} = n_i^2/N_{Eef}$ (допущение 7), при одинаковых условиях тепловой ток в германиевом диоде ($n_i \approx 10^{13}$ см⁻³ при $T = 300$ К) намного больше, чем в кремниевом ($n_i \approx 10^{10}$ см⁻³ при $T = 300$ К).



Из вольтамперной характеристики (II.4.20) следует, что при $V < -3\phi_T$ обратный ток через диод не зависит от напряжения и равен тепловому:

$$I|_{V < -3\phi_T} = -I_S.$$

Примерный вид ВАХ идеализированных диодов показан на рис. II.4.4. Из рисунка очевидно, что ВАХ диода имеет резко асимметричный характер и соответствует качественным результатам параграфа II.2.2.

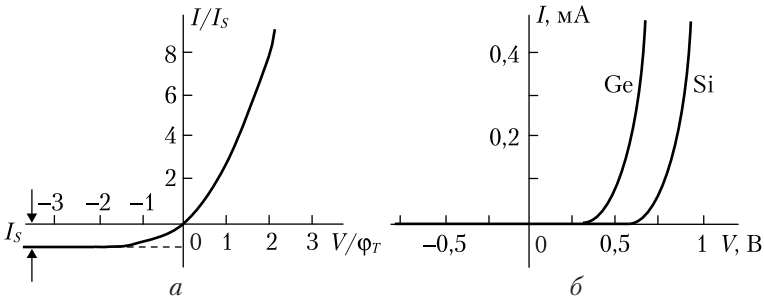


Рис. II.4.4. ВАХ идеализированного диода:
а — крупный масштаб; б — обычный масштаб

В заключение отметим, что согласно уравнениям (II.4.10), (II.4.13), (II.4.16) и (II.4.17) *электронный и дырочный токи на границах ОПЗ пропорциональны избыточным концентрациям соответствующих носителей заряда*. Поэтому экспоненциальная форма ВАХ (II.4.20) непосредственно вытекает из граничных условий Шокли (II.4.13).

II.4.5. Тепловой ток

В диоде с *длинной базой* ($w_B \gg L_B$) из соотношений (II.4.19а), (II.4.22а) получаем

$$I_{Sn} = eSD_B(n_{p0} / L_B), \quad (\text{II.4.23})$$

или с учетом (II.4.12)

$$I_{Sn} = e(n_{p0} / \tau_B)SL_B = eg_BV_B, \quad (\text{II.4.24})$$

где согласно (I.5.2) $g_B = n_{p0} / \tau_B$ — скорость термогенерации электронов в базе; $V_{Bef} = SL_B$ — эффективный объем области базы длиной L_B . Физический смысл выражений (II.4.23)

и (II.4.24) ясен из рис. II.4.3, на котором штриховая кривая *a* изображает распределение электронов в базе при $V < -3\phi_T$ (при этом $I = -I_S$, $I_n(0) = -I_{Sn}$).

Последний сомножитель в (II.4.23) имеет смысл градиента концентрации электронов в базе на границе с *p-n* переходом; таким образом, формула (II.4.23) определяет диффузионный ток электронов в плоскости $x=0$ при $V < -3\phi_T$.

Выражение (II.4.24) показывает, что тепловой ток соответствует собиранию (экстракции) *p-n* переходом неосновных носителей, генерируемых со скоростью g_B в объеме $V_{Bef} = SL_B$. С расстояния, большего L_B , неосновные носители не успевают диффундировать до *p-n* перехода за время жизни и не дают вклада в тепловой ток.

В случае *короткой* базы ($w_B \ll L_B$) такая интерпретация теплового тока несправедлива, так как вследствие малого объема базы тепловая генерация электронов незначительна. Ток I_{Sn} обусловлен инжекцией электронов из контакта базы. С уменьшением толщины базы тепловой ток возрастает. Из соотношений (II.4.19 а), (II.4.22 а) получаем

$$I_{Sn} = eSD_B(n_{p0} / w_B), \quad (\text{II.4.25})$$

Из рис. II.4.3 следует, что последний сомножитель в (II.4.25) имеет смысл градиента концентрации электронов на границе с *p-n* переходом, и формула (II.4.25) определяет диффузионный ток электронов в плоскости $x=0$ при $V < -3\phi_T$.

Практически база может считаться короткой при $w_B / L_B < 0,75$ и длинной при $w_B / L_B > 1,5$.

Для численных расчетов удобно в соотношениях (II.4.19), (II.4.23) и (II.4.25) выразить равновесную концентрацию неосновных носителей через результирующую концентрацию примеси:

$$\begin{cases} n_{p0} = n_i^2 / N_B, \\ p_{n0} = n_i^2 / N_{Eef}. \end{cases}$$

При этом формулы для тепловых токов становятся инвариантными относительно типа проводимости базы и эмиттера

$$I_{SB} = eSn_i^2 / G_B; \quad (\text{II.4.26a})$$

$$I_{SE} = eSn_i^2 / G_E, \quad (\text{II.4.26б})$$

где I_{SB} и I_{SE} — тепловые токи неосновных носителей в базе и эмиттере (для *p*-базы и *n*-эмиттера $I_{SB} = I_{Sn}$, $I_{SE} = I_{Sp}$),

$$G_{B,E} = \frac{N_{B,Eef} w_{B,E}}{D_{B,E}} \cdot \frac{\text{th}(w_{B,E} / L_{B,E})}{w_{B,E} / L_{B,E}}. \quad (\text{II.4.27})$$

— числа Гуммеля в базе и эмиттере. Для предельных случаев $w_{B,E} \ll L_{B,E}$ и $w_{B,E} \gg L_{B,E}$ из (II.4.27) получаем:

$$G_{B,E} = \begin{cases} N_{B,Eef} w_{B,E} / D_{B,E}, & w_{B,E} \ll L_{B,E}, \\ N_{B,Eef} L_{B,E} / D_{B,E}, & w_{B,E} \gg L_{B,E}. \end{cases} \quad (\text{II.4.28a})$$

$$(\text{II.4.28б})$$

Как очевидно из формул (II.4.27), (II.4.28), числа Гуммеля в базе и эмиттере пропорциональны поверхностным (на единицу площади поверхности) концентрациям примеси в соответствующих слоях с толщиной, которая определяется меньшим из значений длины слоя и диффузионной длины.

В большинстве практических случаев выполнено неравенство $N_{Eef} \gg N_B$. При этом $I_{SE} \ll I_{SB} \approx I_S$. Таким образом, *тепловой ток в основном определяется свойствами базы*.

Соотношения (II.4.27) и (II.4.28) показывают, что тепловые токи возрастают при снижении концентрации примеси в базе.

Важным свойством теплового тока является *чрезвычайно сильная температурная зависимость*, связанная в основном с резкой зависимостью от температуры концентрации неосновных носителей n_i . Из соотношений (II.4.26а, б) получаем

$$I_S = I_{SB} + I_{SE} = e S n_i^2 (1 / G_B + 1 / G_E), \quad (\text{II.4.29})$$

$$\text{где} \quad n_i^2(T) = N_c N_v e^{-E_g / kT}, \quad (\text{II.4.30})$$

N_c и N_v — эффективные плотности квантовых состояний в зоне проводимости и валентной зоне, сравнительно слабо зависящие от температуры ($N_c N_v \sim T^3$). Пренебрегая их температурной зависимостью, из (II.4.29) и (II.4.30) получаем

$$I_S \approx I_S(T_0) \exp \left[\frac{E_g}{k} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right], \quad (\text{II.4.31})$$

где $T_0 = 300$ К — номинальная температура окружающей среды. Для температурного диапазона работы полупроводниковых приборов ($-60^\circ\text{C} - +125^\circ\text{C}$) можно с хорошей точностью полагать: $1/T_0 - 1/T \approx \Delta T / T_0^2$, где $\Delta T = T - T_0$. При этом из (II.4.31)

$$I_s(T) \approx I_s(T_0)e^{\Delta T/T_s} = I_s(T_0)e^{\Delta T/T_2}, \quad (\text{II.4.32})$$

$$\text{где} \quad T_s = kT_0^2 / E_g, \quad (\text{II.4.33})$$

$$T_2 = T_s \ln 2 = kT_0^2 \ln 2 / E_g \quad (\text{II.4.34})$$

— *температура удвоения теплового тока* (при повышении температуры на величину T_2 тепловой ток удваивается). Как очевидно из соотношений (II.4.29), (II.4.30) и (II.4.33), при увеличении ширины запрещенной зоны полупроводника E_g величина теплового тока снижается, однако его температурная зависимость возрастает (снижается температура удвоения теплового тока T_2). Для наиболее часто применяемых полупроводников температура удвоения теплового тока составляет

$$T_2 = \begin{cases} 10^\circ\text{C}, & \text{Ge } (E_g = 0,66\text{эВ}), \\ 6,5^\circ\text{C}, & \text{Si } (E_g = 1,12\text{эВ}), \\ 5^\circ\text{C}, & \text{GaAs } (E_g = 1,43\text{эВ}). \end{cases}$$

II.4.6. Температурная зависимость прямой ветви ВАХ диода

При обратном смещении температурная зависимость ВАХ идеализированного диода сводится к температурной зависимости теплового тока, рассмотренной в предыдущем параграфе. При прямом смещении температурную зависимость ВАХ удобно характеризовать температурным коэффициентом напряжения (ТКН)

$$\text{ТКН} = \left. \frac{\partial V}{\partial T} \right|_{dI=0}, \quad (\text{II.4.35})$$

численно равному изменению прямого напряжения на диоде при изменении температуры на 1 градус при постоянном токе диода.

При $V > 3\varphi_T$ ($I \gg I_s$) соотношение (II.4.20) с учетом (II.4.31) дает

$$V = \varphi_T \ln \frac{I}{I_s} = \varphi_T \ln \left(\frac{I}{I_{s0}} e^{E_g/kT} \right) = \frac{E_g}{e} + \frac{kT}{e} \ln \frac{I}{I_{s0}}.$$

Дифференцируя по температуре, получаем



$$\text{ТКН} = \frac{V - V_g}{T}, \quad (\text{II.4.36})$$

где $V_g = E_g / e$ — ширина запрещенной зоны в вольтах.

Для невырожденных полупроводников $V < \Phi_{BE} < V_g$, и величина ТКН отрицательна (рис. II.4.5), поэтому при увеличении температуры прямая ветвь ВАХ диода сдвигается в сторону меньших напряжений. Для наиболее часто применяемых полупроводников значение ТКН при $T = 300 \text{ К}$ и $V = 2V_g / 3$ составляет

$$\text{ТКН} \approx \begin{cases} -0,8 \text{ мВ} / ^\circ\text{C}, & \text{Ge} (V_g = 0,66 \text{ В}), \\ -1,24 \text{ мВ} / ^\circ\text{C}, & \text{Si} (V_g = 1,12 \text{ В}), \\ -1,5 \text{ мВ} / ^\circ\text{C}, & \text{GaAs} (V_g = 1,43 \text{ В}). \end{cases}$$

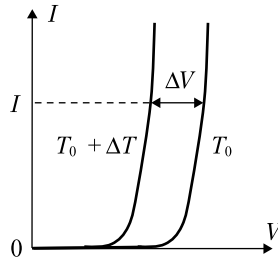


Рис. II.4.5. Изменение прямой ветви ВАХ идеализированного диода при изменении температуры ($\Delta T > 0$)

II.4.7. Характеристические сопротивления диода

Основным показателем качества ВАХ диода является нелинейность его характеристики, позволяющая осуществлять выпрямление электрических сигналов, ограничение их по амплитуде, преобразования частотного спектра и другие нелинейные операции.

Нелинейность ВАХ проявляется в различии *дифференциального* сопротивления $r = dV / dI$ и *статического* сопротивления $R = V / I$. Из ВАХ (II.4.20) находим

$$r = \frac{\Phi_T}{I + I_S}; \quad (\text{II.4.37})$$

$$R = \frac{\Phi_T}{I} \ln(1 + I / I_S). \quad (\text{II.4.38})$$



Очевидно, что в состоянии равновесия ($V = 0$) дифференциальное и статическое сопротивления совпадают.

На прямой ветви ВАХ ($V > 0$) дифференциальное сопротивление меньше статического. При $I \gg I_S$ дифференциальное сопротивление зависит только от тока и температуры, причем эта зависимость одинакова для всех диодов:

$$r = \varphi_T / I. \quad (\text{II.4.39})$$

На обратной ветви ВАХ ($V < 0$) дифференциальное сопротивление больше статического. При $V < -3\varphi_T$ величина дифференциального сопротивления может считаться бесконечно большой, так как $I \approx -I_S$. Статическое сопротивление при этом пропорционально обратному напряжению:

$$R = -V / I_S. \quad (\text{II.4.40})$$

II.4.8. Коэффициенты инжекции и эффективность эмиттера

В плоскости *p-n* перехода ($x = \pm 0$) ток диода имеет электронную и дырочную составляющие (II.4.9). Отношения

$$\left\{ \begin{aligned} \gamma_n &= \frac{I_n(0)}{I_p(0) + I_n(0)} = \frac{1}{1 + I_p(0) / I_n(0)}, \end{aligned} \right. \quad (\text{II.4.41a})$$

$$\left\{ \begin{aligned} \gamma_p &= \frac{I_p(0)}{I_p(0) + I_n(0)} = \frac{1}{1 + I_n(0) / I_p(0)}, \end{aligned} \right. \quad (\text{II.4.41б})$$

показывающие, какие части тока через переход переносятся электронами и дырками, называются *коэффициентами инжекции электронов и дырок*, соответственно. Очевидно, что сумма коэффициентов инжекции электронов и дырок равна 1

$$\gamma_n + \gamma_p = 1. \quad (\text{II.4.42})$$

Коэффициент инжекции носителей, инжектируемых из эмиттера в базу, называется *эффективностью эмиттера*. В случае эмиттера *n*-типа эффективность эмиттера равна коэффициенту инжекции электронов.

Величина эффективности эмиттера не является, как правило, важным параметром полупроводникового диода, однако в биполярном транзисторе имеет решающее значение. Учитывая, что для *p*-базы и *n*-эмиттера $I_{SB} = I_{Sn}$, $I_{SE} = I_{Sp}$, а для *n*-базы и *p*-эмиттера $I_{SB} = I_{Sp}$, $I_{SE} = I_{Sn}$, с учетом (II.4.18а, б) и (II.4.26а, б) получаем



$$\gamma_n = \begin{cases} 1 / (1 + G_B / G_E) & \text{для } p\text{-базы,} \\ 1 / (1 + G_E / G_B) & \text{для } n\text{-базы;} \end{cases}$$

$$\gamma_p = \begin{cases} 1 / (1 + G_E / G_B) & \text{для } p\text{-базы,} \\ 1 / (1 + G_B / G_E) & \text{для } n\text{-базы.} \end{cases}$$

Для эффективности эмиттера всегда справедливо соотношение

$$\gamma = \frac{1}{1 + G_B / G_E}. \quad (\text{II.4.43})$$

В практически важном случае $N_E \gg N_B$, $G_E \gg G_B$ эффективность эмиттера близка к 1, и ВАХ определяется свойствами базы.

Основные выводы

1. Анализ идеализированного диода сводится к решению биполярных уравнений непрерывности в базе и эмиттере без учета электрического поля.

2. ВАХ идеализированного диода описывается уравнением $I = I_S (e^{V/\Phi_T} - 1)$ и полностью определяется единственным параметром — тепловым током I_S .

3. При обратном смещении температурная зависимость ВАХ определяется температурной зависимостью теплового тока и характеризуется температурой T_S или температурой удвоения T_2 , которые определяются соотношениями (II.4.33) и (II.4.34).

4. При прямом смещении температурная зависимость ВАХ характеризуется температурным коэффициентом напряжения (ТКН), величина которого отрицательна.

5. Выпрямляющие свойства диода проявляются в различии дифференциального и статического сопротивлений. При прямом смещении ($I \gg I_S$) дифференциальное сопротивление диода обратно пропорционально току: $r = \Phi_T / I$; при обратном смещении ($I \approx -I_S$) дифференциальное сопротивление диода бесконечно велико.

6. Соотношение дырочного и электронного токов в плоскости перехода определяется эффективностью эмиттера. Эффективность эмиттера увеличивается с увеличением толщины и степени легирования эмиттера и уменьшением толщины и степени легирования базы.

7. В случае, когда эмиттер легирован значительно сильнее базы, эффективность эмиттера близка к 1, инжекция неосновных носителей заряда имеет односторонний характер (в базу из эмиттера), и ВАХ определяется свойствами базы.



II.5. ВАХ реального диода

II.5.1. Особенности ВАХ реального диода

Идеализированная модель диода основана на целом ряде физических допущений, перечисленных в параграфе II.4.1. При работе реального диода в широком диапазоне токов эти допущения могут не выполняться, что приводит к расхождению ВАХ реального диода и ВАХ, описываемой выражением (II.4.20). В области больших токов должны проявляться эффекты, связанные с падением напряжения на сопротивлении базы и эмиттера, а также с повышением уровня инжекции. При высоком уровне инжекции перестают выполняться граничные условия Шокли, определяющие форму ВАХ. Кроме того, при высоком уровне инжекции недопустимо пренебрежение дрейфовым током неосновных носителей заряда, поскольку их концентрация не может считаться малой. Наконец, часть тока в реальном диоде связана с термогенерацией и рекомбинацией электронно-дырочных пар в переходе. Этот ток не был учтен в идеализированной модели диода.

В разделе II.5 будут рассмотрены указанные выше эффекты, выяснены связанные с ними изменения ВАХ диода и определены условия, при которых учет того или иного эффекта становится необходимым.

II.5.2. Термогенерация и рекомбинация носителей заряда в p - n переходе

При выводе ВАХ идеализированного диода мы пренебрегли термогенерацией и рекомбинацией носителей в области p - n перехода, считая, что инжектированные составляющие дырочного и электронного токов одинаковы по обе стороны перехода.

Носители заряда, генерированные в ОПЗ, подхватываются сильным электрическим полем и выбрасываются за пределы перехода: электроны — в n -область, дырки — в p -область. Таким образом, ток термогенерации направлен от n -области к p -области.

Носители заряда, попадающие в ОПЗ из тех областей, где они являются основными, имеют вероятность рекомбинировать в области перехода. Поскольку электроны попадают в ОПЗ со стороны n -области, а дырки — со стороны p -области, с процессом рекомбинации в переходе связан ток, протекающий



от p -области к n -области. При выводе ВАХ мы вычисляли диффузионный ток носителей после их прохождения через ОПЗ. Таким образом, ток, связанный с рекомбинацией в области перехода, учтен не был.

В равновесном состоянии токи термогенерации и рекомбинации в p - n переходе взаимно уравниваются.

При прямом смещении перехода инжекционные потоки носителей резко возрастают, соответственно возрастают концентрации носителей заряда в ОПЗ и ток, связанный с их рекомбинацией. Ток термогенерации, напротив, уменьшается, так как скорость генерации зависит только от температуры, а объем p - n перехода сокращается вследствие уменьшения его ширины. Таким образом, ток рекомбинации увеличивает общий прямой ток диода.

При обратном смещении перехода инжекционные потоки носителей практически прекращаются, соответственно прекращается и ток рекомбинации в переходе. С увеличением обратного смещения ток термогенерации несколько возрастает из-за увеличения объема p - n перехода вследствие его расширения. Таким образом, ток термогенерации увеличивает общий обратный ток диода и делает его зависящим от напряжения.

Как показано в параграфе I.5.4, в моноатомных полупроводниках (Si, Ge) преобладающим является ловушечный механизм рекомбинации Шокли—Рида—Холла. При этом избыточная скорость рекомбинации определяется соотношением (I.5.4)

$$r - g = \frac{np - n_i^2}{\tau_{n0}(n + n_1) + \tau_{p0}(p + p_1)},$$

где $\tau_{n0}, \tau_{p0} = 1 / N_t C_{p,n}$, $C_{n,p}$ — вероятности захвата ловушками электронов и дырок, $n_1 = n_i e^{(E_t - E_i)/kT}$, $p_1 = n_i e^{(E_i - E_t)/kT}$, E_t — энергетический уровень ловушек. Оставаясь качественно верными, результаты значительно упрощаются, если положить

$$C_n = C_p; E_t = E_i.$$

При этом $\tau_{n0} = \tau_{p0} = \tau_0$, $n_1 = p_1 = n_i$, и для избыточной скорости рекомбинации получаем

$$r - g = \frac{np - n_i^2}{\tau_0(n + p + 2n_i)}.$$

Согласно (II.3.2) в ОПЗ $n(x)p(x) = n_i^2 e^{V/\Phi_T}$. Таким образом,



$$r - g = \frac{n_i^2 e^{V/\Phi_T} - n_i^2}{\tau_0(n + p + 2n_i)}. \quad (\text{II.5.1})$$

Поскольку каждая рекомбинирующая в переходе пара вызывает прохождение через ОПЗ заряда e в направлении от p - к n -области, а каждая генерированная пара — прохождение того же заряда в обратном направлении, ток генерации-рекомбинации может быть представлен в виде

$$I_{rg} = eS \int_{-l_p}^{l_n} (r - g) dx. \quad (\text{II.5.2})$$

II.5.3. Ток термогенерации (обратное смещение перехода)

При обратном напряжении $V < -3\Phi_T$ инжекционные потоки электронов и дырок через ОПЗ практически прекращаются и в области перехода выполняются неравенства: $p \ll n_i$, $p \ll n_i$. При этом процессы рекомбинации в ОПЗ отсутствуют (способных рекомбинировать носителей практически нет, $r \ll g$) и из (II.5.1), (II.5.2) получаем

$$r - g = -g = -n_i / 2\tau_0;$$

$$I_{rg} = -I_g = -en_i S l / 2\tau_0. \quad (\text{II.5.3})$$

С учетом (II.3. 9) зависимость тока термогенерации от напряжения имеет вид

$$I_g(V) = I_{g0} \sqrt[k]{(\Phi_{BE} - V) / \Phi_{BE}}, \quad (\text{II.5.4})$$

где

$$I_{g0} = en_i S l_0 / 2\tau_0, \quad (\text{II.5.5})$$

$k = 2$ для ступенчатого (или сильно асимметричного диффузионного) перехода и $k = 3$ для линейного перехода.

Заметим, что ток термогенерации пропорционален собственной концентрации носителей заряда n_i , в то время как тепловой ток пропорционален n_i^2 (II.4.26). В связи с этим значения температур T_S и T_2 , определяющих температурную зависимость обратного тока (II.4.33) (II.4.34), для тока термогенерации вдвое больше, чем для теплового тока.

Интересно сравнить величины тока термогенерации и теплового тока. Полагая эффективность эмиттера $\gamma = 1$



и $I_S = I_{SB}$, из (II.5.5), (II.4.26a) и (II.27) при $L_B^2 = D_B \tau_0$ получаем

$$\frac{I_{g0}}{I_S} = \frac{en_i l G_B}{2en_i^2 \tau_0} = \frac{l N_B w_B}{2n_i \tau_0 D_B} \cdot \frac{\text{th}(w_B / L_B)}{w_B / L_B} = \frac{N_B l}{2n_i L_B} \text{th} \frac{w_B}{L_B}.$$

В германиевых диодах, ввиду высокой собственной концентрации n_i , $I_{g0} \ll I_S$ за исключением области очень низких температур. В кремниевых диодах собственная концентрация носителей заряда значительно меньше, поэтому при комнатной температуре обычно выполняется неравенство $I_{g0} \gg I_S$. С ростом температуры тепловой ток увеличивается быстрее, чем ток термогенерации, поэтому тепловой ток превосходит ток термогенерации при температурах порядка 120 °С. Абсолютная величина тока термогенерации, так же как и теплового тока, в кремниевых диодах значительно меньше, чем в германиевых. Типичные обратные ветви ВАХ германиевого и кремниевых диодов при комнатной температуре представлены на рис. II.5.1.

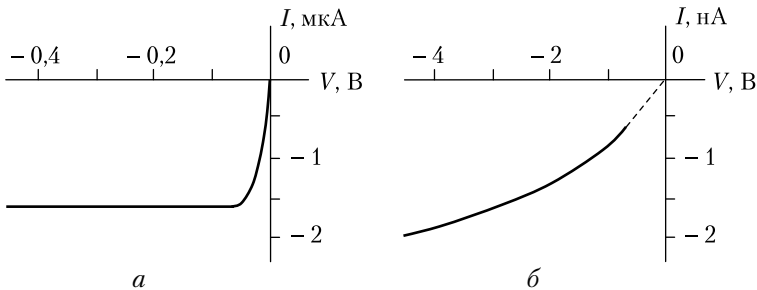


Рис. II.5.1. Обратные ветви ВАХ германиевого (а) и кремниевых (б) диодов

II.5.4. Ток рекомбинации (прямое смещение перехода)

При прямом напряжении $V > 3\phi_T$ концентрация носителей заряда в ОПЗ значительно превышает собственную концентрацию n_i , скорость рекомбинации значительно превосходит скорость термогенерации и в соотношении (II.5.1)

$$r - g = \frac{n_i^2 (e^{V/\phi_T} - 1)}{\tau_0 (n + p + 2n_i)} \gg g. \quad (\text{II.5.6})$$

Учитывая, что согласно (II.3.2) $n(x)p(x) = n_i^2 e^{V/\Phi_T}$, нетрудно убедиться, что сумма $n + p$ в знаменателе правой части (II.5.6) минимальна при $n = p = n_i e^{V/2\Phi_T}$. Таким образом, скорость рекомбинации достигает максимального значения в плоскости $x = x_m$, где $n(x_m) = p(x_m) = n_i e^{V/2\Phi_T}$ (рис. II.5.2)

$$(r - g)_{\max} = \frac{n_i^2 (e^{V/\Phi_T} - 1)}{\tau_0 (n + p + 2n_i)_{\min}} = \frac{n_i^2 (e^{V/\Phi_T} - 1)}{2\tau_0 n_i (e^{V/2\Phi_T} + 1)} \approx \frac{n_i}{2\tau_0} e^{V/2\Phi_T}. \quad (\text{II.5.7})$$

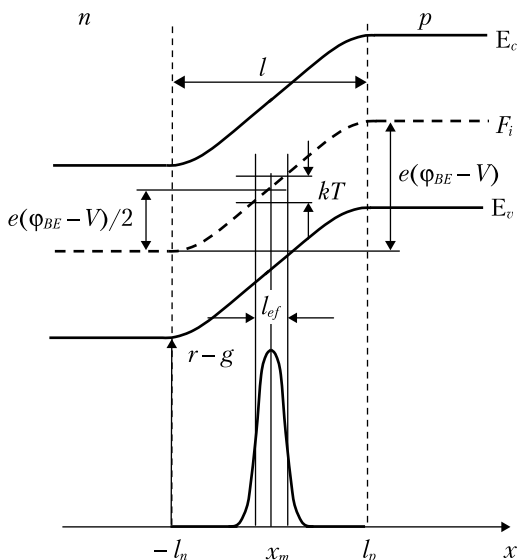


Рис. II.5.2. Изменение скорости рекомбинации электронов и дырок в ОПЗ

Применяя теорему о среднем, можно вычислить ток рекомбинации, используя в соотношении (II.5.2) максимальное значение скорости рекомбинации (II.5.7) и некоторую эффективную длину l_{ef} той части ОПЗ, где скорость рекомбинации близка к максимальной (рис. II.5.2):

$$I_r = eS l_{ef} (r - g)_{\max} = \frac{eS l_{ef} n_i}{2\tau_0} e^{V/2\Phi_T}. \quad (\text{II.5.8})$$

С учетом (II.5.5) получаем

$$I_r = I_{g0} (l_{ef} / l_0) e^{V/2\Phi_T}. \quad (\text{II.5.9})$$

Оценить значение эффективной длины l_{ef} можно, учитывая, что на этой длине электростатическая энергия изменяется на величину $\sim kT$ (см. рис. II.5.2):

$$l_{ef} = \frac{k_r l k T}{e(\varphi_{BE} - V)} = \frac{k_r l \varphi_T}{\varphi_{BE} - V},$$

где $k_r \sim 1$. С учетом (II.3.9) получаем

$$l_{ef} = k_r l_0 \frac{\varphi_T}{\varphi_{BE} - V} \cdot \sqrt[k]{\frac{\varphi_{BE} - V}{\varphi_{BE}}} = \frac{k_r l_0 \varphi_T}{\varphi_{BE}} \cdot \left(\frac{\varphi_{BE}}{\varphi_{BE} - V} \right)^{1-1/k}. \quad (\text{II.5.10})$$

Учитывая слабую зависимость $l_{ef}(V)$ по сравнению с экспоненциальной зависимостью $I_r(V)$ в (II.5.9), ВАХ тока рекомбинации можно представить в виде

$$I_r \approx I_{r0} e^{V/m\varphi_T}, \quad (\text{II.5.11})$$

где $m \approx 2$ — коэффициент неидеальности. Практически значение m несколько меньше 2 вследствие зависимости $l_{ef}(V)$.

Выражение (II.5.11) показывает, что ток рекомбинации слабее нарастает с ростом прямого напряжения ($\sim e^{V/2\varphi_T}$), чем ток инжекции ($\sim e^{V/\varphi_T}$). Поэтому вклад тока рекомбинации в общий ток диода заметен только при малых уровнях тока. В германиевых диодах ток рекомбинации можно обнаружить только при низких температурах (когда значение n_i мало).

Единое аналитическое выражение, описывающее ВАХ тока генерации-рекомбинации в ОПЗ непрерывной и дифференцируемой функцией, получено в работе [7].

В заключение заметим, что сделанные в параграфе I.5.2 допущения о рекомбинации носителей заряда через единственный ловушечный уровень $E_t = F_i$, а также о равенстве коэффициентов захвата ловушками электронов и дырок не меняют качественно полученных результатов. Выражение (II.5.11) для ВАХ тока рекомбинации соответствует действительности и подтверждено многочисленными экспериментами. Его вид ($I_r \sim e^{V/2\varphi_T}$) является следствием того, что скорость рекомбинации максимальна в плоскости $x = x_m$, проходящей через половину высоты потенциального барьера (см. рис. II.5.2).

II.5.5. Сопротивление базы

Поскольку в реальных диодах эмиттер легирован значительно сильнее базы, влияние омического сопротивления эмиттера на ВАХ диода значительно меньше, чем сопротив-



ления базы. Суммарное напряжение, приложенное к диоду, при этом распределяется между *p-n* переходом и базовой областью (см. рис. II.4.1, а):

$$V_D = V + V_B = V + IR_B. \quad (\text{II.5.12})$$

Для диода прямоугольной геометрии сопротивление базы определяется соотношением

$$R_B = \rho_B w_B / S, \quad (\text{II.5.13})$$

где ρ_B — удельное сопротивление базовой области. Если уровень инжекции мал (в базе *p*-типа $n_p \ll N_B$), величина ρ_B зависит только от степени легирования базы

$$\rho_B = \frac{1}{en_n \mu_n + ep_p \mu_p} = \frac{1}{e \mu_p N_B}. \quad (\text{II.5.14})$$

При высоком уровне инжекции инжектированные в базу носители уменьшают ее удельное сопротивление. Если же толщина базы велика ($w_B \gg L_B$), то эти носители сосредоточены в сравнительно малой области базы вблизи перехода, и выражение (II.5.14) остается справедливым. При короткой базе в условиях высокого уровня инжекции формулой (II.5.14) пользоваться нельзя, этот случай будет рассмотрен в параграфе II.5.7. В соответствии с (II.5.12) ВАХ диода имеет вид

$$I = I_S(e^{V/\Phi_T} - 1) = I_S(e^{(V_D - IR_B)/\Phi_T}), \quad (\text{II.5.15})$$

а дифференциальное сопротивление диода при $I \gg I_S$ составляет

$$r = \Phi_T / I + R_B. \quad (\text{II.5.16})$$

Выражение (II.5.16) показывает, что влиянием сопротивления базы можно пренебречь при

$$I \ll I_R = \Phi_T / R_B, \quad (\text{II.5.17})$$

где I_R — ток омического вырождения ВАХ диода. При $I \gg I_R$ ВАХ диода вырождается в линейную (рис. II.5.3).

При токах, больших тока омического вырождения, сопротивление базы существенно влияет на ТКН диода. Так как степень легирования базы обычно невелика ($N_B < 10^{18} \text{ см}^{-3}$), основным механизмом рассеяния носителей заряда в ней является рассеяние на фононах. При этом подвижность основных носителей в базе уменьшается, а сопротивление базы возрастет с ростом температуры. Поэтому при больших токах ТКН может стать положительным (см. рис. II.5.3).



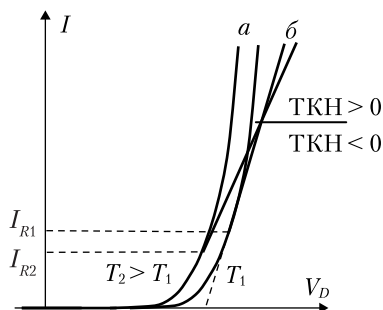


Рис. II.5.3. Влияние сопротивления базы на ВАХ диода:
кривые $a - R_B = 0$; кривые $б - R_B > 0$

II.5.6. Характеристики диода при высоком уровне инжекции

II.5.6.1. Особенности высокого уровня инжекции

При большой плотности тока концентрация неосновных носителей заряда, инжектированных в базу из эмиттера, может превышать концентрацию основных носителей заряда в базе. Этот случай соответствует высокому уровню инжекции. Для базы p -типа условие высокого уровня инжекции имеет вид

$$\delta(0) = n_p(0) / p_{p0} = n_p(0) / N_B \gg 1, \quad (\text{II.5.18})$$

где $\delta(0)$ — уровень инжекции в базе на границе с ОПЗ. Распределение концентраций электронов и дырок в короткой базе при высоком уровне инжекции представлено на рис. II.5.4.

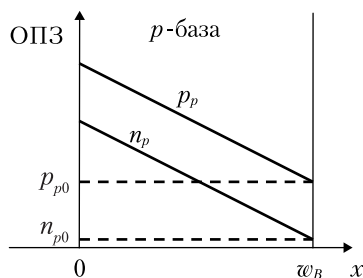


Рис. II.5.4. Распределение носителей заряда в базе при высоком уровне инжекции

Отметим основные особенности работы диода при высоком уровне инжекции

1. Граничное условие Шокли (II.3.6a) для концентраций неосновных носителей заряда на границе с ОПЗ при высоком уровне инжекции не выполняется. Его следует заменить условием (II.3.8).

2. Удельная проводимость базы зависит от уровня инжекции и от координаты x .

3. Заметную роль играет электрическое поле в базе, обусловленное двумя причинами: а) омическим падением напряжения на сопротивлении базы; б) нарушением электронейтральности базы вследствие разницы коэффициентов диффузии электронов и дырок (эффект Дембера).

Сложность задачи вынуждает для получения ВАХ диода при высоком уровне инжекции сделать следующие дополнительные допущения.

1. Длина базы мала ($w_B \ll L_B$). Это допущение оправдано тем, что в диодах с длинной базой при больших токах ВАХ вырождается в линейную (параграф II.5.5) и слабо зависит от инжекционных свойств *p-n* перехода.

2. Высокий уровень инжекции сохраняется во всей базе: $n_p(x) / N_B \gg 1$ при $0 \leq x \leq w_B$. Практически этого не может быть, так как в плоскости $x = w_B$ концентрации носителей принимают равновесные значения (см. рис. II.5.4). Однако для оценочных расчетов будем считать уровень инжекции очень высоким, так что протяженность области базы, где условие $n_p(x) / N_B \gg 1$ не выполнено, много меньше толщины базы w_B .

3. Эмиттер легирован значительно сильнее базы, и эффективность эмиттера близка к 1:

$$\gamma = \gamma_n \approx 1. \quad (\text{II.5.19})$$

II.5.6.2. Распределение носителей заряда в базе

Заметим, что при $n_p(x) \gg p_{p0}$ имеют место приближенные равенства

$$n_p(x) \approx \Delta n_p(x) = \Delta p_p(x) \approx p_p(x). \quad (\text{II.5.20})$$

Избыточные концентрации электронов и дырок при высоком уровне инжекции определяются одним из биполярных уравнений непрерывности (I.7.9), например (I.7.9a), которое в стационарном состоянии ($\partial p_p / \partial t = 0$) имеет вид



$$-\frac{\Delta p_p}{\tau} + D \frac{\partial^2 p_p}{\partial x^2} = 0,$$

или

$$\frac{\partial^2 p_p}{\partial x^2} - \frac{\Delta p_p}{L^2} = 0, \quad (\text{II.5.21})$$

где $L = \sqrt{D\tau}$ — биполярная диффузионная длина; $\tau = \tau_p = \tau_n$.

Электрическое поле не входит явно в уравнение (II.5.21), однако именно оно обеспечивает квазинейтральность базы и, в конечном счете, определяет вид уравнения. Кроме того, ввиду больших концентраций электронов и дырок в базе электрическое поле вызывает значительные дрейфовые токи.

При граничном условии на омическом контакте $\Delta p_p(w_B) = 0$ и $w_B \ll L_B$ решение уравнения (II.5.21) имеет вид, аналогичный (II.4.166)

$$\Delta p_p(x) = \Delta n_p(x) \approx \Delta p_p(0)(w_B - x) / w_B, \quad (\text{II.5.22})$$

где граничное условие в плоскости $x = 0$ определяется выражением (II.3.8) с учетом (II.5.20):

$$\Delta p_p(0) = \Delta n_p(0) = n_i e^{V/2\varphi_T}. \quad (\text{II.5.23})$$

Полные концентрации дырок и электронов примерно равны избыточным

$$p_p(x) \approx n_p(x) \approx \Delta p_p(0)(w_B - x) / w_B. \quad (\text{II.5.24})$$

II.5.6.3. Электрическое поле в базе

Для нахождения электрического поля в базе запишем уравнения для плотностей электронного и дырочного токов с учетом дрейфовых составляющих

$$j_n(x) = eD_n dn_p(x) / dx + en_p(x)\mu_n E(x); \quad (\text{II.5.25a})$$

$$j_p(x) = -eD_p dp_p(x) / dx + ep_p(x)\mu_p E(x). \quad (\text{II.5.25б})$$

Поскольку кинетические коэффициенты $D_{n,p}$ и $\mu_{n,p}$ для электронов и дырок имеют один порядок величины, а при высоком уровне инжекции концентрации электронов и дырок практически одинаковы (II.5.20), абсолютные величины плотностей как диффузионных, так и дрейфовых токов дырок и электронов в правых частях (II.5.25) также не могут различаться сильно:



$$|eD_p dp_p(x) / dx| \approx |eD_n dn_p(x) / dx|$$

$$\text{и } ep_p(x)\mu_p E(x) \approx en_p(x)\mu_n E(x).$$

В то же время согласно допущению (II.5.19) плотность дырочного тока значительно меньше, чем электронного: $j_p(x) \ll j_n(x)$. Такая ситуация возможна только в том случае, когда слагаемые в правой части уравнения (II.5.25б) близки по абсолютной величине и значительно превышают их алгебраическую сумму, т.е.

$$ep_p(x)\mu_p E(x) \approx eD_p dp_p(x) / dx.$$

Отсюда с учетом соотношения Эйнштейна $D_p = \Phi_T \mu_p$ получаем

$$E(x) \approx \frac{\Phi_T}{p_p} \cdot \frac{dp_p}{dx} \approx \frac{\Phi_T}{n_p} \cdot \frac{dn_p}{dx}. \quad (\text{II.5.26})$$

II.5.6.4. ВАХ p - n перехода

Подстановка (II.5.26) в (II.5.25а) с учетом (II.5.24) и (II.5.23) позволяет определить плотность электронного тока

$$j \approx j_n(0) = \frac{e2D_B \Delta n_p(0)}{w_B} = \frac{e2D_B n_i}{w_B} e^{V/2\Phi_T}, \quad (\text{II.5.27})$$

где, как и ранее, D_B — коэффициент диффузии неосновных носителей заряда в базе (в данном случае для p -базы $D_B = D_n$).

Выражение (II.5.27) показывает, что влияние электрического поля на ток неосновных носителей в базе сводится к удвоению их коэффициента диффузии (дрейфовая и диффузионная составляющие токов одинаковы).

II.5.6.5. ВАХ диода

Полное напряжение на диоде складывается из падения напряжения на p - n переходе и на нейтральной области базы

$$V_D = V + V_B. \quad (\text{II.5.28})$$

Последнее находится путем интегрирования соотношения (II.5.26) для поля в базе с учетом граничного условия (II.5.23)

$$V_B = \int_0^{w_B} E(x) dx = \Phi_T \int_{n_p(w_B)}^{n_p(0)} \frac{dn_p}{n_p} = \Phi_T \ln \frac{n_p(0)}{n_{p0}} = \frac{V}{2} + \Phi_T \ln \frac{N_B}{n_i}. \quad (\text{II.5.29})$$



Подстановка (II.5.29) в (II.5.28) дает

$$V = \frac{2}{3} \left(V_D - \phi_T \ln \frac{N_B}{n_i} \right). \quad (\text{II.5.30})$$

Наконец, подставляя (II.5.30) в (II.5.27), с учетом (II.5.28) получим ВАХ диода при высоком уровне инжекции

$$I \approx I_p(0) = I_S^* e^{V_D/3\phi_T}, \quad (\text{II.5.31})$$

где
$$I_S^* = 2eSD_B n_i^{2/3}. \quad (\text{II.5.32})$$

Учитывая, что согласно (II.4.26), (II.4.27) в диоде с короткой базой тепловой ток при низком уровне инжекции определяется соотношением $I_S \approx I_{SB} = eS n_i^2 D_B / N_B w_B$, получаем

$$I_S^* = 2I_S (N_B / n_i)^{2/3}. \quad (\text{II.5.33})$$

ВАХ (II.5.31) показывает, что при высоком уровне инжекции зависимость тока от напряжения получается значительно менее резкой, чем при низком уровне ($I \sim e^{V_D/3\phi_T}$). Это связано с изменением граничного условия при $x = 0$ ($V / \phi_T \rightarrow V / 2\phi_T$), а также с падением напряжения на нейтральной области базы ($V / 2\phi_T \rightarrow V_D / 3\phi_T$). Заметим, что сопротивление базы не вырождает ВАХ диода в линейную, так как при увеличении уровня инжекции сопротивление базы снижается.

II.5.6.6. Эффективность эмиттера

Эффективность эмиттера (в нашем случае — коэффициент инжекции электронов) определяется соотношением (II.4.41a)

$$\gamma = \frac{1}{1 + I_p(0) / I_n(0)}.$$

Поскольку уровень инжекции в эмиттере всегда остается низким, при $V > 3\phi_T$ дырочный ток $I_p(0) \sim e^{V/\phi_T}$. Согласно (II.5.27) при высоком уровне инжекции электронный ток $j_n(0) \sim e^{V/2\phi_T}$. Поэтому с ростом тока (т.е. напряжения V на p - n переходе) эффективность эмиттера снижается.

Заметим, что ВАХ p - n перехода (II.5.27) получена при допущении $\gamma \approx 1$ для очень высокого уровня инжекции (допущения 2 и 3 подпараграфа II.5.6.1). Поэтому полученные для высокого уровня инжекции результаты имеют скорее качественный характер.



II.5.6.7. Границы высокого и низкого уровней инжекции

С учетом (II.5.18), (II.5.13), (II.5.14) и (II.5.17) из соотношения (II.5.27) для диодов с короткой базой получаем

$$\delta = bI / 2I_R,$$

где I_R — ток омического вырождения; b — отношение подвижностей неосновных и основных носителей в базе. Таким образом, единичный уровень инжекции реализуется при $I \approx I_R$. При $I \ll I_R$ диод работает в условиях низкого уровня инжекции, а при $I \gg I_R$ — в условиях высокого уровня инжекции.

В диоде с длинной базой омическое вырождение наступает уже при низком уровне инжекции.

II.5.7. ВАХ реального диода

Особенности прямой ветви ВАХ реального диода поясняются рис. II.5.5. Для удобства ВАХ представлены в полупрологарифмическом масштабе и в безразмерных координатах V_D / Φ_T , I / I_S .

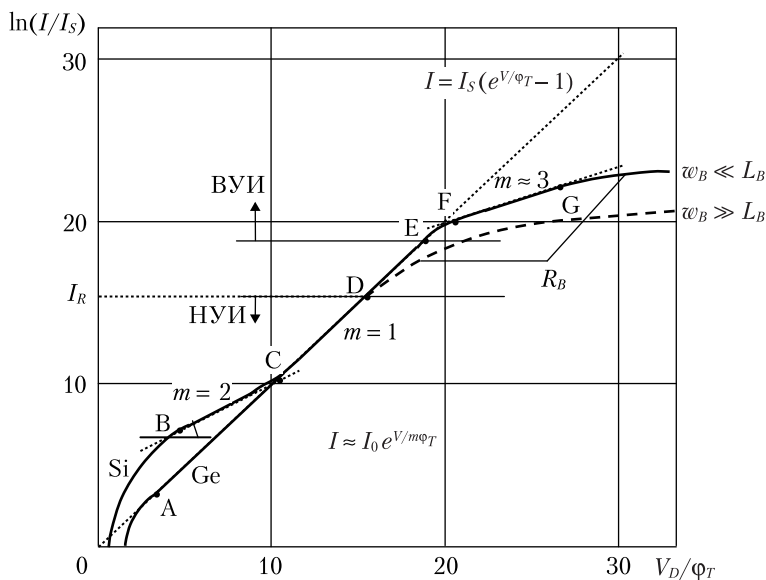


Рис. II.5.5. Особенности прямой ВАХ реального диода

Пунктирная линия изображает ВАХ идеализированного диода

$$I / I_S = e^{V/\Phi_T} - 1,$$

которая при $V > 3\Phi_T$ в выбранном масштабе представляет собой прямую линию с единичным наклоном.

ВАХ германиевого диода при температуре $T = 300$ К практически совпадает с этой линией вплоть до точки D (длинная база, $w_B \gg L_B$) или точки E (короткая база, $w_B \ll L_B$).

В диоде с длинной базой правее точки D ВАХ начинает вырождаться в линейную и в выбранном масштабе представляет собой медленно (логарифмически) нарастающую кривую. Точке D соответствует ток омического вырождения I_R и конец участка низкого уровня инжекции.

В диоде с короткой базой сопротивление базы мало и в выбранном масштабе ВАХ остается примерно линейной вплоть до точки E, соответствующей началу участка высокого уровня инжекции. При высоком уровне инжекции (участок FG) ВАХ описывается уравнением (II.5.31) и изображается прямой линией с тангенсом угла наклона $1/3$. Правее точки G соотношение (II.5.31) становится несправедливым, так как эффективность эмиттера снижается и не может считаться близкой к 1.

ВАХ кремниевого диода при температуре $T = 300$ К левее точки C определяется током рекомбинации в ОПЗ и на участке ВС, в соответствии с уравнением (II.5.11), представляет собой прямую линию с тангенсом угла наклона, близким к $1/2$. В области больших токов поведение кремниевых и германиевых диодов качественно одинаково.

Для практических расчетов ВАХ часто аппроксимируется функцией

$$I = I_0(e^{V_D/m\Phi_T} - 1), \quad (\text{II.5.34})$$

где $m \geq 1$ — коэффициент неидеальности. Разумеется, коэффициент неидеальности остается постоянным лишь в ограниченном диапазоне токов. Величина

$$\frac{1}{m} = \Phi_T \frac{d^2 I / dV^2}{dI / dV}$$

может служить критерием качества прямой ветви ВАХ диода, характеризуя степень ее нелинейности. В этом смысле идеализированный диод имеет наилучшее качество ВАХ, а все эффекты, отличающие ВАХ реального диода (сопротивление базы, ток рекомбинации в ОПЗ, эффекты высокого уровня



инжекции), снижают качество ВАХ. На участке линейного вырождения качество ВАХ наихудшее ($1/m \ll 1$).

На обратной ветви ВАХ ее качество определяется малостью обратного тока. В этом смысле идеализированный диод имеет наилучшее качество, так как его дифференциальное сопротивление $r \rightarrow \infty$ при увеличении обратного напряжения. Ток термогенерации снижает качество обратной ветви ВАХ, так как для него $m \approx 2$.

В заключение отметим, что линейный участок ВАХ в полупериодическом масштабе с тангенсом угла наклона $1/3$ при высоком уровне инжекции (участок FG на рис. II.5.5) редко удается наблюдать экспериментально, так как его теоретическое обоснование содержит слишком много допущений.

Основные выводы

1. В кремниевых диодах при малых токах существенную роль играют процессы генерации и рекомбинации носителей заряда в области p - n перехода. Первый из этих процессов приводит к увеличению обратного тока и его зависимости от напряжения. Второй процесс увеличивает прямой ток через диод, причем зависимость его от напряжения становится менее резкой, чем в идеализированной модели (коэффициент неидеальности $m \approx 2$).

2. При токах, больших тока омического вырождения, часть приложенного к диоду напряжения выделяется на сопротивлении базы. При этом в диодах с длинной базой ВАХ диода приближается к линейной. Сопротивление базы влияет на величину ТКН. При больших токах величина ТКН может стать положительной.

3. При больших прямых токах концентрация неосновных носителей в базе на границе с p - n переходом может превысить равновесную концентрацию основных носителей, что соответствует высокому уровню инжекции. При этом нарушаются граничные условия Шокли в базе, становится существенной дрейфовая составляющая тока неосновных носителей в базе, а в диодах с короткой базой уменьшается величина сопротивления базы. При высоком уровне инжекции зависимость тока от напряжения становится менее резкой, чем в идеализированной модели. Коэффициент неидеальности ВАХ p - n перехода близок к 2, а коэффициент неидеальности ВАХ диода — к 3. В диодах с короткой базой высокий уровень инжекции соответствует токам, большим тока омического вырождения.

4. Для практических расчетов ВАХ удобно аппроксимировать функцией

$$I = I_0(e^{V_D/m\Phi_T} - 1),$$

где коэффициент неидеальности m можно считать постоянным в ограниченном диапазоне токов.

II.6. Пробой p - n перехода

II.6.1. Механизмы пробоя p - n перехода

Экспериментально установлено, что при повышении обратного напряжения до некоторой величины обратный ток диода резко возрастает. Это явление называется *пробоем перехода*, а соответствующее напряжение — *напряжением пробоя*.

Существуют четыре основных механизма пробоя:

1) тепловой; 2) лавинный; 3) туннельный; 4) поверхностный.

Последний вид пробоя связан с изменением свойств полупроводника вблизи поверхности и сильно зависит от состояния поверхности. Если полупроводник в местах выхода перехода на поверхность покрыт диэлектриком достаточно высокого качества, поверхностный механизм пробоя не проявляется. Поверхностный пробой мы в дальнейшем рассматривать не будем.

Для определения напряжения пробоя следует рассчитать пробивные напряжения для каждого механизма в отдельности и выбрать из них наименьшее значение.

II.6.2. Лавинный пробой

При лавинном пробое носители заряда, ускоряющиеся в электрическом поле p - n перехода, могут набрать энергию, достаточную для образования электронно-дырочной пары. Новые носители, ускоряясь в поле перехода, производят новые акты ионизации, что в конечном итоге приводит к лавинному характеру умножения носителей и резкому увеличению обратного тока.

В состоянии равновесия процессы ударной ионизации компенсируются процессами ударной (Оже) рекомбинации (параграф I.5.3). При отрицательном напряжении преобладает ударная ионизация (рис. II.6.1).

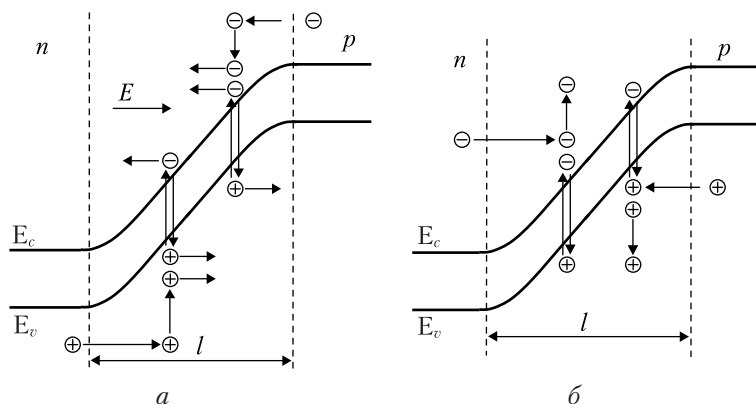


Рис. II.6.1. Размножение носителей заряда в *p-n* переходе:
 а — ударная ионизация; б — ударная (Оже) рекомбинация

Лавинный пробой происходит при таком напряжении, при котором каждый носитель, проходя через область объемного заряда, производит путем последовательных столкновений в среднем одну электронно-дырочную пару.

Действительно, пусть коэффициент размножения носителей в переходе есть M . Это значит, что каждый носитель заряда, проходя через переход, создает в среднем M электронно-дырочных пар (мы полагаем, что коэффициенты размножения электронов и дырок одинаковы). Поскольку суммарный путь, проходимый вторичными носителями, равен ширине перехода (см. рис. II.6.1), каждая ионизованная пара также создает на своем пути M пар. Результирующий ток запишется в виде

$$I = I_0(1 + M + M^2 + \dots) = I_0 / (1 - M), \quad (\text{II.6.1})$$

где I_0 — ток, созданный первичными носителями. При $M \rightarrow 1$ ток стремится к бесконечности, что соответствует лавинному пробую.

Пусть $\alpha(E)$ — коэффициент ударной ионизации, определяемый как среднее число электронно-дырочных пар, образованных одним носителем на единице пути. Тогда условие пробоя может быть записано в виде

$$M = \int_l \alpha[E(x)] dx = 1. \quad (\text{II.6.2})$$

Вообще говоря, $\alpha_n \neq \alpha_p$, однако при пробое количество электронов и дырок почти одинаково (носители размножаются парами, а их количество много больше исходного), поэтому в качестве α может быть использовано некоторое среднее значение между α_n и α_p .

Зависимость $\alpha(E)$ определяется вероятностью того, что на длине свободного пробега λ носитель наберет в поле энергию, равную пороговой энергии ионизации E_I . Если считать, что энергия ионизации набирается без единого столкновения, то вероятность этого процесса выражается соотношением

$$W_I = \frac{1}{\tau_{\text{ст}}} e^{-E_I/eE\lambda},$$

где $1/\tau_{\text{ст}}$ — частота столкновений; λ — длина свободного пробега. Таким образом,

$$\alpha(E) = \alpha_0 e^{-E_I/E}.$$

Для кремния параметры имеют следующие значения:

$$\alpha_0 = 1,1 \cdot 10^6 \text{ см}^{-1}; \quad E_I = 1,65 \cdot 10^6 \text{ В/см.}$$

При этом критерий пробоя (II.6.2) записывается в виде

$$\alpha_0 \int_{l(V_{brI})} e^{-E_I/E(x, V_{brI})} dx = 1, \quad (\text{II.6.3})$$

где V_{brI} — напряжение лавинного пробоя.

Зависимость коэффициента размножения носителей в переходе от напряжения описывается эмпирической формулой

$$M(V) \approx (V / V_{brI})^v, \quad (\text{II.6.4})$$

где для кремниевых переходов

$$v = \begin{cases} 3, & \text{ступенчатый переход, } p\text{-база;} \\ 5, & \text{ступенчатый переход, } n\text{-база;} \\ 4, & \text{линейный переход.} \end{cases} \quad (\text{II.6.4})$$

Для ступенчатых и резко асимметричных диффузионных переходов поле E и ширина перехода l являются функциями приложенного напряжения и результирующей концентрации примеси в базе N_B , поэтому условие (II.6.3) можно рассматривать как трансцендентное уравнение, определяющее зависимость $V_{brI}(N_B)$. С увеличением степени легирования базы напряжение лавинного пробоя уменьшается, так как увеличение поля в переходе резко увеличивает вероятность ионизации пар (рис. II.6.2).



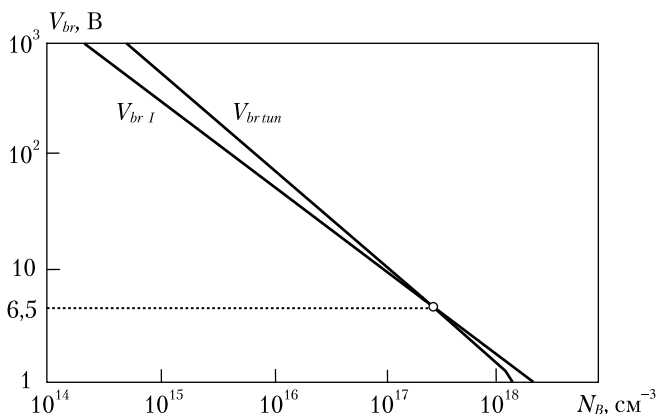


Рис. II.6.2. Зависимость напряжения пробоя от концентрации примеси в базе (ступенчатый переход, Si, 300 К)

Для всех изученных полупроводников при $T = 300$ К [6]

$$V_{brI} \approx V_0 (V_g / V_{g0})^{3/2} (N_0 / N_B)^{3/4},$$

где $V_g = E_g / e$; $V_{g0} = 1,1$ В; $V_0 = 60$ В; $N_0 = 10^{16}$ см $^{-3}$.

Учитывая, что концентрация примеси в базе определяет удельное сопротивление базовой области ρ_B , эту зависимость можно представить в виде $V_{brI}(\rho_B)$. Хорошим приближением является степенная функция [6]

$$V_{brI}(B) = A \rho_B^b \text{ (Ом} \cdot \text{см)}. \quad (\text{II.6.5})$$

Для кремниевого диода с *n*-базой: $A = 86$, $b = 0,64$.

Для кремниевого диода с *p*-базой: $A = 23$, $b = 0,75$.

Параметры линейного *p-n* перехода определяются градиентом концентрации примеси в переходе N' . Для всех изученных полупроводников при $T = 300$ К [6]

$$V_{brI} \approx V_0 (V_g / V_{g0})^{6/5} (N'_0 / N'_B)^{2/5},$$

где $N'_0 = 10^{20}$ см $^{-4}$, а остальные параметры те же.

Вид ВАХ в предпробойной области описывается полуэмпирической формулой

$$I / I_0 = 1 / [1 - (V / V_{brI})^v]. \quad (\text{II.6.6})$$

Для германиевой базы *n*-типа и кремниевой базы *p*-типа $v = 3$; для германиевой базы *p*-типа и кремниевой базы *n*-типа $v = 5$.

Номограммы для расчета напряжений лавинного пробоя диффузионных переходов приведены в работе [6].

II.6.3. Туннельный (зенеровский) пробой

При туннельном пробое увеличение обратного тока с ростом напряжения происходит за счет увеличения вероятности туннельного прохождения носителей через потенциальный барьер, создаваемый запрещенной зоной полупроводника. Из рис. II.6.3 очевидно, что при обратном смещении перехода происходит перекрытие энергетических областей, соответствующих валентной зоне p -области и зоне проводимости n -области ($E_{vp} > E_{cn}$). В результате становится вероятным туннелирование электронов из валентной зоны p -области в зону проводимости n -области. Эта вероятность увеличивается с увеличением обратного напряжения, так как уменьшается ширина барьера по основанию d (высота барьера постоянна и равна ширине запрещенной зоны E_g), а также увеличивается диапазон перекрытия разрешенных зон. Основной является первая причина.

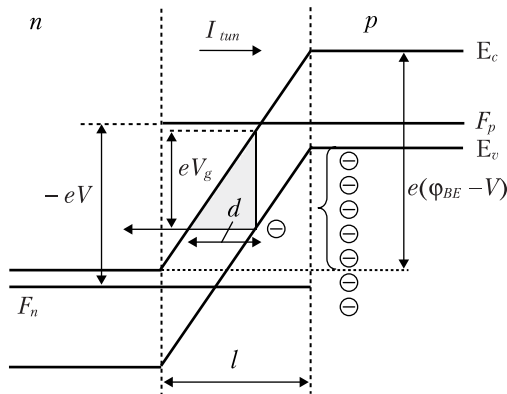


Рис. II.6.3. Энергетическая диаграмма p - n перехода при $V < 0$, поясняющая механизм туннельного тока через p - n переход

Таким образом, имеет место монотонный рост обратного тока при увеличении напряжения. Напряжение туннельного пробоя может быть определено лишь условно, если критерием пробоя считать превышение обратным током некоторой заданной величины.

Туннельная прозрачность барьера очень резко (экспоненциально) зависит от его ширины. Поэтому туннельный ток I_{tun} резко зависит от напряжения, что и дает право говорить о «пробое».

Приближенно можно считать, что пробой происходит при напряжении $V_{br\ tun}$, при котором ширина барьера по основанию достигает критической величины

$$d = d_{br}.$$

Если считать, что в ОПЗ энергетические уровни представляют собой прямые линии (что соответствует усреднению напряженности электрического поля), то из подобия треугольников на рис. II.6.3 получаем

$$d = lV_g / (\varphi_{BE} - V).$$

При пробое ($V = -V_{br\ tun}$, $l = l_{br\ tun}$) выполнено соотношение:
 $d_{br} = l_{br} V_g / (\varphi_{BE} + V_{br\ tun})$.

Для ступенчатого перехода

$$l_{br} = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_{BE} + V_{br\ tun})}{eN_B}}$$

$$\text{и } d_{br} = \frac{V_g}{\varphi_{BE} + V_{br\ tun}} \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_{BE} + V_{br\ tun})}{eN_B}} = \sqrt{\frac{2\varepsilon\varepsilon_0 V_g^2}{eN_B (\varphi_{BE} + V_{br\ tun})}}.$$

Отсюда

$$V_{br\ tun} = \frac{2\varepsilon\varepsilon_0 V_g^2}{eN_B d_{br}^2} - \varphi_{BE}. \quad (\text{II.6.7a})$$

Для линейного перехода

$$l_{br} = \sqrt[3]{\frac{12\varepsilon\varepsilon_0(\varphi_{BE} + V_{br\ tun})}{eN'}}$$

$$\text{и } d_{br} = \frac{V_g}{\varphi_{BE} + V_{br\ tun}} \sqrt[3]{\frac{12\varepsilon\varepsilon_0(\varphi_{BE} + V_{br\ tun})}{eN'}} = \sqrt[3]{\frac{12\varepsilon\varepsilon_0 V_g^3}{eN'(\varphi_{BE} + V_{br\ tun})^2}}.$$

Отсюда

$$V_{br\ tun} = \left(\frac{12\varepsilon\varepsilon_0}{eN'} \right)^{1/2} \left(\frac{V_g}{d_{br}} \right)^{3/2} - \varphi_{BE}. \quad (\text{II.6.7б})$$

Формулы (II.6.7) дают правильные результаты при $d_{br} \approx 35$ нм.

В кремниевых переходах при $N_B > 2 \cdot 10^{18} \text{ см}^{-3}$ (для ступенчатого перехода) или $N' > 7 \cdot 10^{24} \text{ см}^{-4}$ (для линейного перехода) напряжение туннельного пробоя близко к нулю.

Номограммы для определения напряжения туннельного пробоя приведены в работе [6].

II.6.4. Особенности лавинного и туннельного механизмов пробоя

Как при туннельном, так и при лавинном пробое пробивное напряжение снижается с уменьшением ширины перехода, т.е. с повышением степени легирования базы. Однако для туннельного пробоя зависимость пробивного напряжения от степени легирования базы более резкая, чем для лавинного (см. рис. II.6.2).

Практически в переходах с напряжением пробоя до 5 В имеет место туннельный пробой, а с напряжением пробоя более 7 В — лавинный. В области пробивных напряжений от 5 до 7 В могут сосуществовать оба механизма пробоя.

Ввиду неограниченного размножения носителей при лавинном пробое ВАХ имеет очень резкий излом при $V = V_{br I}$ (рис. II.6.4). ВАХ туннельного тока имеет более «мягкий» характер (рис. II.6.4).

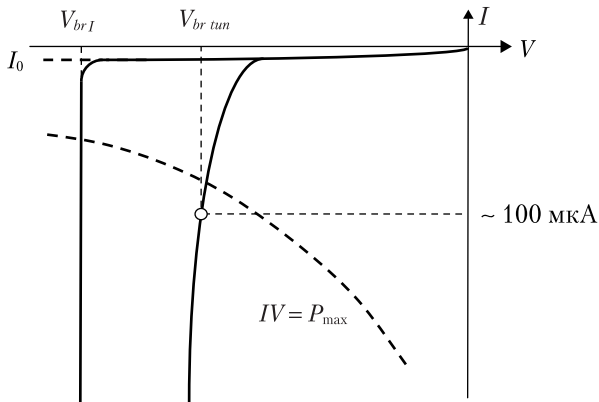


Рис. II.6.4. Сравнение ВАХ лавинного и туннельного пробоя.
Пунктир — кривая допустимой мощности

Различны и температурные характеристики этих двух видов пробоя.

Лавинный механизм пробоя реализуется при сравнительно малых концентрациях примеси (рис. II.6.2), когда рассеяние носителей заряда происходит в основном на фононах, число которых с увеличением температуры растет. Поэтому с ростом температуры уменьшается длина свободного пробега носите-

лей, падает вероятность ионизации атомов и, как следствие, возрастает напряжение лавинного пробоя.

При туннельном пробое с ростом температуры туннельный ток возрастает вследствие сужения (хотя и слабого) запрещенной зоны, а условно определяемое напряжение пробоя снижается.

В заключение отметим, что оба рассмотренных механизма пробоя являются обратимыми: если поглощаемая диодом мощность не превосходит максимально допустимой величины $P_{\max}$, при которой наступает тепловое разрушение образца, его характеристики восстанавливаются после снижения напряжения. Благодаря этому малая величина дифференциального сопротивления диодов в режиме лавинного или туннельного пробоя используется в специальных приборах — *кремниевых стабилизаторах*, предназначенных для стабилизаторов напряжения.

II.6.5. Тепловой пробой

При протекании обратного тока в *p-n* переходе выделяется тепло, в результате чего повышается температура перехода. Увеличение температуры приводит к возрастанию обратного тока и еще большему нагреву диода. Такая обратная связь может привести к неограниченному увеличению тока — пробую.

Рассмотрим тепловой баланс обратного смещенного *p-n* перехода ($V < 0$). В состоянии теплового равновесия поглощаемая диодом мощность

$$P_{\text{погл}} = VI \quad (\text{II.6.8})$$

должна быть равна рассеиваемой мощности

$$P_{\text{рас}} = \Delta T / R_T, \quad (\text{II.6.9})$$

где $I < 0$ — обратный ток; R_T — тепловое сопротивление диода; $\Delta T = T - T_0$ — температура перегрева; T_0 — температура окружающей среды.

Поскольку в кремниевых диодах обратный ток весьма мал и не вызывает обычно заметного разогрева перехода, тепловой пробой в основном имеет место в германиевых диодах, для которых в соответствии с (II.4.32)

$$I = -I_S = -I_S(T_0)e^{\Delta T/T_S}.$$

Условие теплового баланса дает



$$P = \Delta T / R_T = -I_S(T_0) V e^{\Delta T / T_S}. \quad (\text{II.6.10})$$

Графическое решение этого уравнения изображено на рис. II.6.5 для четырех значений напряжения: $0 > V_1 > V_2 > V_3 > V_4$. При $V = V_1$ точка A_1 соответствует состоянию устойчивого равновесия, а точка B_1 — неустойчивого равновесия. Действительно, при флуктуации температуры, например, в сторону ее увеличения, рассеиваемая мощность в состоянии A_2 становится больше поглощаемой, что приводит к уничтожению флуктуации. В состоянии B_2 рассеиваемая мощность оказывается меньше поглощаемой, что приводит к увеличению флуктуации. Зависимость $I(V)$, построенная на основании графического решения уравнения (II.6.10), имеет участок отрицательного дифференциального сопротивления, соответствующий неустойчивым состояниям, типа точки B_1 (рис. II.6.6). При напряжении V_3 дифференциальное сопротивление диода равно нулю. Абсолютная величина этого напряжения и есть напряжение теплового пробоя V_T , так как ток при этом неограниченно нарастает.

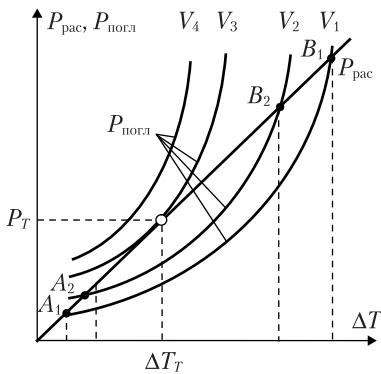


Рис. II.6.5. Тепловой баланс обратнo-смещенного p - n перехода

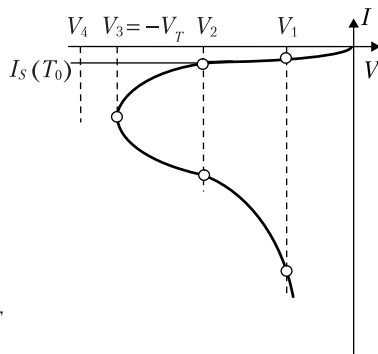


Рис. II.6.6. ВАХ теплового пробоя

Очевидно, при $V = -V_T$ должны выполняться равенства

$$\begin{cases} P_{\text{погл}} = P_{\text{рас}}, \\ \partial P_{\text{погл}} / \Delta T = \partial P_{\text{рас}} / \Delta T. \end{cases}$$

Отсюда

$$1 / R_T = I_S(T_0) V_T \frac{1}{T_S} e^{\Delta T_T / T_S},$$

где ΔT_T — перегрев перехода при пробое.

Решая полученную систему уравнений относительно ΔT_T и V_T , находим

$$\begin{cases} \Delta T_T = T_S, & (\text{II.6.11}) \end{cases}$$

$$\begin{cases} V_T = T_S / 2,72 R_T I_S(T_0). & (\text{II.6.12}) \end{cases}$$

Выражение (II.6.12) показывает, что тепловой пробой характерен для сравнительно мощных, в первую очередь германиевых, диодов с повышенным значением теплового тока. Для повышения напряжения теплового пробоя необходимо улучшать условия теплоотвода (снижать тепловое сопротивление).

Основные выводы

1. В полупроводниковых приборах низкой мощности основными механизмами пробоя *p-n* перехода являются лавинный и туннельный.

2. Лавинный пробой происходит при напряжении, для которого коэффициент размножения электронно-дырочных пар в переходе равен 1. Напряжение лавинного пробоя увеличивается при уменьшении степени легирования базы и при повышении температуры.

2. Туннельный пробой характерен для узких сильно легированных переходов. При увеличении температуры напряжение туннельного пробоя уменьшается.

3. В кремниевых диодах при напряжении пробоя, большем 7 В, вероятным является лавинный механизм пробоя, при напряжении пробоя, меньшем 5 В, — туннельный.

4. Малые дифференциальные сопротивления диода в режиме лавинного или туннельного пробоя позволяют использовать диоды как стабилизаторы напряжения.

5. При тепловом пробое *p-n* перехода обратная ветвь ВАХ диода имеет участок отрицательного дифференциального сопротивления. Тепловой пробой характерен для мощных германиевых диодов с плохим теплоотводом. Для повышения пробивного напряжения необходимо улучшать качество теплоотвода и применять приборы с малым обратным током.



II.7. Частотные и импульсные свойства полупроводникового диода

II.7.1. Механизмы инерционности диода

Как и всякий электронный прибор, полупроводниковый диод обладает инерционностью. При работе в режиме переключения из проводящего (открытого) в непроводящее (закрытое) состояние инерционность диода проявляется в следующем:

1) после подачи на открытый диод запирающего напряжения его высокое обратное сопротивление устанавливается не мгновенно, а с некоторой временной задержкой;

2) после подачи на закрытый диод отпирающего напряжения инжекция носителей заряда через переход также начинается с некоторой временной задержкой. После отпираания перехода прямое напряжение на нем устанавливается в течение некоторого конечного времени.

При работе в режиме малого сигнала инерционность диода проявляется в том, что его адмиттанс (полная проводимость) имеет реактивную составляющую.

Инерционность диода связана с двумя физическими процессами накопления заряда в различных областях диода.

Во-первых, при изменении напряжения на p - n переходе изменяется его ширина и, следовательно, полное количество положительного и отрицательного заряда в области перехода. Поскольку изменение заряда конечным током можно осуществить только за конечное время, напряжение на переходе не может измениться мгновенно.

Во-вторых, при изменении напряжения на диоде изменяются заряды неосновных носителей, инжектированных из эмиттера в базу (и, в меньшей степени, из базы в эмиттер). Изменение этих зарядов также не может произойти мгновенно.

Таким образом, инерционность диода обусловлена накоплением зарядов в области p - n перехода и в квазинейтральных областях базы и эмиттера. Первый процесс учитывается введением в эквивалентную схему диода *барьерной емкости p - n перехода*, второй — *диффузионных емкостей базы и эмиттера*.

II.7.2. Барьерная емкость p - n перехода

На рис. II.7.1, *а* схематически представлено распределение зарядов в области перехода и границы перехода при напряжении V (сплошные линии). При изменении напряжения



на малую величину dV ширина перехода изменяется. Если $dV > 0$, переход сужается и заштрихованные на рисунке области становятся электронейтральными. В результате заряд в эмиттерной области перехода уменьшается на величину $+dQ$, а заряд в базовой области перехода увеличивается на ту же величину (разумеется, переход в целом остается электронейтральным).

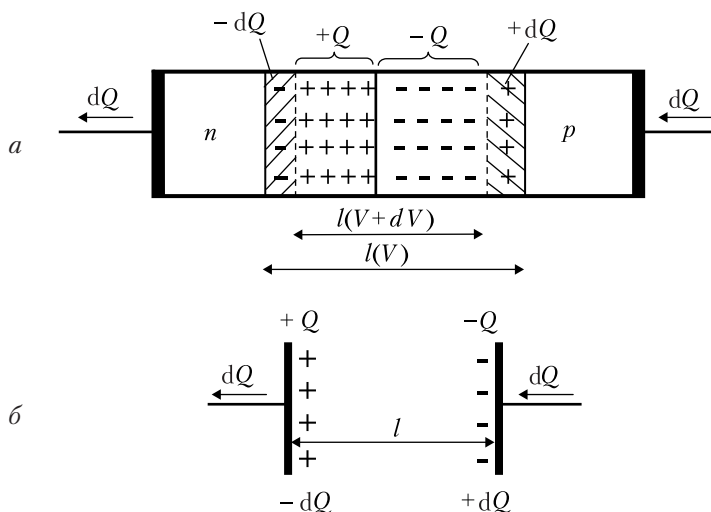


Рис. II.7.1. Распределение зарядов в *p-n* переходе (а) и в плоском конденсаторе (б)

Нейтрализация заштрихованных областей перехода осуществляется дополнительными (не связанными с протеканием постоянного тока) дырками и электронами, поступившими в эмиттер и базу, соответственно, из внешней цепи.

Таким образом, при изменении напряжения на диоде на величину dV в цепи диода протекает заряд dQ , изменяющий заряд эмиттерной и базовой областей перехода на величины $-dQ$ и $+dQ$, соответственно. Аналогичные процессы протекают в плоском конденсаторе (рис. II.7.1, б), электрическая емкость которого определяется соотношением

$$C = \frac{Q}{V} = \frac{dQ}{dV} = \frac{\epsilon \epsilon_0 S}{l},$$

где ϵ — диэлектрическая проницаемость изолятора; l — расстояние между обкладками.

Учесть накопление заряда в области перехода можно путем введения в эквивалентную схему диода *барьерной емкости* (рис. II.7.2). Эта емкость подключается параллельно безынерционному элементу D, ВАХ которого соответствует ВАХ диода без учета сопротивления базы, введенного в схему отдельно. Единственное отличие барьерной емкости перехода от емкости плоского конденсатора заключается в том, что в плоском конденсаторе заряды $\pm Q$ сосредоточены в плоскостях внутренних поверхностей обкладок, и расстояние между ними l остается неизменным. В барьерной емкости заряды $\pm Q$ занимают весь переход, однако *приращение* этих зарядов $\pm dQ$ также сосредоточены в плоскостях, ограничивающих p - n переход и отстоящих друг от друга на расстояние $l(V)$. Таким образом, *дифференциальная барьерная емкость* p - n перехода определяется так же, как емкость плоского конденсатора:

$$C(V) = \frac{dQ}{dV} = \frac{\epsilon \epsilon_0 S}{l(V)}, \quad (\text{II.7.1})$$

где l — ширина ОПЗ.

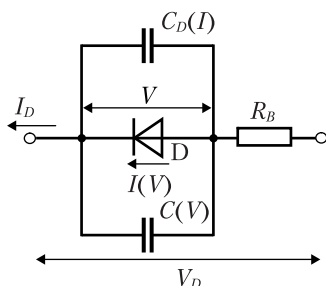


Рис. II.7.2. Эквивалентная схема полупроводникового диода

Зависимость этой емкости от напряжения на переходе согласно (II.3.9) имеет вид

$$C(V) = C_0 \sqrt[k]{\varphi_{BE} / (\varphi_{BE} - V)}, \quad (\text{II.7.2})$$

где

$$C_0 = \epsilon \epsilon_0 S / l_0$$

— равновесное значение барьерной емкости (при $V = 0$), $k = 2$ для ступенчатого перехода, $k = 3$ — для линейного.

Интегральное значение барьерной емкости перехода определяется как отношение избыточного (по сравнению с равновесным) заряда в переходе к напряжению

$$\bar{C}(V) = \frac{Q(V) - Q(0)}{V}. \quad (\text{II.7.3})$$

Интегральная и дифференциальная емкости связаны между собой соотношением

$$\bar{C}(V) = \frac{1}{V} \int_0^V C(V) dV. \quad (\text{II.7.4})$$

С учетом (II.7.2), (II.7.4)

$$\frac{\bar{C}(V)}{C(V)} = \frac{k\phi_{BE}}{(k-1)V} \left(\sqrt[k]{\frac{\phi_{BE} - V}{\phi_{BE}}} - \frac{\phi_{BE} - V}{\phi_{BE}} \right). \quad (\text{II.7.5a})$$

При $V \rightarrow 0$ значения интегральной и дифференциальной емкостей совпадают. При $-V \gg \phi_{BE}$ из (II.7.5a) получаем

$$\frac{\bar{C}(V)}{C(V)} = \frac{\kappa}{\kappa - 1}. \quad (\text{II.7.5б})$$

Отметим, что для расчета электронных схем более важной является дифференциальная емкость. Действительно, ток смещения через емкость есть

$$I_C = \frac{dQ}{dt} = \frac{dQ}{dV} \cdot \frac{dV}{dt} = C(V) \frac{dV}{dt}.$$

В дальнейшем без специальных оговорок под термином «барьерная емкость» мы будем понимать дифференциальную барьерную емкость.

Интегральная емкость может быть использована для оценки длительности переходного процесса ее заряда

$$\Delta t \approx \frac{\Delta Q}{I_C} = \frac{Q(V_2) - Q(V_1)}{\bar{I}_C} = \frac{\bar{C}(V_2)V_2 - \bar{C}(V_1)V_1}{\bar{I}_C},$$

где $\bar{I}_C$ — средний ток через емкость; V_1 и V_2 — начальное и конечное значения напряжения на емкости.

Вольт-фарадная характеристика барьерной емкости представлена на рис. II.7.3. Пунктирная кривая соответствует формуле (II.7.2), которая получена в приближении полного обеднения ОПЗ носителями заряда, когда $I(V) \xrightarrow{V \rightarrow \phi_{BE}} 0$, и

$C(V) \xrightarrow{V \rightarrow \phi_{BE}} \infty$. Реально приближение полного обеднения перестает выполняться при $V \gtrsim \phi_{BE} / 2$. При $V \rightarrow \phi_{BE}$ ширина ОПЗ имеет порядок дебаевской длины экранирования в базе $L_{DB} = \sqrt{\epsilon \epsilon_0 \phi_T / e N_B}$ и $C \rightarrow C_{\max} \approx \epsilon \epsilon_0 S / L_{DB}$ (см. рис. II.7.3).



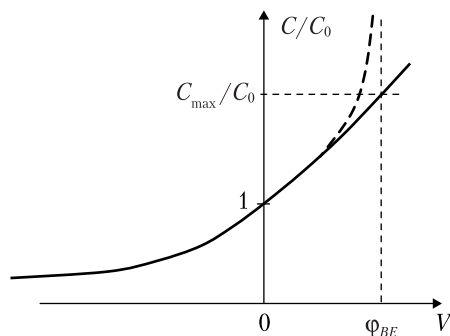


Рис. II.7.3. Зависимость барьерной емкости p - n перехода от напряжения

II.7.3. Диффузионные емкости диода

На рис. II.7.4 изображены распределения концентраций носителей заряда в базе и эмиттере при положительном напряжении на переходе.

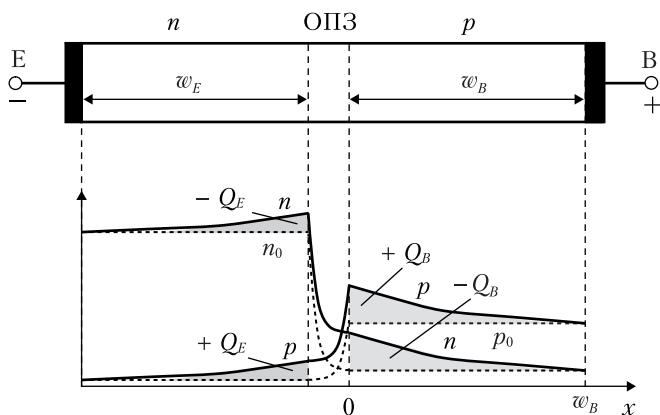


Рис. II.7.4. Заряды в диффузионных емкостях базы и эмиттера при $V > 0$

База и эмиттер являются электрически нейтральными, причем нейтральность имеет локальный характер (в базе $\Delta n_p = \Delta p_p$). Однако избыточные электроны в базе появляются вследствие их инжекции из эмиттера (слева направо). Избыточные дырки, компенсирующие заряд электронов, вытягиваются из омического контакта базы (справа налево) электрическим

полем, созданным инжектированными электронами. Таким образом, накопление зарядов избыточных электронов ($-Q_B$) и избыточных дырок ($+Q_B$) в базе связано с протеканием заряда Q_B в цепи диода. В эмиттере также накапливаются избыточные носители заряда $+Q_E$ и $-Q_E$, однако в большинстве случаев ими можно пренебречь из-за одностороннего характера инжекции (см. параграф II.3.3).

Влияние процессов накопления этих зарядов на цепь диода аналогично влиянию некоторой электрической емкости C_D , включенной параллельно переходу (см. рис. II.7.2). Эта емкость называется *диффузионной емкостью* диода.

В отличие от барьерной емкости, моделирующей пространственное разделение заряда в переходе, диффузионная емкость описывает накопление положительных и отрицательных зарядов, не разделенных в пространстве. Локальная электрическая нейтральность в базе и эмиттере устанавливается за очень малое время (порядка максвелловского времени релаксации в материалах базы и эмиттера).

Диффузионную емкость базы можно оценить соотношением

$$C_B \approx \frac{dQ_B}{dV} = \frac{dQ_B}{dI_n(0)} \cdot \frac{dI_n(0)}{dV} = \gamma \frac{dQ_B}{dI_n(0)} \cdot \frac{dI}{dV} = \frac{\gamma}{r} \cdot \frac{dQ_B}{dI_n(0)}, \quad (\text{II.7.6})$$

где $\gamma = \gamma_n$ — эффективность эмиттера; $r = \Phi_T / (I + I_S)$ — дифференциальное сопротивление *p-n* перехода (II.4.37); $I_n(0)$ — электронный ток на границе ОПЗ с базой.

Равенство (II.7.6) не является вполне точным, так как при изменении напряжения на переходе на величину dV изменение заряда dQ_B связано не только с инжекцией (или экстракцией) носителей через переход, но и с процессами генерации и рекомбинации пар в базе. Эти последние процессы не связаны с протеканием тока в цепи диода, следовательно, формула (II.7.6) дает завышенное значение диффузионной емкости. Кроме того, для определения заряда Q_B в (II.7.6) мы используем *стационарное* распределение носителей заряда, которое отличается от мгновенных значений при быстрых изменениях напряжения на *p-n* переходе. Необходимые уточнения сделаны в параграфе II.7.4.

Заряд Q_B определяется соотношением

$$Q_B = eS \int_0^{w_B} \Delta n_p(x) dx. \quad (\text{II.7.7})$$



В диоде с длинной базой распределение избыточных электронов имеет экспоненциальный характер (рис. II.7.5, а)

$$\Delta n_p(x) = \Delta n_p(0)e^{-x/L_B},$$

и
$$Q_B = eS\Delta n_p(0)L_B. \quad (\text{II.7.8})$$

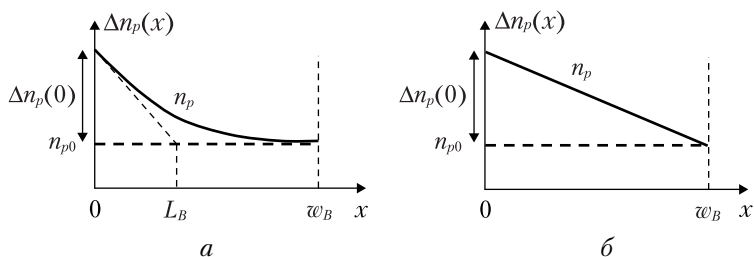


Рис. II.7.5. Распределение неосновных носителей заряда в диффузионной емкости базы диода с длинной (а) и короткой (б) базой

Учитывая, что электронный ток имеет диффузионный механизм

$$I_n(0) = eSD_B \frac{dn_p}{dx}(0) = eSD_B \frac{n_p(0)}{L_B}, \quad (\text{II.7.9})$$

из (II.7.8) и (II.7.9) получаем

$$\frac{dQ_B}{dI_n(0)} = \frac{Q_B}{I_n(0)} = \frac{L_B^2}{D_B} = \tau_B. \quad (\text{II.7.10})$$

Подстановка (II.7.10) в (II.7.6) дает

$$C_{DB} = \frac{\gamma\tau_B}{r} = \frac{\gamma\tau_B(I + I_s)}{\Phi_T}. \quad (\text{II.7.11})$$

В диоде с короткой базой распределение избыточных электронов имеет линейный характер (рис. II.7.5, б)

$$\Delta n_p(x) = \Delta n_p(0)(w_B - x) / w_B.$$

При этом

$$Q_B = eS\Delta n_p(0)w_B / 2; \quad (\text{II.7.12})$$

$$I_n(0) = eSD_B \frac{dn_p}{dx}(0) = eSD_B \frac{n_p(0)}{w_B}. \quad (\text{II.7.13})$$

Из (II.7.12) и (II.7.13) получаем

$$\frac{dQ_B}{dI_n(0)} = \frac{Q_B}{I_n(0)} = \frac{w_B^2}{2D_B}. \quad (\text{II.7.14})$$

В случае короткой базы инжектированные из эмиттера неосновные носители (в нашем случае дырки) пролетают до базового контакта, не успев рекомбинировать. Поэтому отношение $Q_B / I_n(0)$ в (II.7.14) имеет смысл *среднего времени диффузии неосновных носителей через базу*:

$$T_B = w_B^2 / 2D_B. \quad (\text{II.7.15})$$

Подстановка (II.7.14) в (II.7.6) дает

$$C_{DB} = \frac{\gamma T_B}{r} = \frac{\gamma T_B (I + I_S)}{\Phi_T}. \quad (\text{II.7.16})$$

Соотношения (II.7.11) и (II.7.16) показывают, что диффузионная емкость базы определяется током через диод и временем жизни неосновных носителей в базе (при $w_B \gg L_B$) либо средним временем диффузии неосновных носителей через базу (при $w_B \ll L_B$). С хорошей точностью можно полагать

$$C_{DB} = \frac{\gamma(I + I_S)}{\Phi_T(T_B^{-1} + \tau_B^{-1})}. \quad (\text{II.7.17a})$$

Аналогичное соотношение для диффузионной емкости эмиттера имеет вид

$$C_{DE} = \frac{(1 - \gamma)(I + I_S)}{\Phi_T(T_E^{-1} + \tau_E^{-1})}, \quad (\text{II.7.17б})$$

где $T_E = w_E^2 / 2D_E$ — среднее время диффузии неосновных носителей через эмиттер.

Полная диффузионная емкость диода есть сумма диффузионных емкостей базы и эмиттера:

$$C_D = C_{DB} + C_{DE}. \quad (\text{II.7.18})$$

Если эффективность эмиттера γ близка к 1, то $C_{DE} \ll C_{DB}$ и $C_D \approx C_{DB}$.

Диффузионная емкость закрытого диода ($V < -3\Phi_T$; $I \approx -I_S$) равна нулю. При $V > 3\Phi_T$ $I \gg I_S$ и диффузионная емкость пропорциональна току через *p-n* переход.



II.7.4. Частотные свойства диода на малом переменном сигнале

Выше отмечалось, что соотношение (II.7.6) для диффузионной емкости не является точным. Строго говоря, для ее определения следует найти малосигнальный адмиттанс диода и проанализировать его мнимую часть.

Положим, что к p - n переходу приложено напряжение V , представляющее собой сумму постоянного напряжения смещения V^0 и малого гармонического сигнала $\dot{V}^1(t)$

$$V(t) = V^0 + \dot{V}^1(t), \quad (\text{II.7.19})$$

$$\text{где} \quad \dot{V}^1(t) = V^1 e^{i\omega t}, \quad (\text{II.7.20})$$

причем

$$|\dot{V}^1| \ll \Phi_T. \quad (\text{II.7.21})$$

Положим также, что уровень инжекции остается низким, а эффективность эмиттера близка к 1, так что свойства диода определяются только процессами, происходящими в базовой области.

Для анализа характеристик диода в этом случае необходимо рассмотреть нестационарное биполярное уравнение непрерывности в p -базе

$$\frac{\partial n_p}{\partial t} = -\frac{\Delta n_p}{\tau_B} + D_B \frac{\partial^2 n_p}{\partial x^2}. \quad (\text{II.7.22})$$

Ввиду малости переменной составляющей $\dot{V}^1(t)$ приложенного напряжения концентрация электронов в базе и ток диода могут быть также представлены в виде суммы постоянных и малых переменных составляющих

$$n_p(x, t) = n_p^0(x) + \dot{n}_p^1(x, t) = n_p^0(x) + \dot{n}_p^1(x) e^{i\omega t}; \quad (\text{II.7.23})$$

$$I(t) = I^0 + \dot{I}^1(t) = I^0 + \dot{I}^1 e^{i\omega t}. \quad (\text{II.7.24})$$

При подстановке (II.7.23) в (II.7.22) уравнение (II.7.22) распадается на два, одно из которых описывает стационарное распределение электронов $n_p^0(x)$, а второе — переменную составляющую $\dot{n}_p^1(x)$. Решение стационарного уравнения исследовалось в разделе II.4. Уравнение для переменной составляющей имеет вид

$$\frac{\partial \dot{n}_p^1}{\partial t} = i\omega \dot{n}_p^1 = -\frac{\dot{n}_p^1}{\tau_B} + D_B \frac{\partial^2 \dot{n}_p^1}{\partial x^2}. \quad (\text{II.7.25})$$



Оно может быть представлено в следующей форме:

$$\frac{\partial^2 \dot{n}_p^1}{\partial x^2} - \frac{\dot{n}_p^1}{\dot{L}_B^2} = 0, \quad (\text{II.7.26})$$

$$\text{где} \quad \dot{L}_B = L_B / \sqrt{1 + i\omega\tau_B} \quad (\text{II.7.27})$$

— комплексная диффузионная длина неосновных носителей в базе.

Уравнение (II.7.26) по форме совпадает со стационарным уравнением (II.4.8а), отличаясь от него только заменой L_B на $\dot{L}_B$. Поэтому решение уравнения (II.7.26) имеет вид, аналогичный (II.4.15а) с той же заменой,

$$\dot{n}_p^1(x) = \dot{n}_p^1(0) \frac{\text{sh}[(w_B - x) / \dot{L}_B]}{\text{sh}(w_B / \dot{L}_B)}. \quad (\text{II.7.28})$$

Для определения граничной концентрации $\dot{n}_p^1(0)$ воспользуемся граничным условием Шокли (II.3.6) с учетом (II.7.19)

$$n_p(0) = n_{p0} e^{V/\Phi_T} = n_{p0} e^{V^0/\Phi_T} e^{\dot{V}^1/\Phi_T} \approx n_{p0} e^{V^0/\Phi_T} [1 + (V_1 / \Phi_T) e^{i\omega t}]. \quad (\text{II.7.29})$$

Последнее приближение в (II.7.29) основано на разложении экспоненциального множителя $e^{\dot{V}^1/\Phi_T}$ в степенной ряд при условии (II.7.21). Таким образом, граничное условие записывается в виде

$$n_p(0) = n_p^0(0) + \dot{n}_p^1(0), \quad (\text{II.7.30})$$

$$\text{где} \quad n_p^0(0) = n_{p0} e^{V^0/\Phi_T}, \quad (\text{II.7.31})$$

$$\dot{n}_p^1(0) = n_p^1(0) e^{i\omega t}, \quad (\text{II.7.32})$$

$$n_p^1(0) = n_p^0(0) V^1 / \Phi_T. \quad (\text{II.7.33})$$

Переменная составляющая тока через диод определяется выражениями (II.4.18а) и (II.4.19а) с заменой $L_B \rightarrow \dot{L}_B$ и $n_{pj} = n_p(0) \rightarrow \dot{n}_p^1(0)$

$$\dot{j}^1 = \frac{eSD_B \dot{n}_p^1(0)}{\dot{L}_B \text{th}(w_B / \dot{L}_B)}. \quad (\text{II.7.34})$$

Подставляя (II.7.32) в (II.7.34), с учетом (II.7.31), (II.7.33) получаем

$$\dot{j}^1 \approx \frac{eSD_B n_{p0}(0)}{L_B \text{th}(w_B / L_B)} \cdot e^{V^0/\Phi_T} \cdot \frac{L_B \text{th}(w_B / L_B)}{\dot{L}_B \text{th}(w_B / \dot{L}_B)} \cdot \frac{\dot{V}^1}{\Phi_T} \approx$$

$$\approx \frac{\dot{V}^1(I + I_S)}{\Phi_T} \cdot \frac{L_B \text{th}(w_B / L_B)}{\dot{L}_B \text{th}(w_B / \dot{L}_B)} = \frac{\dot{V}^1}{r} \cdot \frac{L_B \text{th}(w_B / L_B)}{\dot{L}_B \text{th}(w_B / \dot{L}_B)}. \quad (\text{II.7.35})$$

Таким образом, адмиттанс диода составляет

$$\dot{Y} = \frac{\dot{I}_1}{\dot{V}_1} = \frac{1}{r} \cdot \frac{L_B \text{th}(w_B / L_B)}{\dot{L}_B \text{th}(w_B / \dot{L}_B)}. \quad (\text{II.7.36})$$

В случае диода с длинной базой ($w_B \gg L_B > |\dot{L}_B|$) гиперболические тангенсы в (II.7.36) могут считаться равными 1:

$$\dot{Y} \approx \frac{1}{r} \cdot \frac{L_B}{\dot{L}_B} = \frac{\sqrt{1 + i\omega\tau_B}}{r}. \quad (\text{II.7.37})$$

В области не слишком высоких частот ($\omega\tau_B \ll 1$) функцию (II.7.37) можно разложить в ряд по степеням $\omega\tau_B$, ограничиваясь линейным членом

$$\dot{Y} \approx \frac{1}{r} + \frac{i\omega\tau_B}{2r} = \frac{1}{r} + i\omega C_{DB}, \quad (\text{II.7.38})$$

$$\text{где} \quad C_{DB} = \tau_B / 2r \quad (\text{II.7.39})$$

— диффузионная емкость базы.

Сравнение (II.7.39) и (II.7.11) показывает, что расчет диффузионной емкости диода с длинной базой с использованием соотношения (II.7.6) завышает значение диффузионной емкости вдвое.

В диоде с короткой базой ($w_B \ll |\dot{L}_B| < L_B$), разлагая числитель и знаменатель (II.7.36) в ряд по степеням w_B / L_B и $w_B / \dot{L}_B$ с точностью до двух первых членов, получаем:

$$\dot{Y} \approx \frac{1}{r} \left[1 - \frac{1}{3} \cdot \left(\frac{w_B^2}{\dot{L}_B^2} \right) \right] \approx \frac{1}{r} + i\omega C_{DB}, \quad (\text{II.7.40})$$

$$\text{где} \quad C_{DB} = \frac{w_B^2 \tau_B}{3r L_B} = \frac{2T_B}{3r} \quad (\text{II.7.41})$$

— диффузионная емкость базы.

Сравнение (II.7.41) и (II.7.16) показывает, что расчет диффузионной емкости диода с короткой базой при использовании соотношения (II.7.6) завышает значение диффузионной емкости в 1,5 раза.

На более высоких частотах, когда неравенство $\omega\tau_B \ll 1$ несправедливо, диффузионная емкость и активная проводимость диода зависят от частоты. Нетрудно показать, что с ростом частоты диффузионные емкости и сопротивление диода уменьшаются.



II.7.5. Импульсные свойства диода

Накопление зарядов в барьерной и диффузионной емкостях диода приводит к тому, что переключение диода из проводящего состояния в закрытое и обратно невозможно осуществить мгновенно. Процессы переключения диода рассмотрим на примере схемы, изображенной на рис. II.7.6. Если ключ находится в положении 1, через диод течет некоторый прямой (положительный) ток от источника ЭДС E_1 , определяемый величиной E_1 , сопротивлением резистора R_1 и ВАХ диода. Когда ключ находится в положении 2, диод закрыт и через него протекает обратный ток от источника ЭДС E_2 .

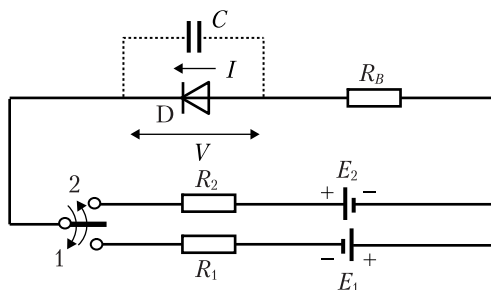


Рис. II.7.6. Электрическая схема переключения диода

Рассмотрим подробнее процессы, происходящие при переключении диода из проводящего состояния в закрытое и обратно.

Процесс выключения диода. Пусть в исходном состоянии ключ в течение длительного времени находился в положении 1, а в момент времени $t = 0$ мгновенно перебрасывается в положение 2. Тогда при $t < 0$ ток диода составляет

$$I = I_1 = \frac{E_1 - V_1}{R_1 + R_B} \approx \frac{E_1}{R_1 + R_B}, \quad (\text{II.7.42})$$

где V_1 — прямое напряжение на *p-n* переходе диода, R_B — сопротивление базы диода. Приближение в (II.7.42) основано на допущении о том, что $E_1 \gg V_1$.

Будем полагать, что изменение заряда в барьерной емкости открытого диода невелико и действием барьерной емкости можно пренебречь вплоть до момента запирающего перехода.

Распределение электронов в базе диода, временные диаграммы тока через переход и напряжения на диоде представлены на рис. II.7.7.

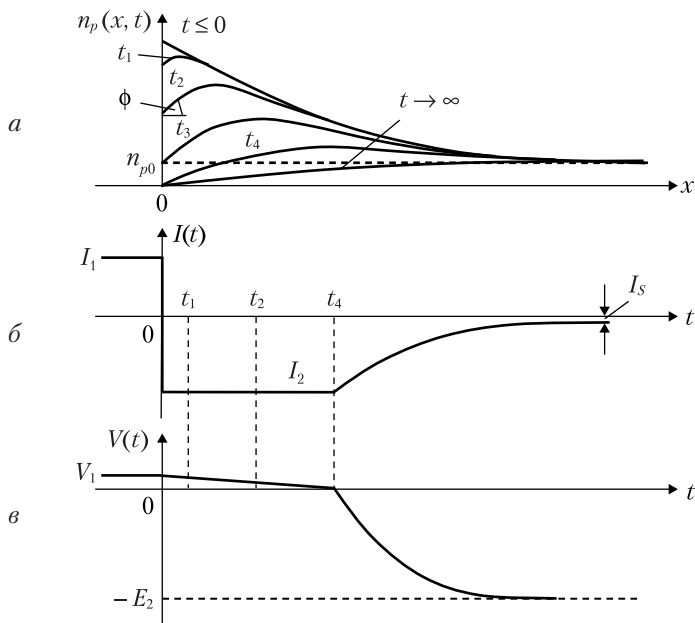


Рис. II.7.7. Процесс выключения диода:

a — распределение дырок в базе; *б* — осциллограмма тока через *p-n* переход; *в* — осциллограмма напряжения на *p-n* переходе

Для диода с длинной базой при $t \leq 0$ распределение электронов экспоненциально. После переброса ключа в положение 2 избыточные электроны в базе рассасываются током от источника ЭДС E_2 , а также рекомбинируют в базе. При этом распределение электронов в базе для последовательных моментов времени $0 < t_1 < t_2 < t_3 < t_4$ показано на рис. II.7.7, *a*. Момент времени $t = t_3$ соответствует равновесной концентрации электронов на границе с *p-n* переходом ($V = 0$), а момент времени $t = t_4$ — нулевой граничной концентрации электронов ($V \approx -3\phi_T$). Следовательно, с течением времени от $t = 0$ до $t = t_4$ напряжение на переходе изменяется от $V = V_1$ до $V = -3\phi_T$. При $E_1 \gg V_1$ ток через переход в течение этого временного интервала остается практически постоянным

$$I = -I_2 = \frac{-E_2 - V}{R_2 + R_B} \approx \frac{-E_2}{R_2 + R_B}, \quad (\text{II.7.43})$$

На осциллограмме тока (рис. II.7.7, *б*) наблюдается характерная «полочка». Поскольку величина тока определяется

градиентом концентрации электронов на границе с переходом ($x = 0$), в течение времени $0 \leq t \leq t_4$ этот градиент остается постоянным (постоянный угол ϕ на рис. II.7.7, *a*).

При $t > t_4$ градиент концентрации электронов при $x = 0$ не может более оставаться постоянным и уменьшается плавно до величины, соответствующей стационарному распределению электронов при закрытом переходе (рис. II.7.7, *a*). При этом ток диода уменьшается до величины $I \approx -I_S$, а напряжение на диоде падает до установившегося значения $V \approx -E_2$.

Длительность «полочки» тока уменьшается с увеличением отношения I_2 / I_1 , а также с уменьшением времени жизни неосновных носителей в базе (диод с длинной базой) или времени их пролета через базу (диод с короткой базой).

В течение 2-й стадии запираания диода ($t > t_4$) *p-n* переход закрыт и барьерная емкость диода заряжается через сопротивление $R_2 + R_B$ с постоянной времени $C(R_2 + R_B)$.

Процесс включения диода. Пусть в исходном состоянии ключ (см. рис. II.7.6) долгое время находился в положении 2, а в момент времени $t = 0$ мгновенно перебрасывается в положение 1. При $t \leq 0$ диод закрыт и его барьерная емкость C заряжена до напряжения $V \approx -E_2$ (рис. II.7.8, *б*). После переброса ключа барьерная емкость стремится перезарядиться до напряжения $V = E_1$, причем процесс перезарядки происходит приблизительно экспоненциально с постоянной времени $\tau = \bar{C}(R_1 + R_B)$, где $\bar{C}$ — усредненное значение барьерной

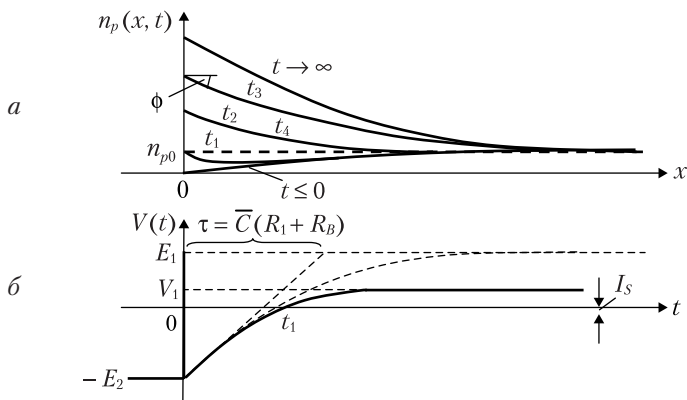


Рис. II.7.8. Процесс включения диода:

a — распределение дырок в базе; *б* — осциллограмма напряжения на *p-n* переходе

емкости. В момент $t = t_1$ диод отпирается и напряжение на нем с течением времени стремится к стационарному значению $V = V_1$. При $V \leq E_1$ ток диода после его отпирания остается практически постоянным и равным

$$I = I_1 = E_1 / (R_1 + R_B). \quad (\text{II.7.44})$$

Накопление заряда в базе диода после момента отпирания иллюстрирует рис. II.7.8, *a*, где $0 < t_1 < t_2 < t_3$. Градиент концентрации электронов на границе с переходом постоянен (постоянный угол ϕ на рис. II.7.8, *a*) и определяется величиной тока I_1 .

Таким образом, при включении диода его малое прямое сопротивление устанавливается не мгновенно, а лишь после перезаряда барьерной емкости до напряжения $V = 0$.

Из рис. II.7.7, *a* и II.7.8, *a* очевидно, что в течение переходных процессов переключения диода распределение неосновных носителей в базе существенно отличается от стационарного. Это обстоятельство объясняет различные значения диффузионной емкости, полученные на основе стационарного распределения носителей (см. параграф II.7.3) и на основе малосигнального анализа (см. параграф II.7.4).

Основные выводы

1. Инерционность диода проявляется в наличии барьерной и диффузионной емкостей, сунтирующих p - n переход.

2. Барьерная емкость связана с накоплением положительных и отрицательных зарядов примесных ионов в p - и n -областях ОПЗ при изменении напряжения. Величина барьерной емкости зависит от напряжения на переходе и обратно пропорциональна ширине перехода.

3. Диффузионная емкость связана с накоплением избыточных носителей заряда в базе. Она проявляется при прямом включении диода, причем ее величина пропорциональна току через диод (при $I \gg I_S$). В диоде с длинной базой диффузионная емкость пропорциональна времени жизни неосновных носителей заряда в базе, а в диоде с короткой базой — времени пролета носителей через базу.

4. При выключении диода его высокое обратное сопротивление устанавливается с некоторой задержкой. Эта задержка определяется временем, необходимым для рассасывания избыточных носителей заряда в базе (т.е. для заряда диффузионной емкости).

5. При включении диода его низкое прямое сопротивление также устанавливается с задержкой, определяемой временем перезаряда барьерной емкости до нулевого напряжения.



Литература

1. *Shockley, W.* Theory of p - n junctions in semiconductors and p - n junction transistors / W. Shockley // Bell Syst. Techn. J. — 1949. — Vol. 28. — P. 435.
2. *Shockley, W.* Electrons and holes in semiconductors / W. Shockley, N. Y. : D. Van Nostrand, Princeton, — 1950.
3. *Sah, C. T.*, Carrier generation and recombination in p - n junctions and p - n junction characteristics / C. T. Sah, R. N. Noyce, W. Shockley // Proc. IRE. — 1957. — Vol. 45. — P. 1228.
4. *Moll, J. I.* The evaluation of the theory of the current-voltage characteristics / J. I. Moll // Proc. IRE. — 1958. — Vol. 46. — P. 1076.
5. *Старосельский, В. И.* Расчет барьерной емкости диффузионных p - n переходов / В. И. Старосельский // Электронная техника : сер. Полупроводниковые приборы. — 1969. — Вып. 2. — С. 94–105.
6. *Sze, S. M.* Physics of Semiconductor Devices (3-rd edition) / S. M. Sze, K. Ng. Kwok. — Wiley-Interscience, 2007. — P. 815.
7. *Старосельский, В. И.* Моделирование тока генерации-рекомбинации носителей заряда в p - n переходе / В. И. Старосельский // Микроэлектроника. — 1994. — Т. 23. — № 2. — С. 58–64.



Глава III

ТРАНЗИСТОРЫ СО СТРУКТУРОЙ МЕТАЛЛ—ДИЭЛЕКТРИК— ПОЛУПРОВОДНИК

Полевой транзистор со структурой металл—диэлектрик—полупроводник (МДПТ) является основным активным прибором сверхбольших интегральных схем (СБИС). Принцип работы МДПТ впервые предложен Дж. Лилиенфельдом [1] и О. Хейлом [2] в начале 30-х гг. XX в. В конце 1940-х гг. теория МДПТ была развита У. Шокли и Дж. Пирсоном [3]. Первый кремниевый МДПТ был изготовлен в 1960 г. (Канг, Аталла, [4,5]). Характеристики МДПТ экспериментально и теоретически исследованы в работах Х. Ихантоллы и О. Молла [6], К. Са [7]. Теория МДПТ базируется на классической работе У. Шокли [8], развитой ранее для полевых транзисторов с управляющим барьерным переходом (гл. VII).

III.1. Эффект поля в структурах МДП

III.1.1. Структура МДП

Структура металл—диэлектрик—полупроводник (МДП) является основой целого ряда полупроводниковых приборов и, в частности, элементов интегральных микросхем. Схематично структура МДП представлена на рис. III.1.1, где в качестве полупроводника использован кремний p -типа проводимости. В качестве диэлектрика в кремниевых ИМС чаще всего используется SiO_2 (структура МОП), однако возможно применение и других диэлектриков (Si_3N_4 , SiN , Al_2O_3 , TiO_2 , AlN и др.). Вместо металлического слоя может быть использован слой сильно легированного вырожденного полупроводника (поликристаллический кремний n^+ - или p^+ -типа). Полупроводник имеет



достаточно большую толщину (десятки и сотни микрон) и является *подложкой*. На тыльной стороне подложки сформирован омический контакт. Толщина диэлектрика обычно лежит в диапазоне $d = 0,02 \div 0,1$ мкм. Металлический электрод называется *затвором*. Как полупроводниковый прибор структура МДП представляет собой двухполюсник с электродами затвора G (gate) и подложки B (bulk). Напряжение V между затвором и подложкой будем считать положительным, если потенциал затвора V_G выше потенциала подложки V_B :

$$V = V_G - V_B.$$



Рис. III.1.1. Структура МДП

Для определенности будем полагать, что подложка имеет p -тип проводимости и имеет нулевой потенциал (заземлена) (см. рис. III.1.1). При этом

$$V_B = 0; \quad V = V_G.$$

III.1.2. Идеальная структура МДП в условиях термодинамического равновесия

На рис. III.1.2 представлена энергетическая диаграмма идеальной структуры МДП при $V = 0$. Энергетический уровень E_0 соответствует потенциальной энергии свободного электрона. Разность между энергией E_0 и энергией электрона E_c на дне зоны проводимости называется *истинной работой выхода* из полупроводника или его *электронным сродством*

$$\psi = E_0 - E_c.$$

Электронное сродство зависит только от материала полупроводника и не зависит от степени его легирования.

Разности между энергией E_0 и энергиями Ферми F в полупроводнике и металле соответствуют *термодинамическим работам выхода* из полупроводника и металла затвора

$$\Phi_B = E_0 - F;$$

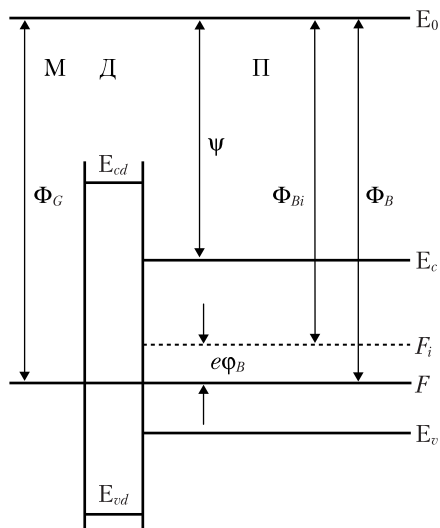


Рис. III.1.2. Энергетическая диаграмма идеальной структуры МДП при $V = 0$

$$\Phi_G = E_0 - F.$$

Термодинамическая работа выхода из полупроводника зависит от степени его легирования.

В идеальной структуре МДП удовлетворяются следующие условия.

1. Термодинамические работы выхода из полупроводника и затвора одинаковы:

$$\Phi_B = \Phi_G.$$

2. Перенос заряда через диэлектрик отсутствует при любых напряжениях V .

3. Плотность заряда в диэлектрике, а также поверхностная плотность заряда на границе диэлектрик-полупроводник равны нулю.

4. Полупроводник подложки легирован однородно.

При этих условиях все энергетические уровни на рис. III.1.2 расположены горизонтально (*плоские зоны*) и плотность заряда во всей структуре равна нулю. Подложка здесь и в дальнейшем будет считаться невырожденной; концентрации носителей заряда определяются соотношениями

$$\begin{cases} n = n_i \exp \frac{F_n - F_i}{kT}, \\ p = n_i \exp \frac{F_i - F_p}{kT}. \end{cases} \quad \begin{matrix} \text{(III.1.1a)} \\ \text{(III.1.1б)} \end{matrix}$$

В состоянии термодинамического равновесия ($V = 0$) $F_n = F_p = F$ (см. рис. III.1.2) и соотношения (III.1.1) можно представить в виде

$$\begin{cases} n_0 = n_i \exp (\mp \phi_B / \phi_T), \\ p_0 = n_i \exp (\pm \phi_B / \phi_T), \end{cases} \quad \begin{matrix} \text{(III.1.1в)} \\ \text{(III.1.1г)} \end{matrix}$$

где значение $\phi_B = |F_i - F| / e$ (см. рис. III.1.2) определяется степенью легирования полупроводника. По определению значение ϕ_B считается положительным для любого типа проводимости подложки. Таким образом, верхний из знаков «+ –» в соотношениях (III.1.1) соответствует подложке p -типа, а нижний — подложке n -типа. В дальнейшем мы будем придерживаться этого правила, ограничиваясь рассмотрением подложки p -типа.

III.1.3. Эффект поля в идеальной структуре МДП

Эффектом поля в структуре МДП называется изменение поверхностной проводимости полупроводника под действием постоянного внешнего напряжения V . При этом возможны три основные ситуации, которые поясняются рис. III.1.3, *a, б, в*.

1. Приложенное напряжение отрицательно ($V < 0$, см. рис. 1.3, *a*). Электрическое поле в диэлектрике постоянно (вследствие отсутствия зарядов) и направлено против оси x . Так как поверхностная плотность заряда на границе диэлектрик—полупроводник равна нулю, электрическая индукция $D = \epsilon E$ изменяется непрерывно. Поле в полупроводнике экранируется избыточными дырками, концентрация которых вблизи поверхности полупроводника повышается. Такой режим называется *режимом обогащения* (имеется в виду обогащение приповерхностного слоя полупроводника основными носителями заряда). На рис. III.1.3, *a* показано распределение плотности заряда ρ в структуре. Положительный заряд избыточных дырок в полупроводнике (поверхностная плотность $Q_{sp} > 0$) уравновешен отрицательным зарядом электронов на поверхности затвора (поверхностная плотность $Q_{sG} < 0$). Электрическое поле проникает в полупроводник на

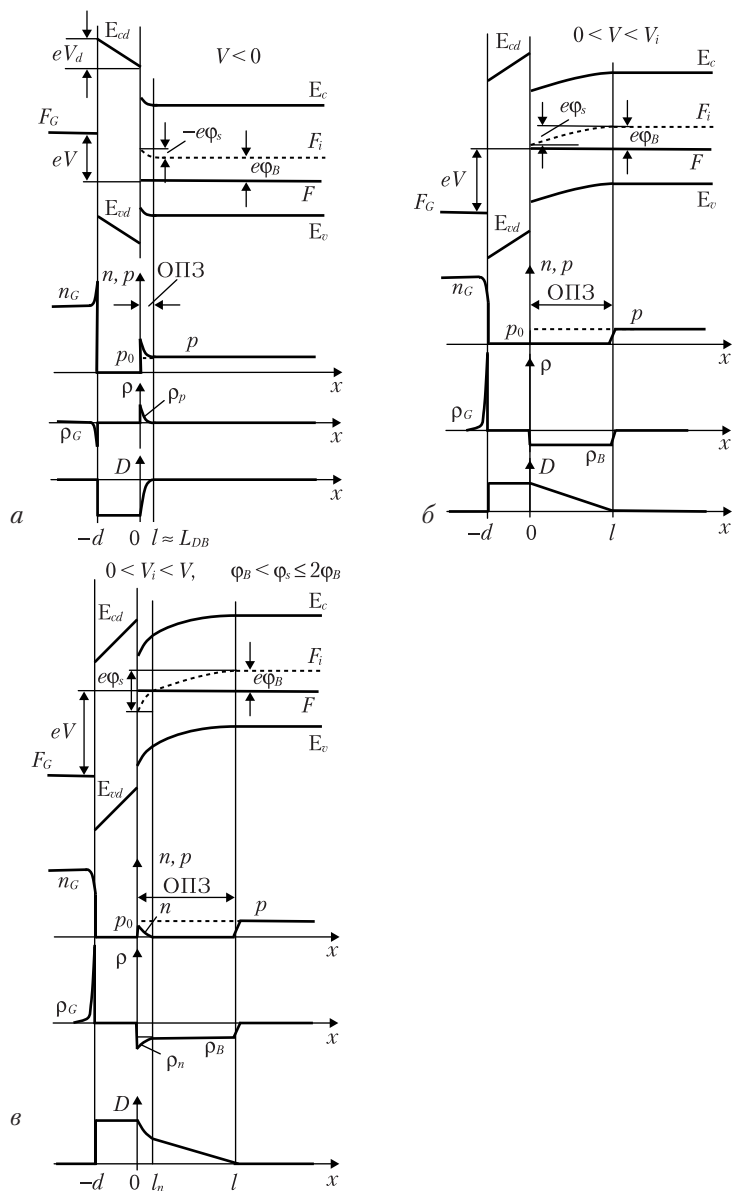


Рис. III.1.3. Энергетические диаграммы МДП-структуры и распределения концентраций носителей заряда, плотности заряда и электрической индукции:

a — режим обогащения; *b* — режим обеднения; *v* — режим инверсии

глубину порядка дебаевской длины экранирования в подложке

$$L_{DB} = \sqrt{\frac{\varepsilon \varepsilon_0 \Phi_T}{e N_B}},$$

где N_B — концентрация примеси в подложке; ε — диэлектрическая проницаемость полупроводника. Поскольку ток в структуре отсутствует, квазиуровни Ферми в полупроводнике совпадают ($F_n = F_p = F$) и располагаются горизонтально (напомним, что плотности токов пропорциональны градиентам соответствующих квазиуровней Ферми). В соответствии с соотношением (III.1.16) энергетическая диаграмма полупроводника искривляется таким образом, что уровни E_c , E_v и F_i оказываются изогнутыми вверх вблизи границы с диэлектриком (см. рис. III.1.3, а). Разность между уровнями Ферми в металле и полупроводнике пропорциональна приложенному напряжению:

$$F_G - F = eV. \quad (\text{III.1.3})$$

Потенциал поверхности полупроводника при заземленной подложке (*поверхностный потенциал* φ_s) определяется изгибом уровня F_i :

$$e\varphi_s = F_i(\infty) - F_i(0). \quad (\text{III.1.4})$$

и отрицателен (см. рис. III.1.3, а).

2. Приложенное напряжение положительно и не превышает некоторой величины V_i , называемой *напряжением инверсии*,

$$0 < V < V_i$$

(см. рис. III.1.3, б). В этом режиме энергетическая диаграмма искривляется в противоположную сторону (уровни E_c , E_v и F_i изогнуты вниз вблизи поверхности полупроводника). Соотношения (III.1.3) и (III.1.4) остаются справедливыми, причем величина поверхностного потенциала положительна и не превышает величины φ_B :

$$0 < \varphi_s < \varphi_B = (F_i - F) / e. \quad (\text{III.1.5})$$

Очевидно, что приповерхностный слой полупроводника обеднен основными носителями (*режим обеднения*). При условии (III.1.5) уровень электростатической энергии остается выше уровня Ферми $F_i > F$, поэтому концентрация неосновных носителей (электронов) весьма мала ($n < n_i$).



Если $\varphi_s < \varphi_T$ (изгиб зон меньше kT), степень обеднения приповерхностного слоя полупроводника невелика и режимы обеднения и обогащения отличаются только направлением изгиба зон и знаком избыточной концентрации дырок. Толщина обедненного слоя определяется дебаевской длиной экранирования L_{DB} . В случае $\varphi_s \gg 3\varphi_T$ ситуация качественно меняется. Приповерхностный слой практически полностью обеднен основными носителями (их концентрация много меньше, чем концентрация ионов примеси в подложке N_B), как это показано на рис. III.1.3, б. Плотность объемного заряда в обедненном слое (область пространственного заряда — ОПЗ) определяется концентрацией оголенных ионов, как это имеет место в p - n переходе (см. рис. III.1.3, б):

$$\rho_B = -eN_B.$$

Электрическое поле в ОПЗ изменяется линейно (см. рис. III.1.3, б), а толщина ОПЗ определяется соотношением

$$l = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_s}{eN_B}}. \quad (\text{III.1.6a})$$

Соотношение (III.1.6a) аналогично соотношению (II.2.25), определяющему ширину резко несимметричного p - n перехода с заменой контактной разности потенциалов φ_{BE} на поверхностный потенциал φ_s . Величина φ_s и, следовательно, толщина обедненной области l возрастают с увеличением напряжения V .

При увеличении напряжения V до некоторой величины V_i выполняется условие

$$\varphi_s(V_i) = \varphi_B.$$

В этом случае уровень Ферми на границе диэлектрик—полупроводник ($x = 0$) совпадает с уровнем электростатической энергии F_i , что соответствует равенствам

$$p(0) = n(0) = n_i.$$

Напряжение V_i называется *напряжением инверсии*.

3. Приложенное напряжение превышает напряжение инверсии ($V > V_i$, см. рис. III.1.3, в). В этом режиме $\varphi_s > \varphi_B$ и в приповерхностном слое полупроводника $0 \leq x \leq x_i$ уровень электростатической энергии расположен ниже уровня Ферми (см. рис. III.1.3, в). В соответствии с (III.1.1) в этой области концентрация неосновных носителей больше, чем основных ($n > p$), т.е. инвертируется тип проводимости подложки.



Этот режим называется *режимом инверсии*. При условии $2\varphi_B > \varphi_s > \varphi_B$ из (III.1.1) получаем

$$p_0 = N_B > n(0) > n_i > p(0).$$

Такой режим называется *режимом слабой инверсии* (поверхностная концентрация неосновных носителей $n(0)$ остается меньше равновесной концентрации основных носителей в подложке $p_0 = N_B$). В режиме слабой инверсии практически во всей ОПЗ $0 < x < l$ концентрации носителей заряда остаются много меньшими, чем в подложке, поэтому толщина ОПЗ определяется соотношением (III.1.6a).

Концентрация электронов максимальна на поверхности ($x=0$) и резко убывает при $x > 0$. Согласно (III.1.1a) снижение $n(x)$ в $e = 2,72$ раз по сравнению с $n(0)$ происходит в плоскости $x = l_n$, где потенциал полупроводника отличается от потенциала поверхности на величину φ_T . Величина l_n может быть найдена из условия

$$E_s l_n = \varphi_T,$$

где

$$E_s = E(0) = \frac{1}{\epsilon \epsilon_0} \int_0^l \rho_B dx = \sqrt{\frac{2eN_B \varphi_s}{\epsilon \epsilon_0}}$$

— напряженность поля на поверхности. Таким образом,

$$l_n = \varphi_T \sqrt{\frac{\epsilon \epsilon_0}{2eN_B \varphi_s}} \quad (\text{III.1.6б})$$

и

$$\frac{l_n}{l} = \frac{\varphi_T}{2\varphi_s} \ll 1.$$

При $\varphi_s > 2\varphi_B$ имеет место *сильная инверсия* ($n(0) > p_0 = N_B$). Условие $\varphi_s = 2\varphi_B$ выполняется при некотором значении приложенного напряжения V_{t0} , которое называется *пороговым напряжением МДП-структуры*.

При увеличении напряжения до значения V_{t0} ширина ОПЗ l возрастает в соответствии с (III.1.6a) вследствие увеличения поверхностного потенциала до $\varphi_s = 2\varphi_B$, достигая значения

$$l_{t0} = l(V_{t0}) = \sqrt{\frac{2\epsilon \epsilon_0 \cdot 2\varphi_B}{eN_B}}. \quad (\text{III.1.7})$$

Дальнейшее увеличение напряжения (переход в область сильной инверсии) не приводит к заметному расширению



ОПЗ, так как тонкий инверсионный слой экранирует ОПЗ от электростатического воздействия со стороны металлического затвора. При повышении напряжения электрическое поле увеличивается только в диэлектрике.

Толщина инверсионного слоя в режиме сильной инверсии имеет порядок дебаевской длины экранирования L_{DB} .

III.1.4. Вольтфарадные характеристики идеальной структуры МДП

Свойства структуры МДП как электрического двухполюсника (см. рис. III.1.1) подобны свойствам конденсатора. С этой точки зрения представляет интерес анализ вольтфарадной характеристики (CV -характеристики) структуры МДП, представляющей собой зависимость удельной дифференциальной емкости от напряжения

$$C_s(V) = \frac{dQ_s}{dV},$$

где $Q_s = Q_{sG} = -(Q_{sB} + Q_{sn})$, Q_{sG} , Q_{sB} и Q_{sn} — поверхностные плотности заряда на затворе, заряда ионов примеси в ОПЗ и заряда электронов в инверсионном слое соответственно.

Эквивалентная схема структуры МДП представлена на рис. III.1.4, где R_B — сопротивление подложки; $C_G = SC_{sG}$ — емкость затвора; $C_B = SC_{sB}$ — емкость полупроводникового слоя ОПЗ; C_{sG} и C_{sB} — соответствующие удельные емкости, причем

$$\begin{aligned} C_{sB} &= \epsilon \epsilon_0 / l, \\ C_{sG} &= \epsilon_d \epsilon_0 / d = \epsilon \epsilon_0 / d_{ef}, \end{aligned} \quad (III.1.8)$$

где $d_{ef} = \epsilon d / \epsilon_d$ — эффективная толщина диэлектрика; ϵ_d — его диэлектрическая проницаемость.

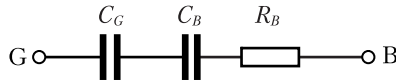


Рис III.1.4. Эквивалентная схема стуктуры МДП

Значение C_{sB} определяется толщиной обедненного или обогащенного приповерхностного слоя в полупроводнике. В режиме обогащения ($V > 0$) толщина приповерхностного слоя близка к дебаевской длине экранирования в подложке и

$$C_{sB} \approx \epsilon \epsilon_0 / L_{DB}.$$

Для точного определения зависимости $C_{sB}(V)$ необходимо найти распределения $p(x)$ и $n(x)$ путем решения уравнения Пуассона совместно с уравнениями (III.1.1). Решение можно найти в работе [9, §7.2.1]. При $\phi_s = 0$ все энергетические уровни в полупроводнике расположены горизонтально (см. рис. III.1.2). Условию $\phi_s = 0$ соответствует напряжение V_{FB} , называемое *напряжением плоских зон*. В идеальной структуре МДП $V_{FB} = 0$.

В режиме сильной инверсии $l = l_{t0} \gg L_{DB}$ и

$$C_{sB} = \epsilon \epsilon_0 / l_{t0}. \quad (\text{III.1.9})$$

Полное значение емкости структуры МДП соответствует последовательному соединению емкостей C_G и C_B (см. рис. III.1.4):

$$C_s(V) = \frac{C_{sG} C_{sB}(V)}{C_{sG} + C_{sB}(V)}. \quad (\text{III.1.10})$$

Вольтфарадная характеристика идеальной структуры МДП, измеренная на достаточно низких частотах, представлена на рис. III.1.5 (кривая *a*). При больших отрицательных напряжениях толщина обогащенного слоя мала и $C_{sB} \gg C_{sG}$, поэтому $C_s \approx C_{sG}$. При $N_B = p_0 \approx 10^{15} \text{ см}^{-3}$ значение L_{DB} в соответствии с (III.1.2) составляет около 0,1 мкм и сравнимо с толщиной диэлектрика. Поэтому полная удельная емкость плоских зон с учетом (III.1.8)—(III.1.10) составляет

$$C_{sFB} = \frac{\epsilon_d \epsilon_0 C_{sG} C_{sB}(V)}{d + L_{DB} \epsilon_d / \epsilon} < C_{sG}. \quad (\text{III.1.11})$$

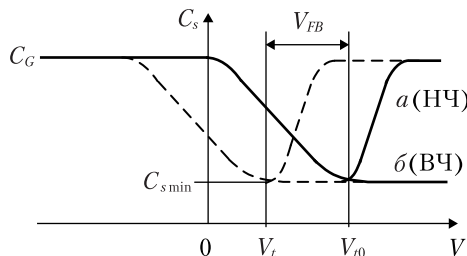


Рис III.1.5. CV-характеристики идеальной (сплошные линии) и реальной (штриховые линии) структуры МДП:

a — низкочастотная; *б* — высокочастотная

При $0 < V < V_{t0}$ толщина ОПЗ возрастает с ростом напряжения (III.1.6а), что приводит к снижению емкости C_{sB} и соответственно C_s .

При $V = V_{t0}$ емкость C_{sB} достигает минимального значения C_{sBt} , соответствующего толщине ОПЗ $l = l_{t0} \gg L_{DB}$. С учетом (III.1.7) и (III.1.9) получаем

$$C_{sBt} = \sqrt{\frac{eN_B \epsilon \epsilon_0}{4\phi_B}}; \quad (\text{III.1.12a})$$

$$C_s(V_{t0}) = C_{s\min} = \frac{C_{sG} C_{sBt}}{C_{sG} + C_{sBt}}. \quad (\text{III.1.12б})$$

При $V > V_{t0}$ изменение напряжения dV приводит только к изменению заряда dQ_{ns} в инверсионном слое (см. рис. III.1.3, в), а заряд акцепторов dQ_{sB} ОПЗ остается постоянным. Поэтому емкость полупроводника C_{sB} определяется малой толщиной слоя сильной инверсии ($C_{sB} \ll C_{sG}$) и полная емкость структуры приближается к емкости затвора.

Нарастание емкости при $V > V_{t0}$ возможно только в том случае, когда концентрация электронов в инверсионном слое успевает следовать за изменениями приложенного к структуре малого переменного напряжения dV , с помощью которого осуществляется измерение емкости. Для этого период измерительного сигнала должен быть много большим, чем время генерации и рекомбинации неосновных носителей (время жизни τ). Таким образом, вольтфарадная характеристика, соответствующая кривой *a* на рис. III.1.5, реализуется при ее измерении на частотах $f \ll \frac{1}{2} \pi \tau$.

В случае $f \gg \frac{1}{2} \pi \tau$ мы получаем высокочастотную вольтфарадную характеристику (кривая *б* на рис. III.1.5). При $V > V_{t0}$ под действием измерительного сигнала происходит изменение заряда акцепторов вследствие изменения на величину dV ширины ОПЗ (как в *p-n* переходе). При этом заряд инверсионного слоя не успевает измениться за период сигнала и емкость полупроводника определяется соотношениями (III.1.12).

III.1.5. Особенности эффекта поля в реальной структуре МДП

В реальной структуре МДП условия 1—3 (см. параграф III.1.1), соответствующие идеальной структуре, обычно не выполняются. Различие термодинамических работ выхода из металла и полупроводника приводит к тому, что энергетиче-



ские зоны в полупроводнике оказываются изогнутыми даже при $V = 0$. В результате CV -характеристики (см. рис. III.1.5) сдвигаются по оси напряжения, а напряжение плоских зон сдвигается от нуля на величину контактной разности потенциалов металл—полупроводник

$$\Phi_{GB} = (\Phi_G - \Phi_B) / e. \quad (\text{III.1.13})$$

К тому же эффекту приводит наличие зарядов в диэлектрике и на границе диэлектрик—полупроводник. Поверхностная плотность заряда Q_{ss} на этой границе индуцирует равную по модулю поверхностную плотность заряда на поверхности металла и создает в диэлектрике встроенное электрическое поле. С учетом (III.1.8) напряженность встроенного поля составляет

$$E_{ss} = Q_{ss} / \epsilon_d \epsilon_0 = Q_{ss} / C_{sG} d.$$

Это приводит к сдвигу напряжения плоских зон на величину $E_{ss} d$:

$$V_{FBss} = -Q_{ss} / C_{sG}. \quad (\text{III.1.14})$$

Заряд в диэлектрике с объемной плотностью $\rho_d(x)$ создает в нем встроенное поле с напряженностью

$$E_p(x) = \frac{1}{\epsilon_d \epsilon_0} \int_{-d}^x \rho_d(x) dx.$$

Интегрируя это поле по всей толщине диэлектрика, получаем соответствующий сдвиг напряжения плоских зон

$$V_{FBp} = - \int_{-d}^0 E_p(x) dx = \frac{1}{\epsilon_d \epsilon_0} \int_{-d}^0 dx \int_{-d}^x \rho_d(x) dx = \frac{1}{\epsilon_d \epsilon_0} \int_{-d}^0 x \rho_d(x) dx. \quad (\text{III.1.15})$$

Последнее выражение получено интегрированием по частям.

Результирующий сдвиг напряжения плоских зон (см. рис. III.1.5) с учетом (III.1.8) и (III.1.12) составляет

$$V_{FB} = \Phi_{GB} + V_{FBss} + V_{FBp} = \Phi_{GB} - Q_{ss ef} / C_{sG}, \quad (\text{III.1.16})$$

где
$$Q_{ss ef} = Q_{ss} + \int_{-d}^0 x \rho_d(x) dx \quad (\text{III.1.17})$$

— эффективное значение поверхностной плотности заряда на границе диэлектрик—полупроводник.

Поверхностный заряд Q_{ss} частично связан с поверхностными состояниями, локализованными на границе диэлектрик—



полупроводник, энергетические уровни которых расположены в глубине запрещенной зоны полупроводника (быстрые поверхностные состояния). Эти поверхностные состояния могут перезаряжаться при изменении поверхностного потенциала ϕ_s , что приводит к зависимости эффективной емкости C_{sg} от напряжения. Другая составляющая Q_{ss} связана с фиксированным зарядом в тонком (1–5 нм) промежуточном слое диэлектрика вблизи границы диэлектрик–полупроводник, который отличается по свойствам от основного объема диэлектрика.

Объемный заряд в диэлектрике $\rho_d(x)$ связан с дефектами различного рода, а также с наличием примесных ионов. Последние (например, ионы Na^+ в SiO_2) могут перемещаться при стрессовых термополевых нагрузках, что также ведет к зависимости C_{sg} от напряжения. В дальнейшем этой зависимостью мы пренебрежем. Неоднородность легирования подложки также приводит к изменению напряжения плоских зон и к некоторому изменению вида CV -характеристики. Эти эффекты будут рассмотрены ниже.

Условие малости сквозного тока через диэлектрик обычно выполняется достаточно хорошо за исключением случаев очень тонкого диэлектрика и больших напряжений. Механизмы переноса заряда через диэлектрик рассмотрены в работе [9, §7.3.4]. При больших напряжениях может произойти пробой диэлектрика. Критическая напряженность электрического поля в диэлектрике, соответствующая пробоем, составляет около 10^7 В/см.

III.1.6. Пороговое напряжение структуры МДП

Пороговое напряжение является одним из основных параметров структуры МДП, а также полевых транзисторов, созданных на ее основе. На рис. III.1.6 представлена энергетическая диаграмма реальной структуры МДП, а также распределения плотности заряда $\rho(x)$, электрической индукции

$$D(x) = \frac{1}{\epsilon_0} \int_x^l x \rho(x) dx \text{ и потенциала } \phi(x) \text{ при } V = V_{t0}.$$

Штриховая линия показывает реальное распределение плотности объемного заряда в диэлектрике $\rho_d(x)$. В соответствии с (III.1.17) его можно заменить эффективным значением поверхностной плотности заряда Q_{sef} в плоскости $x = 0$. При



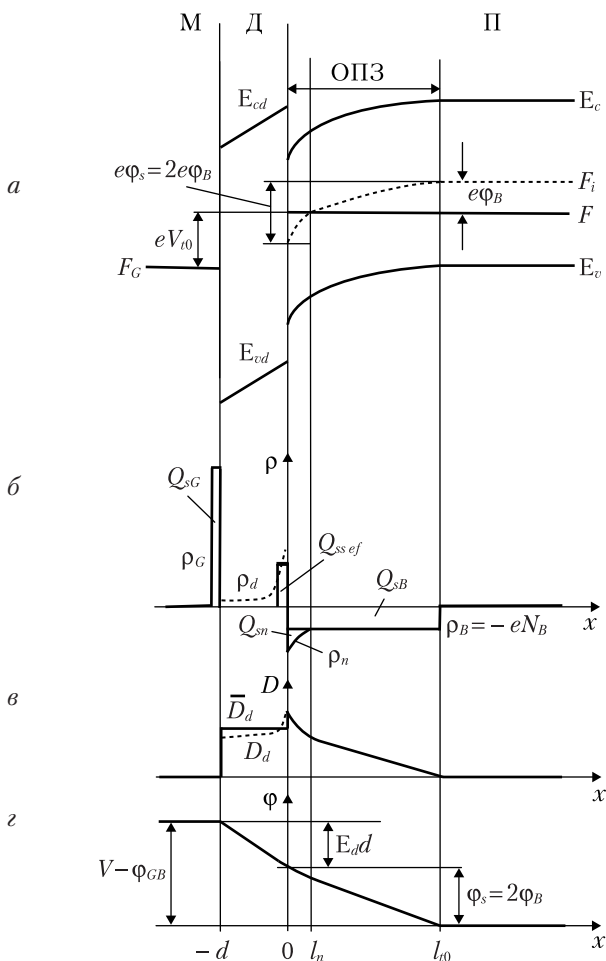


Рис. III.1.6. Энергетическая диаграмма реальной структуры МДП при $V = V_{t0}$ (а) и распределения плотности объемного заряда (б), электрической индукции (в) и потенциала (г)

этом реальное распределение поля в диэлектрике (пунктир $D_d(x)$ на рис. III.1.6, в) заменяется постоянным полем $\bar{D}_d = \epsilon_d E_d$. В плоскости $x=0$ функция $D_d(x)$ терпит разрыв на величину $Q_{ss\ ef} / \epsilon_0$.

Из рис. III.1.6, а, г следует, что приложенное внешнее напряжение V может быть представлено в виде

$$V = \varphi_{GB} + \varphi_s + \bar{E}_d d,$$

где контактная разность потенциалов φ_{GB} определяется соотношением (III.1.13), а средняя напряженность электрического поля в диэлектрике составляет

$$\bar{E}_d = Q_{sG} / \epsilon_d \epsilon_0.$$

С учетом суммарной электрической нейтральности структуры $Q_{sG} + Q_{ssef} + Q_{sB} + Q_{sn} = 0$ получаем

$$V = \varphi_{GB} + \varphi_s - (Q_{ssef} + Q_{sB} + Q_{sn}) / C_{sG}. \quad (\text{III.1.18})$$

При пороговом напряжении $V = V_{t0}$ поверхностный потенциал составляет $\varphi_s = \pm 2\varphi_B$ (см. параграф III.1.4), а поверхностная (при $x = 0$) концентрация неосновных носителей (электронов) в полупроводнике равна концентрации примеси N_B (см. рис. III.1.3, б). Толщина инверсионного слоя ($\sim L_{DB}$) много меньше пороговой ширины ОПЗ l_{t0} (см. параграф III.1.3), поэтому в (III.1.18) $Q_{sn} \ll Q_{sB}$. С учетом этих соображений из (III.1.18) получаем

$$V_{t0} = \varphi_{GB} \pm 2\varphi_B - [Q_{ssef} + Q_{sB0}] / C_{sG}. \quad (\text{III.1.19a})$$

Напомним, что в слагаемом $\pm 2\varphi_B$ верхний знак соответствует подложке p -типа, нижний — n -типа. Сравнение (III.1.19a) и (III.1.16) показывает, что пороговое напряжение складывается из напряжения плоских зон, поверхностного потенциала $\varphi_s(V_{t0}) = \pm 2\varphi_B$, а также напряжения на емкости затвора C_{sG} от заряда ионов ОПЗ с поверхностной плотностью Q_{sB0} .

Значение φ_{GB} определяется соотношением (III.1.13), где в соответствии с рис. III.1.2

$$\Phi_B = \Phi_{Bi} \pm e\varphi_B,$$

где Φ_{Bi} — термодинамическая работа выхода из собственного полупроводника. Таким образом,

$$V_{t0} = \varphi_{GBi} \pm \varphi_B - [Q_{ssef} + Q_{sB0}] / C_{sG}, \quad (\text{III.1.19б})$$

где

$$\varphi_{GBi} = (\Phi_G - \Phi_{Bi}) / e \quad (\text{III.1.20})$$

— контактная разность потенциалов металл—собственный полупроводник. Значения φ_{GBi} для некоторых структур металл—диэлектрик—кремний приведены в табл. III.1.1.



Таблица III.1.1

**Значения контактной разности потенциалов
металл—*i*-Si ($T = 300$ K)**

Материал затвора	Φ_{GB^i} , В	Материал затвора	Φ_{GB^i} , В
Al, Ag	– 0,30	W	0,00
Au	+ 0,53	Pt	+ 1,30
Ni	+ 0,70	n^+ -Si*	– 0,56
Ti	– 0,65	p^+ -Si*	+ 0,56

Если в качестве затвора используется вырожденный полукремний (Si*), то уровень Ферми в затворе практически совпадает с одной из границ запрещенной зоны. При этом

$$\Phi_G \approx \Phi_{Bi} \pm E_g / 2,$$

где $E_g = 1,12$ эВ — ширина запрещенной зоны кремния, верхний знак (+) соответствует подложке n -типа, нижний — подложке p -типа.

Значение Φ_B определяется соотношениями (III.1.1в, г):

$$\Phi_B = \Phi_T \ln(N_B / n_i). \quad (\text{III.1.21})$$

Обычно концентрация примеси в подложке лежит в диапазоне $N_B = (10^{15} - 10^{17}) \text{ см}^{-3}$. При этом

$$\Phi_B = (0,30 - 0,42) \text{ В}.$$

Значение Q_{sef} зависит от материала диэлектрика, его толщины, качества поверхности диэлектрик—полупроводник и ориентации подложки. Вместо параметра Q_{sef} принято использовать эффективную плотность поверхностных состояний

$$N_{sef} = Q_{sef} / e. \quad (\text{III.1.22})$$

Для случая, когда в качестве диэлектрика используется термический окисел SiO₂ (структура МОП — металл—оксид—полупроводник), типичные значения N_{sef} приведены в табл. III.1.2.

Таблица III.1.2

Типичные значения N_{sef} для поверхности Si—SiO₂

Ориентация подложки,	N_{sef} , см ⁻²
< 100 >	+ 9 · 10 ¹⁰
< 110 >	+ 2 · 10 ¹¹
< 111 >	+ 5 · 10 ¹¹



Значение параметра Q_{sB0} для случая однородно легированной подложки ($N_B(x) = \text{const}$) составляет

$$Q_{sB0} = \mp e N_B l_{t0} = \mp \sqrt{2\epsilon\epsilon_0 \cdot 2\phi_B \cdot e N_B}, \quad (\text{III.1.23})$$

где для определения толщины ОПЗ l_{t0} использовано соотношение (III.1.7).

Практический интерес представляет случай, когда полупроводник подложки дополнительно легирован с поверхности (рис. III.1.7). При этом результирующая концентрация примеси составляет

$$N_B(x) = N_B + \Delta N_B(x),$$

где $\Delta N_B(x) \xrightarrow{x \rightarrow \infty} 0$.

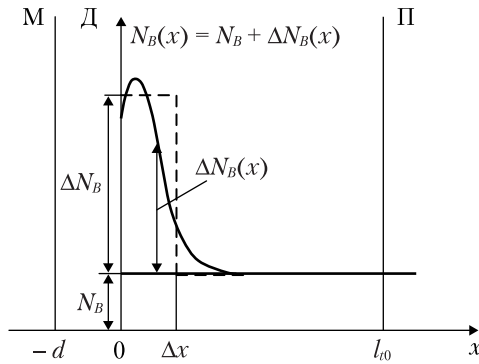


Рис. III.1.7. Примесный профиль МДП-структуры при поверхностном подлегировании подложки

Если эффективная толщина подлегированного слоя мала по сравнению с пороговой толщиной ОПЗ (на рис. III.1.7 $\Delta x \ll l_{t0}$), его влияние на пороговое напряжение V_{t0} сводится к изменению эффективного значения поверхностной плотности заряда $Q_{s\text{ef}}$ на величину

$$\Delta Q_{s\text{ef}} = e \Delta D_B,$$

где $\Delta D_B = \int_0^{\infty} \Delta N_B(x) dx$ — поверхностная доза подлегивающей примеси. При этом согласно (III.1.19) изменение порогового напряжения составляет

$$\Delta V_{t0B} = \Delta Q_{s\text{ef}} / C_{sd} = \pm e \Delta D_B / C_{sG}, \quad (\text{III.1.24})$$

где знак «+» соответствует донорной примеси, знак «-» — акцепторной.

Подлегирирование поверхности (обычно методом ионной имплантации) применяется при необходимости для точной регулировки порогового напряжения.

Свойства структуры МДП при разных режимах работы для полупроводника p -типа сведены в табл. III.1.3. Для полупроводника n -типа следует изменить знаки неравенств для напряжений и поверхностных зарядов, а также осуществить замены $n \rightleftharpoons p$.

Основные выводы

1. Эффект поля в структуре МДП состоит в изменении величины поверхностной проводимости полупроводника и, возможно, ее типа под действием приложенного напряжения.

2. Изменение приложенного напряжения V вызывает изменение поверхностного потенциала полупроводника ϕ_s , который определяется как изменение потенциала в ОПЗ. При напряжении плоских зон $V = V_{FB}$ полупроводник остается однородным во всем объеме до поверхности. При напряжении инверсии $V = V_i$ концентрации электронов и дырок на поверхности одинаковы: $n(0) = p(0) = n_i$. При пороговом напряжении $V = V_{t0}$ в приповерхностной области возникает инверсионный слой, причем поверхностная концентрация неосновных носителей равна концентрации основных носителей в объеме полупроводника и возрастает при $V > V_{t0}$ для подложки p -типа, при $V < V_{t0}$ для подложки n -типа.

3. В зависимости от приложенного напряжения можно выделить три основных режима работы: обеднения, обогащения и инверсии. Свойства структуры МДП при разных режимах работы сведены в табл. III.1.3.

4. Напряжение плоских зон зависит от контактной разности потенциалов металл—полупроводник, эффективной плотности поверхностного заряда на границе диэлектрик—полупроводник, а также от удельной емкости диэлектрика и определяется соотношением (III.1.16).

5. Пороговое напряжение складывается из напряжения плоских зон, поверхностного потенциала $\phi_s(V_{t0}) = \pm 2\phi_B$ и напряжения на емкости диэлектрика от заряда ионов ОПЗ с поверхностной плотностью $Q_{sB}(V_{t0})$, а также заряда поверхностных состояний и заряда в диэлектрике с эффективной поверхностной плотностью $Q_{ss\,ef}$. Величина порогового напряжения для однородно легированного полупроводника определяется соотношениями (III.1.19)—(III.1.23), (III.1.7).

6. Точная регулировка порогового напряжения осуществляется подлегирированием тонкого поверхностного слоя полупроводника.



Таблица III.13

Свойства структуры МДП (полупроводник p -типа) в различных электрических режимах

Режим	Поверхностный потенциал φ_s	Поверхностные концентрации носителей	Ширина ОПЗ l	Заряды в ОПЗ	Удельная емкость структуры
Обогащение ($V < V_{FB}$)	$\varphi_s < 0$	$p(0) > N_B$; $n(0) < n_0 < n_i$	$l \approx l_{DB}$	$Q > 0$; $Q_{sp} \approx 0$; $Q_{sn} = 0$ $Q_{sB} = 0$	$C_{sG} > C_s > C_{sFB}$
Состояние плоских зон ($V = V_{FB}$)	$\varphi_s = 0$	$p(0) = p_0 = N_B$; $n(0) = n_0 = n_i^2 / N_B$	$l = 0$	$Q = 0$; $Q_{sp} = 0$; $Q_{sn} = 0$ $Q_{sB} = 0$	$C_s = C_{sFB}$
Обеднение ($V > V_{FB}$)	$0 < \varphi_s < \varphi_B$	$n_i < p(0) < N_B$; $n_0 < n(0) < n_i$	$l \gg l_{DB}$, растет с ростом V	$Q \approx 0$; $Q_{sp} \approx 0$; $Q_{sn} < 0$ $Q_{sB} < 0$	$C_{sFB} > C_s$, снижается с ростом V
Слабая инверсия ($V_{r0} > V > V_i$)	$\varphi_B < \varphi_s < 2\varphi_B$	$p(0) < n_i$; $N_B > n(0) > n_i$	$l_{r0} > l \gg l_{DB}$	$Q \approx 0$; $Q_{sp} \approx 0$; $Q_{sn} < 0$ $Q_{sB} < 0$	$C_{sFB} > C_s > C_{sG0}$
Пороговый режим ($V = V_{r0}$)	$\varphi_s = 2\varphi_B$	$p(0) < n_i$; $n(0) = N_B$	$l_{r0} = l \gg l_{DB}$	$Q \approx 0$; $-Q_{sn} \ll -Q_{sB} = eN_B l_{r0}$	$C_s = C_{smin}$
Сильная инверсия ($V > V_{r0}$)	$\varphi_s > 2\varphi_B$	$p(0) \ll n_i$; $n(0) > N_B$	$l_{r0} = l \gg l_{DB}$	$Q \approx 0$; $-Q_{sn} < 0$; $Q_{sB} = -eN_B l_{r0}$	При $f \ll 1/2 \pi \tau$: $C_s = C_{sG}$; при $f \gg 1/2 \pi \tau$: $C_s = C_{smin}$



III.2. Устройство и принцип действия МДП-транзистора

III.2.1. Принцип действия МДПТ

Эффект поля в структуре МДП лежит в основе действия одного из важнейших полупроводниковых приборов — МДП-транзистора (МДПТ). В разделе III.1 показано, что в режиме сильной инверсии приповерхностная область полупроводника в структуре МДП изменяет тип проводимости, причем поверхностная концентрация неосновных носителей превышает концентрацию основных носителей в объеме. В МДПТ эта поверхностная область называется *каналом*. Тип проводимости канала противоположен типу проводимости подложки.

Для определенности в дальнейшем мы будем рассматривать МДП-транзисторы с n -каналом за исключением специально оговоренных случаев. Как показано в параграфе III.1.3, толщина канала h_c имеет порядок дебаевской длины экранирования в подложке L_{DB} . Заметим, что канал электрически изолирован от подложки обедненной областью пространственного заряда (ОПЗ, рис. III.1.6). Поверхностная проводимость канала зависит от поля в диэлектрике и, следовательно, от напряжения между затвором и каналом. Поэтому сопротивление канала в направлении, параллельном границе с диэлектриком, зависит от потенциала затвора. Изменяя потенциал затвора, можно управлять током в цепи канала при постоянном напряжении, приложенном к каналу. Таким образом, в МДП-транзисторе канал является управляемым резистором, сопротивление которого контролируется потенциалом затвора относительно одной из границ канала.

Ток в канале МДП-транзистора переносится главным образом неосновными для подложки носителями заряда. Поэтому МДП-транзисторы относятся к классу *униполярных* транзисторов, в которых принципиальное значение имеет поведение только одного типа носителей.

III.2.2. Устройство и разновидности МДПТ

Простейшая структура МДП-транзистора с n -каналом изображена на рис. III.2.1. Активная часть прибора представляет собой структуру МДП: затвор (G) — подзатворный диэлектрик толщиной d — подложка (B). Канал образуется

между специально сформированными n^+ -областями стока (D) и истока (S), на поверхности которых расположены омические контакты. Области стока и истока могут быть сформированы, например, с помощью селективной диффузии доноров или селективной ионной имплантации через окна в диэлектрике.

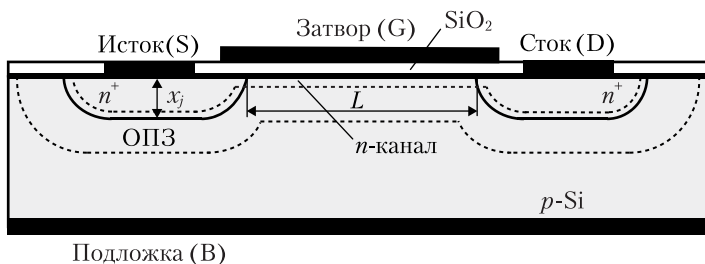


Рис. III.2.1. Простейшая структура n -канального МДПТ

В качестве диэлектрика в кремниевых ИС обычно используется термический оксид кремния SiO_2 . В этом случае прибор называется МОП-транзистором (металл—оксид—полупроводник). Четвертым электродом является электрод подложки В. Как очевидно из рисунка, МДП-транзистор является симметричным прибором: электроды стока и истока взаимозаменяемы.

Электрический режим работы транзистора определяется следующими внешними напряжениями:

$$V_{GS} = V_G - V_S \text{ — напряжение затвор—исток;}$$

$$V_{GD} = V_G - V_D \text{ — напряжение затвор—сток;}$$

$$V_{DS} = V_D - V_S = V_{GS} - V_{GD} \text{ — напряжение сток—исток;}$$

$$V_{BS} = V_B - V_S \text{ — напряжение подложка—исток,}$$

где V_G, V_S, V_D и V_B — потенциалы затвора, истока, стока и подложки, соответственно.

По типу проводимости канала МДП-транзисторы делятся на n -канальные и p -канальные. В p -канальных транзисторах все слои полупроводника имеют обратный тип проводимости.

Из двух электродов стока и истока стоком считается электрод с более высоким потенциалом: всегда $V_{DS} \geq 0$ (для p -канального МДПТ $V_{DS} \leq 0$).

При $V_{DS} = 0$ канал существует при $V_{GS} > V_t$ и отсутствует при $V_{GS} < V_t$, где V_t — пороговое напряжение МДПТ. Понятие порогового напряжения уточняется ниже; временно будем

считать, что пороговое напряжение МДПТ совпадает с пороговым напряжением структуры МДП, определенным в параграфе III.1.3. В n -канальном транзисторе при $V_t < 0$ канал является *встроенным*, т.е. существует в состоянии равновесия. При $V_t > 0$ канал является *индуцированным*, т.е. создается только при напряжении $V_{GS} > V_t > 0$. В p -канальном транзисторе канал является встроенным при $V_t > 0$ и индуцированным при $V_t < 0$.

Длина канала L определяется как расстояние между металлургическими переходами сток—подложка и исток—подложка на поверхности кристалла (см. рис. III.2.1).

В нормальном режиме работы МДПТ $V_{BS} \leq 0$ (для p -канального МДПТ $V_{BS} \geq 0$). При этом p - n переходы исток—подложка, сток—подложка и канал—подложка оказываются закрытыми, что обеспечивает взаимную изоляцию МДП транзисторов в ИС.

Таким образом, МДП-транзисторы различаются по типу проводимости канала (n - и p -канальные), а также по знаку порогового напряжения — со встроенным каналом (*нормально открытые*) и с индуцированным каналом (*нормально закрытые*). Обозначения МДПТ различного типа в электрических схемах приведены на рис. III.2.2.

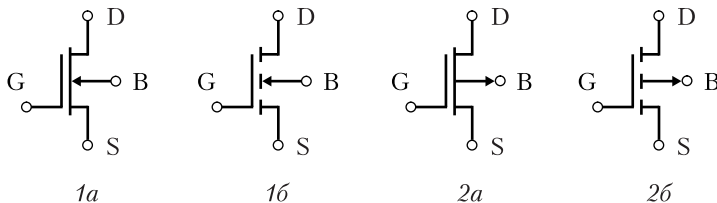


Рис. III.2.2. Схемные обозначения МДПТ:

- 1а — n -канальный со встроенным каналом ($V_{DS} > 0$, $V_t < 0$);
- 1б — n -канальный с индуцированным каналом ($V_{DS} > 0$, $V_t > 0$);
- 2а — p -канальный со встроенным каналом ($V_{DS} < 0$, $V_t > 0$);
- 2б — p -канальный с индуцированным каналом ($V_{DS} < 0$, $V_t < 0$)

Схемотехнические возможности МДПТ значительно расширяются при использовании в рамках единой ИС как n -канальных, так и p -канальных транзисторов. Простейшая структура такой *комплементарной* транзисторной пары показана на рис. III.2.3. n -канальный МДПТ сформирован в p -подложке 1 с концентрацией акцепторов $\approx 10^{15} \text{ см}^{-3}$ так же, как в структуре на рис. III.2.1 (n^+ -области стока и истока 2). Для формирования p -канального МДПТ создается специальный n -карман 3 с концентрацией донорной примеси $\approx 10^{17} \text{ см}^{-3}$, который является

подложкой p -канального МДПТ. Области его стока и истока формируются в виде ионно-имплантированных p^+ -областей 4. Контакт к n -карману 3 выполнен в виде n^+ -области 5. Нанесение диэлектрика 7 (SiO_2) на слаболегированную поверхность p -подложки 1 приводит к искривлению энергетических уровней таким образом, что на поверхности образуются паразитные инверсионные области — n -каналы, которые могут соединять n^+ -области стока и истока n -канальных МДПТ между собой, а также с n -карманами p -канальных МДПТ. Для предотвращения этого вредного эффекта вокруг p -канальных МДПТ формируются специальные противоканальные p^+ -области 6. Эти области могут также использоваться для создания контактов к общей p -подложке 1.

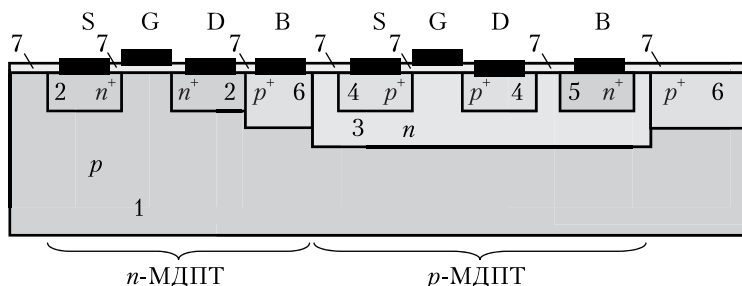


Рис. III.2.3. Структура комплементарной пары МДПТ:

- 1 — p -подложка; 2 — n^+ -области стока и истока n -МДПТ;
- 3 — n -карман; 4 — p^+ -области стока и истока p -МДПТ;
- 5 — контактная область к n -карману (подложке p -МДПТ);
- 6 — противоканальные p^+ -области; 7 — SiO_2

III.2.3. Качественный анализ характеристик МДПТ

МДП-транзистор представляет собой четырехполюсный прибор (исток, сток, затвор, подложка), однако в большинстве практических случаев электроды подложки и истока объединяются ($V_{BS} = 0$). Наилучшие усилительные свойства МДПТ проявляются в случае, когда входным электродом является затвор, а выходным — сток. При этом входным напряжением является напряжение V_{GS} , а выходным током — ток стока I_D . Поскольку входной ток (ток затвора I_G) практически равен нулю, статические ВАХ описываются функциями $I_D(V_{GS}, V_{DS})$ или $I_D(V_{GS}, V_{GD})$.

Графически статические характеристики ВАХ МДПТ принято представлять *выходными ВАХ* — семейством кривых $I_D(V_{DS})$ с параметром V_{GS} , или *проходными ВАХ* — семейством кривых $I_D(V_{GS})$ с параметром V_{DS} .

Проведем вначале качественный анализ ВАХ n -канального МДП-транзистора, включенного по схеме с общим истоком. На рис. III.2.4 схематично представлена структура транзистора и показаны границы ОПЗ (штриховые линии) и канала для различных напряжений V_{DS} . Заметим сразу, что границы канала изображены условно, так как концентрация электронов в канале уменьшается в направлении оси x непрерывно и канал (область сильной инверсии) не имеет четкой границы. «Толщина» канала на рис. III.2.4 соответствует его поверхностной проводимости, которая повышается с увеличением концентрации электронов в канале и, следовательно, с увеличением напряжения V_{GS} .

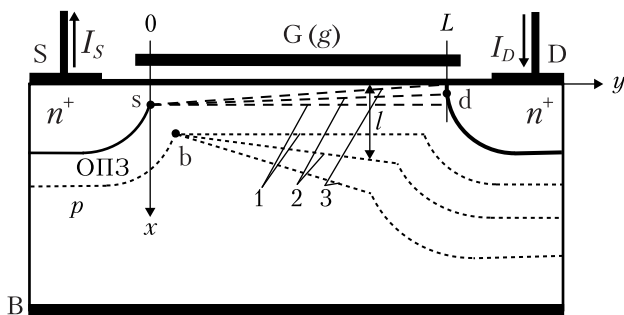


Рис. III.2.4. Границы ОПЗ-канал (штриховые линии) и ОПЗ-затвор (пунктир) при $V_{gs} > V_t$ и различных напряжениях V_{ds} :

1 — $V_{ds1} = 0$; 2 — $0 < V_{ds2} < V_{ds s}$; 3 — $V_{ds3} \geq V_{ds s}$

Контактные n^+ -области стока и истока, а также области подложки и затвора имеют некоторые сопротивления, на которых при протекании токов выделяются напряжения. Поэтому потенциалы V_D , V_S , V_G и V_B электродов затвора, истока, стока и подложки отличаются от потенциалов «внутренних» точек, находящихся непосредственно на границе с каналом, диэлектриком и ОПЗ. На рис. III.2.4 показаны «внутренние» точки стока (d), истока (s) и подложки (b), а также (условно) затвора (g). Напряжения между этими «внутренними» точками V_{gs} , V_{gd} , V_{ds} и V_{bs} также отличаются от «внешних» напряжений V_{GS} , V_{GD} , V_{DS} и V_{BS} между электродами транзи-

стора. Эти различия будут учтены позже при синтезе эквивалентных схем МДПТ. Пока же мы временно будем говорить о характеристиках «внутреннего» транзистора, используя для напряжений обозначения V_{gs} , V_{gd} , V_{ds} и V_{bs} .

Выходные ВАХ МДПТ для фиксированного напряжения $V_{bs} = 0$ изображены на рис. III.2.5. При $V_{bs} = 0$ ток через p - n переход, образованный между n^+ -областью истока и подложкой, равен нулю. Поскольку потенциал стока выше потенциала истока, p - n переход между n^+ -областью стока и подложкой закрыт. Далее мы пренебрежем весьма малым обратным током этого перехода, а также током через ОПЗ между каналом и подложкой. При этом в стационарном состоянии ток через электрод подложки $I_B = 0$, и ток стока равен току истока (на рис. III.2.4 $I_D = I_S$).

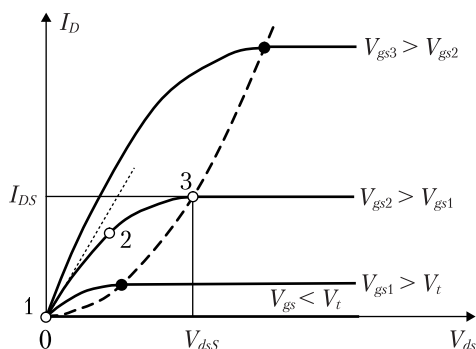


Рис. III.2.5. Выходные ВАХ МДПТ при $V_{bs} = 0$.

Штриховая линия — граница крутой и пологой областей ВАХ

Рассмотрим вначале случай, когда напряжение затвористок V_{gs} превышает пороговое напряжение ($V_{gs} = V_{gs2} > V_t$ на рис. III.2.5). При $V_{ds} \rightarrow 0$ (точка 1 на рис. III.2.5) канал является эквипотенциальным и, следовательно, однородным по всей длине L (штриховая линия 1 на рис. III.2.4). Ширина ОПЗ l , границы которой изображены пунктиром на рис. III.2.4, не зависит от координаты y . Поверхностная проводимость канала

$$\sigma_{Cs}(y) = e \int_0^{h_c(y)} \mu_n n(x, y) dx, \quad (\text{III.2.1})$$

(μ_n — подвижность электронов в канале) зависит только от напряжения $V_{gs} = V_{gd}$ и не зависит от y . Поэтому ток в цепи

сток-исток определяется законом Ома (увеличивается пропорционально напряжению V_{ds} , точка 1 и пунктирная прямая на рис. III.2.5).

При увеличении напряжения V_{ds} потенциал канала возрастает по направлению оси x на рис. III.2.4. Ширина ОПЗ зависит от напряжения между каналом и подложкой, поэтому ОПЗ расширяется по направлению от истока к стоку (пунктир 2 на рис. III.2.4). Канал также становится существенно неоднородным — его поверхностная проводимость уменьшается по направлению к стоку вследствие уменьшения напряжения затвор—канал (штриховая линия 2 на рис. III.2.4). Поэтому сопротивление канала возрастает с ростом напряжения V_{ds} и ВАХ становится нелинейной (участок между точками 1 и 2 на рис. III.2.5).

Когда напряжение сток—исток достигает значения

$$V_{ds} = V_{dsS} = V_{gs} - V_t \quad (\text{III.2.2})$$

(точка 3 на рис. III.2.5), напряжение затвор—сток становится равным пороговому: $V_{gd} = V_{gdS} = V_t$. При этом поверхностная проводимость канала на границе со стоком ($x = L$) становится равной нулю

$$\sigma_c(L) = e \int_0^{h_c(y)} \mu_n n(x, y) dx = 0 \quad (\text{III.2.3})$$

и происходит *перекрытие канала* на границе со стоком (штриховая линия 3 на рис. III.2.4). Дифференциальное сопротивление канала $\partial V_{ds} / \partial I_D$ становится бесконечным, и при дальнейшем увеличении напряжения V_{ds} ток стока остается постоянным и равным максимальному для заданного напряжения V_{gs} значению I_{DS} (см. рис. III.2.5).

Напряжение V_{dsS} называется *напряжением насыщения*, а ток I_{DS} — *током насыщения*. Значение I_{DS} зависит только от напряжения V_{gs} (вопрос о возможности протекания тока при перекрытом канале будет обсуждаться в параграфе III.3.2).

При увеличении напряжения V_{gs} начальное сопротивление канала (при $V_{ds} \rightarrow 0$) уменьшается. Поэтому выходная ВАХ при $V_{gs} = V_{gs3} > V_{gs2}$ расположена выше кривой при $V_{gs} = V_{gs2}$ (см. рис. III.2.5) и ток I_{DS} возрастает. Напряжение насыщения V_{dsS} также повышается в соответствии с (III.2.2). При уменьшении напряжения V_{gs} напряжение насыщения V_{dsS} и ток насыщения I_{DS} снижаются.

Участок ВАХ между точками 1 и 3 на рис. III.2.5 ($0 \leq V_{ds} \leq V_{dsS}$) называется *крутой областью* ВАХ, а участок



правее точки 3 — *пологой областью* ВАХ. Крутая и пологая области ВАХ разделены на рис. III.2.5 штриховой линией.

При напряжении затвор—канал, меньшем порогового напряжения ($V_{gs} < V_t$), ток в цепи сток—исток отсутствует ($I_D \approx 0$, см. рис. III.2.5). Эта область ВАХ называется *областью отсечки* тока.

В отличие от простой структуры МДП, рассмотренной в разделе III.1, МДП-транзистор имеет дополнительные электроды стока и истока. Введение этих электродов приводит к двум принципиальным различиям в физических свойствах этих приборов.

1. Образование или исчезновение инверсионного слоя в структуре МДП под действием приложенного напряжения происходит за счет процессов генерации и рекомбинации носителей заряда в приповерхностной области подложки. Длительность этих процессов имеет порядок времени жизни неосновных носителей заряда τ в подложке, величина которого довольно значительна ($\tau = 10^{-5} - 10^{-7}$ с). С этим, в частности, связаны различия в CV -характеристиках, измеренных на низких и высоких частотах (см. параграф III.1.4).

В n -канальном МДПТ при изменении напряжения затвор—канал от значения $V_{gs} < V_t$ до значения $V_{gs} > V_t$ (включение МДПТ) канал образуется за счет инжекции в приповерхностную область p -подложки электронов из n^+ -областей истока и (если напряжение V_{ds} мало) стока. При выключении МДПТ при малом напряжении V_{ds} канал исчезает за время экстракции электронов из канала в n^+ -области стока и истока под действием диффузии, а при значительном напряжении V_{ds} — за счет их вытягивания в сток электрическим полем. Длительности этих процессов значительно меньше и в зависимости от длины канала L могут составлять $10^{-9} - 10^{-12}$ с. Благодаря этому возможно создание МДПТ очень высокого быстродействия.

2. Пороговое напряжение МДПТ отличается от порогового напряжения МДП-структуры, если напряжение исток—подложка отлично от нуля. Эта вторая особенность МДПТ рассмотрена в следующем параграфе.

III.2.4. Пороговое напряжение МДПТ и влияние потенциала подложки

Отметим, что при обратном смещенном p - n переходе исток—подложка ($V_{bs} < 0$) даже при нулевом напряжении сток—исток ($V_{ds} = 0$) МДПТ не является равновесной системой,



так как через переходы исток—подложка, сток—подложка и канал—подложка протекают (хотя и малые) обратные токи. В результате уровень Ферми F расщепляется на квазиуровни Ферми для электронов F_n и дырок F_p , расстояние между которыми определяется напряжением между подложкой и каналом, равным при $V_{ds} = 0$ напряжению V_{bs} :

$$F_n - F_p = eV_{bs}.$$

На рис. III.2.6 представлены энергетические диаграммы n -канального МДПТ в сечении $0 < y < L$ (см. рис. III.2.4) для случая, когда напряжение сток—исток равно нулю, а напряжение затвор—исток равно пороговому напряжению транзистора: $V_{gs} = V_t$. Рисунок III.2.6, *a* соответствует нулевому напряжению подложка—исток ($V_{bs} = 0$), а рис. III.2.6, *б* — закрытому переходу подложка—исток ($V_{bs} < 0$). Диаграмма на рис. III.2.6, *a* в точности соответствует рис. III.1.6 для МДП-структуры. Поверхностный потенциал при пороговом напряжении, как и для МДП-структуры, составляет

$$\phi_s = 2\phi_B,$$

поэтому при $V_{bs} = 0$ пороговая ширина ОПЗ под затвором и пороговое напряжение МДПТ определяются теми же соотношениями (III.1.7), (III.1.19а, б) и (III.1.23), что и для МДП-структуры.

$$l_{t0} = l(V_{t0}) = \sqrt{\frac{2\epsilon\epsilon_0 \cdot 2\phi_B}{eN_B}}, \quad (\text{III.2.4})$$

$$V_{t0} = \phi_{GB} \pm 2\phi_B - (Q_{sef} + Q_{sB0}) / C_{sG} = \quad (\text{III.2.5a})$$

$$= \phi_{GBi} \pm \phi_B - (Q_{sef} + Q_{sB0}) / C_{sG}, \quad (\text{III.2.5б})$$

$$\text{где} \quad Q_{sB0} = \mp eN_B l_{t0} = \mp \sqrt{2\epsilon\epsilon_0 \cdot 2\phi_B \cdot eN_B}. \quad (\text{III.2.6})$$

По определению порогового напряжения при $V_{gs} = V_t$ выполнены равенства $n(0) = N_B$ и $F_i(0) - F_n(0) = e\phi_B$. Рисунок III.2.6, *б* показывает, что при $V_{gs} = V_t$ и $V_{bs} < 0$ поверхностный потенциал возрастает до значения

$$\phi_s = 2\phi_B - V_{bs}.$$

Поэтому в соотношениях (III.2.4)–(III.2.6), определяющих пороговую ширину ОПЗ и пороговое напряжение, следует сделать замену: $2\phi_B \rightarrow 2\phi_B - V_{bs}$. При этом вместо (III.2.4)–(III.2.6) для n -канального МДПТ получаем

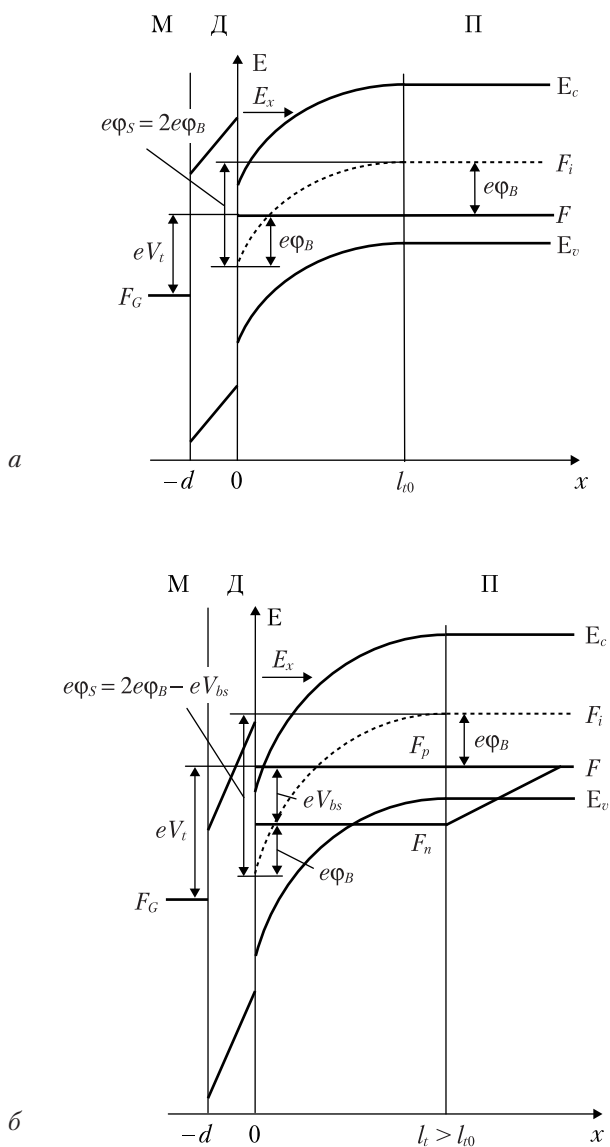


Рис. III.2.6. Энергетические диаграммы *n*-канального МДПТ
при $V_{gs} = V_t$, $V_{ds} = 0$:
a - $V_{bs} = 0$; *б* - $V_{bs} < 0$

$$l_t(V_{bs}) = l_{t0} \sqrt{1 - V_{bs} / 2\phi_B}. \quad (\text{III.2.7})$$

$$V_t(V_{bs}) = \phi_{GB} + 2\phi_B - [Q_{ss\text{ef}} + Q_{sB}(V_{bs})] / C_{sG} = \quad (\text{III.2.8a})$$

$$= \phi_{GBi} + \phi_B - [Q_{ss\text{ef}} + Q_{sB}(V_{bs})] / C_{sG}, \quad (\text{III.2.8б})$$

$$\text{где } Q_{sB}(V_{bs}) = -eN_B l_t = -\sqrt{2\epsilon\epsilon_0 \cdot (2\phi_B - V_{bs})} \cdot eN_B. \quad (\text{III.2.9})$$

Сравнение соотношений (III.2.7)–(III.2.9) и (III.2.4)–(III.2.6) показывает, что

$$V_t(V_{bs}) = V_{t0} + K_B \left(\sqrt{2\phi_B - V_{bs}} - \sqrt{2\phi_B} \right), \quad (\text{III.2.10})$$

$$\text{где } K_B = \sqrt{2\epsilon\epsilon_0 e N_B} / C_{sG} \quad (\text{III.2.11})$$

— коэффициент подложки, зависящий от свойств подложки (N_B) и свойств диэлектрика (C_{sG}). В n -канальных транзисторах при увеличении запирающего напряжения V_{bs} пороговое напряжение возрастает, в p -канальных — снижается. Во всех случаях *запирающее напряжение V_{bs} снижает ток стока при заданных напряжениях V_{gs} и V_{ds}* .

В технической литературе потенциал истока часто принимается за 0 (заземленный исток). При этом полученный результат можно сформулировать как *зависимость порогового напряжения от потенциала подложки* (точнее, от напряжения подложка—исток).

Физическая причина эффекта влияния подложки состоит в том, что подложка влияет на поверхностную концентрацию носителей заряда в канале (и, следовательно, на проводимость канала) так же, только в меньшей степени, как затвор. Действие потенциала подложки поясняется рис. III.2.7. Затвор воздействует на канал через емкость затвор—канал с удельным значением $C_{sG} = \epsilon\epsilon_0 d_f$, а подложка — через емкость подложка—канал, удельное значение которой при $V_{bs} = 0$ составляет

$$C_{sBt} \approx \epsilon\epsilon_0 / l_{t0}.$$

Учитывая, что ширина ОПЗ l_{t0} определяется соотношением (III.2.4), находим, что отношение

$$\frac{C_B}{C_G} = \frac{d_{ef}}{l_{t0}} = \frac{\sqrt{2\epsilon\epsilon_0 e N_B}}{C_{sG} \sqrt{8\phi_B}} = \frac{K_B}{\sqrt{8\phi_B}} \quad (\text{III.2.12})$$

пропорционально коэффициенту подложки. Таким образом, коэффициент подложки характеризует относительную степень влияния подложки на проводимость канала. Это



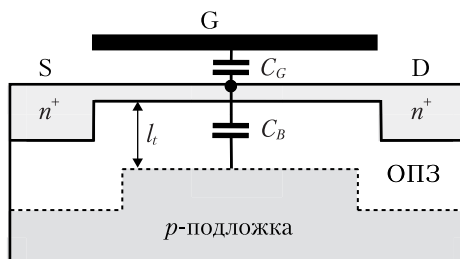


Рис. III.2.7. Влияние потенциала подложки на проводимость канала

влияние обычно является вредным фактором, так как в ИС транзисторы на общей подложке могут иметь различные потенциалы истоков и, следовательно, различные характеристики. Коэффициент подложки уменьшается при уменьшении толщины диэлектрика (возрастает C_G) и при уменьшении концентрации примеси в подложке (увеличивается ширина ОПЗ и снижается C_B).

Отметим, что пороговое напряжение является *параметром* транзистора, т.е. *числом*, которое определяется (и измеряется) при $V_{ds} = 0$. При $V_{ds} > 0$ точки канала, находящиеся на различном расстоянии от истока, имеют различные потенциалы и *локальное пороговое напряжение* внутри прибора зависит от координаты y (см. рис. III.2.4). Это обстоятельство существенно усложняет формулы для ВАХ МДПТ.

Пороговое напряжение является одним из важнейших параметров МДПТ, от воспроизводимости которого зависит возможность реализации различных схемотехнических решений в ИМС, а также величина напряжения питания. Поэтому необходимо располагать способами регулирования порогового напряжения в процессе изготовления ИС. Из соотношений (III.2.4)–(III.2.6) следует, что величина V_{t0} зависит от контактной разности потенциалов металл–собственный полупроводник ϕ_{GBi} , концентрации примеси в подложке N_B , а также от толщины d и диэлектрической проницаемости ϵ_d подзатворного диэлектрика (параметр C_{sG}) и эффективной плотности поверхностных состояний $N_{ss\,ef}$.

Таблица III.1.1 показывает, что подбор металла затвора позволяет варьировать пороговое напряжение в диапазоне около одного вольта, однако в ИМС выбор материала затвора ограничен технологическими соображениями. Использо-

ние поликремниевого затвора позволяет изменять пороговое напряжение на величину, соответствующую ширине запрещенной зоны (1,12 В) в зависимости от типа проводимости. Заметим, что поликремний n^+ -типа является единственным материалом, позволяющим получить отрицательное значение Φ_{GBi} .

Выбор диэлектрика также в значительной степени определяется технологическими возможностями. Наиболее часто применяется собственный термический оксид SiO_2 , имеющий диэлектрическую проницаемость $\epsilon_d \approx 3,9$. Высокой диэлектрической проницаемостью обладает нитрид кремния Si_3N_4 ($\epsilon_d \approx 7,5$). Величина удельной емкости диэлектрика C_{sG} определяет степень проявления эффекта поля, который лежит в основе действия МДП-транзистора. Поэтому регулировка порогового напряжения за счет изменения C_{sG} вряд ли целесообразна. Возможности регулировки порогового напряжения за счет вариации концентрации примеси в подложке N_B также ограничены. При $N_B < 10^{15} \text{ см}^{-3}$ свойства полупроводника чувствительны к влиянию температуры, а также существенно возрастают обратные токи через p - n переходы. При $N_B > 10^{17} \text{ см}^{-3}$ значительно увеличивается коэффициент подложки. Вариация концентрации примеси N_B в указанном диапазоне меняет значение V_{t0} на $\approx 0,06 \text{ В}$ за счет изменения величины Φ_B на 0,36—0,6 В.

Значительно больший эффект дает поверхностное легирование подзатворной области подложки (см. параграф III.1.6). Доза легирования тонкого поверхностного слоя $\Delta D_B = 10^{12} \text{ см}^{-2}$ обеспечивает изменение порогового напряжения на величину 2—3 В.

В заключение заметим, что высокая воспроизводимость порогового напряжения требует возможно большего снижения плотности поверхностных состояний N_{ss} . Достаточно сказать, что идея создания МДПТ, высказанная Шокли в 1948 г. [2], долгое время не могла быть реализована только из-за отсутствия надежных способов подготовки поверхности полупроводника перед нанесением диэлектрика. При достижимом в то время значении $N_{ss} \approx 5 \cdot 10^{12} \text{ см}^{-2}$ и толщине диэлектрика $d \approx 1 \text{ мкм}$ вклад слагаемого $Q_{ss} / C_{sG} = eN_{ss} / C_{sG}$ в значение порогового напряжения составляет $\approx 100 \text{ В}$. Учитывая низкую воспроизводимость величины N_{ss} , легко понять, что производство МДПТ в таких условиях было невозможным.



Основные выводы

1. Принцип действия МДП-транзистора основан на модуляции электрического сопротивления инверсионного приповерхностного слоя полупроводника (канала) напряжением затвор—исток.

2. На выходных ВАХ МДПТ можно выделить область отсечки, крутую область и пологую область. В пологой области ВАХ ток стока почти не зависит от напряжения сток—исток, что объясняется перекрытием канала на границе со стоком.

3. Пороговое напряжение МДПТ зависит от напряжения подложка—исток и определяется соотношением (III.2.10). Степень влияния потенциала подложки характеризуется коэффициентом подложки K_B , который определяется соотношением (III.2.11).

4. Эффективными способами регулировки порогового напряжения являются применение поликремниевого затвора n^+ - или p^+ -типов, а также поверхностное легирование подзатворной области подложки.

III.3. Анализ идеализированного транзистора

III.3.1. Допущения модели идеализированного МДПТ

Структура n -канального МДПТ схематично показана на рис. III.3.1, *a*.

Наиболее простая идеализированная модель транзистора предложена Шокли и Пирсоном [3]. Модель основана на следующих допущениях.

1. Обратные токи утечки p - n переходов и перехода канал—подложка пренебрежимо малы, поэтому ток стока и ток истока одинаковы и равны току в канале I , который не зависит от координаты y (см. рис. III.3.1, *a*):

$$I_D = I_S = I. \quad (\text{III.3.1})$$

2. Толщина канала много меньше эффективной толщины диэлектрика

$$h_c \ll d_{ef} = \epsilon_d d / \epsilon. \quad (\text{III.3.2})$$

3. Ток в канале переносится только носителями заряда, которые в подложке являются неосновными (для n -канального МДПТ — электронами). Роль основных носителей пренебрежимо мала (униполярный прибор).

4. Ток в канале в направлении оси y (см. рис. III.2.4) имеет чисто дрейфовый механизм:

$$I = I_{dr} \gg I_{dif}. \quad (\text{III.3.3})$$



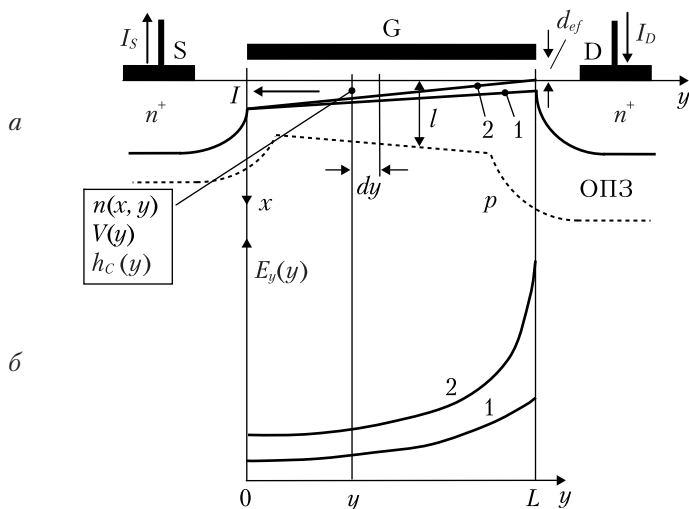


Рис. III.3.1. Схематичная структура n -канального МДПТ (а) и распределения напряженности продольного электрического поля E_y в канале (б):

1 — крутая область ВАХ; 2 — пологая область ВАХ

5. Подвижность носителей заряда в канале μ постоянна (не зависит от электрического поля).

6. Ширина ОПЗ l под затвором не зависит от координаты.

7. При напряжении затвор—исток ниже порогового напряжения проводимость канала равна нулю и ток в канале отсутствует.

8. Продольное электрическое поле много меньше поперечного (приближение плавного канала)

$$E_y \ll E_x. \quad (\text{III.3.4})$$

9. Длина канала L и ширина канала Z (в направлении, перпендикулярном плоскости, см. рис. III.3.1) достаточно велики.

10. Потенциал канала $V(y)$ не зависит от координаты x .

Отметим, что допущения 5 и 8 предполагают достаточно большую длину канала и не слишком большие напряжения сток—исток. В этом случае отсутствует разогрев носителей заряда в канале продольным полем, активизирующий новые механизмы их рассеяния (например, на оптических фононах) и снижающий подвижность.

Допущение 6 исключает зависимость локального порогового напряжения от координаты y (см. параграф III.2.4). Таким образом, локальное значение порогового напряжения $V_t(y)$ совпадает с пороговым напряжением транзистора и зависит только от напряжения подложка—исток:

$$V_t(V_{bs}, y) = V_t(V_{bs}). \quad (\text{III.3.5})$$

Допущение 10 позволяет приписать границам канала с n^+ -областями ($y = 0$ и $y = L$ на рис. III.2.4) потенциалы «внутренних» точек стока и истока:

$$V(0) = V_s; \quad (\text{III.3.6a})$$

$$V(L) = V_d. \quad (\text{III.3.6b})$$

Наконец, допущение 9 исключает необходимость учета краевых эффектов и позволяет (вместе с допущениями 8 и 10) свести задачу анализа статических характеристик МДПТ к квазиодномерной.

Принятые допущения позволяют получить простые аналитические формулы для описания ВАХ МДПТ, а также выяснить его инерционные свойства. В современных конструкциях МДПТ допущения 2, 5, 6—10 могут нарушаться. Допущения 1 и 4 также не выполняются в режиме малых токов. Необходимые поправки будут сделаны далее в более точных моделях (см. разделы III.4, III.5).

III.3.2. Характеристики идеализированного транзистора

Для определенности вывод ВАХ проведем для n -канального МДПТ. В соответствии с допущениями 4 и 1 в параграфе III.3.1 ток в канале I имеет дрейфовую природу и равен току стока I_D . При допущении 5 напряжение dV на участке канала длиной dy (см. рис. III.3.1) определяется законом Ома и составляет

$$dV(y) = I_D dR_c(y), \quad (\text{III.3.7})$$

где $dR_c(y) = dy / Z\sigma_{cs}(y)$, $\sigma_{cs}(y) = \int_0^{h_c(y)} \mu_n en(x, y) dx$ — поверх-

ностная проводимость канала. Учитывая, что подвижность электронов в канале постоянна (допущение 5), а интеграл



$\int_0^{h_c(y)} en(x, y) dx$ представляет собой поверхностную плотность заряда электронов в канале $-Q_{sn}$ (см. рис. III.1.6), получаем

$$dR_C(y) = -dy / \mu_n Z Q_{sn}(y). \quad (\text{III.3.8})$$

В режиме *отсечки* ($V_{gs} < V_t$) канал отсутствует и $Q_{sn} = 0$.

В *крутой* области ВАХ ($V_{gs} \geq V_t$, $V_{gd} \geq V_t$) заряд электронов в канале Q_n полностью экранирует полупроводниковую область под затвором от изменений электрического поля в диэлектрике. Поэтому поверхностная плотность заряда $Q_{sn}(y)$ линейно зависит от напряжения затвор—канал $V_g - V(y)$:

$$-Q_{sn}(y) = C_{sg}[V_g - V(y) - V_t]. \quad (\text{III.3.9})$$

Последнее равенство идентично формуле для плоского конденсатора с той разницей, что обеспечивает локальное нулевое значение Q_{sn} при $V_g - V(y) = V_t$.

На границе с *пологой* областью ВАХ ($V_{gs} \geq V_t$, $V_{gd} = V_t$) канал на границе со стоком ($y = L$) перекрывается: $Q_{sn}(L) = 0$, а потенциал канала согласно (III.3.6б) достигает при $y = L$ максимального значения $V(L) = V_d$. В полой области ВАХ ($V_{gs} \geq V_t$, $V_{gd} < V_t$) соотношение (III.3.9) несправедливо, так как дает для прилежащей к стоку области канала ($y \rightarrow L$) отрицательные значения заряда $-Q_{sn}(y)$, что соответствует отрицательной концентрации электронов в канале и физически невозможно.

Таким образом, с учетом (III.3.6а, б) получаем

$$-Q_{sn}(y) = \begin{cases} 0, & V_{gs} < V_t, \\ C_{sg}[V_{gs} - V_t - V(y) + V_s], & V_{gs, gd} \geq V_t. \end{cases} \quad (\text{III.3.10б})$$

В полой области ВАХ ($V_{gs} \geq V_t$, $V_{gd} < V_t$) функция $Q_{sn}(y)$ остается такой же, как при $V_{gd} = V_t$, и не изменяется при изменении напряжения $V_{gd} < V_t$.

Отметим, что уравнение (III.3.10б) *локально* связывает поверхностную плотность заряда электронов в канале Q_{sn} с потенциалом канала $V(y)$. Такая локальная связь обусловлена приближением плавного канала 8. Если оно не выполняется, то в точке с координатой y значение $Q_{sn}(y)$ зависит не только от потенциала $V(y)$, но и от потенциалов канала в некоторой окрестности точки y . С такой ситуацией мы столкнемся в более сложной модели МДПТ.

Подстановка (III.3.10б) в (III.3.8), а затем в (III.3.7) дает дифференциальное уравнение, связывающее при $V_{gs} \geq V_t$ потенциал канала $V(y)$ с током стока

$$I_D dy / L = \beta [V_{gs} - V_t - V(y) + V_s] dV, \quad (\text{III.3.11})$$

где

$$\beta = Z \mu_n C_{sg} / L. \quad (\text{III.3.12})$$

Интегрирование левой части уравнения (III.3.11) по y от 0 до y , а правой — по $V(y)$ от $V(0) = V_s$ до $V(y)$ дает распределение потенциала $V(y)$ в канале

$$I_D y / L = \beta [V(y) - V_s] \cdot [V_{gst} - [V(y) - V_s] / 2]; \quad (\text{III.3.13})$$

$$V(y) - V_s = V_{gst} - \sqrt{V_{gst}^2 - 2y I_D / \beta L} \quad (\text{III.3.14})$$

и абсолютной величины напряженности продольного электрического поля в канале

$$E_y(y) = \frac{dV(y)}{dy} = \frac{I_D / \beta L}{\sqrt{V_{gst}^2 - 2I_D y / \beta L}}. \quad (\text{III.3.15})$$

В уравнениях (III.3.13)–(III.3.15) использовано обозначение

$$V_{gst} = V_{gs} - V_t,$$

которое далее будет принято повсеместно, поскольку во всех формулах, относящихся к характеристикам МДПТ, фигурирует не абсолютная величина напряжения затвор–исток, а ее превышение над пороговым значением. В некоторых случаях удобно также использовать обозначение

$$V_{gdt} = V_{gd} - V_t.$$

Уравнения (III.3.11), (III.3.13)–(III.3.15) справедливы при работе МДПТ в крутой области ВАХ ($V_{gst} \geq 0$, $V_{gdt} \geq 0$). ВАХ транзистора в *крутой* области может быть получена из уравнения (III.3.13) при $y = L$, $V(y) = V(L) = V_d$:

$$I_D = \beta V_{ds} (V_{gst} - V_{ds} / 2) = \quad (\text{III.3.16a})$$

$$= \beta (V_{gst}^2 - V_{gdt}^2) / 2. \quad (\text{III.3.16б})$$

В режиме *отсечки* $I_D = 0$ (допущение 7) в *пологой* области ВАХ функции $Q_{sn}(y)$, $V(y)$ и, следовательно, ток I_D остаются такими же, как при $V_{gdt} = 0$ и не изменяются при изменении напряжения $V_{gd} < V_t$. Таким образом, ВАХ МДПТ



для произвольного режима работы можно представить в виде

$$I_D = \begin{cases} 0, & V_{gst} < 0, \\ \beta V_{ds} (V_{gst} - V_{ds} / 2), & V_{gst} \geq 0, V_{ds} \leq V_{dsS} = V_{gst}, \\ I_{DS} = \beta V_{gst}^2 / 2, & V_{gst} \geq 0, V_{ds} > V_{dsS}, \end{cases} \quad (\text{III.3.17a})$$

$$\text{или } I_D = \begin{cases} 0, & V_{gs} < V_t; \\ \beta (V_{gst}^2 - V_{gdt}^2) / 2, & V_{gst} \geq 0, V_{gdt} \geq 0, \\ I_{DS} = \beta V_{gst}^2 / 2, & V_{gst} \geq 0, V_{gdt} < 0. \end{cases} \quad (\text{III.3.17б})$$

Ток I_{DS} в пологой области ВАХ не зависит от напряжений V_{ds} , V_{gd} и называется *током насыщения*. Напомним, что в (III.3.17a) принято $V_{ds} \geq 0$, а в (III.3.17б) — $V_{gd} < V_{gs}$. Если эти эквивалентные неравенства не выполнены, электроды стока и истока следует поменять местами.

Соотношения (III.3.17a) описывают ВАХ МДПТ при включении по схеме с *общим истоком* (ОИ), когда входным напряжением V_{in} является напряжение затвор–исток V_{gs} , входным током I_{in} — ток затвора I_G (равный нулю), выходным напряжением V_{out} — напряжение сток–исток V_{ds} и выходным током I_{out} — ток стока I_D (рис. III.3.2, а).

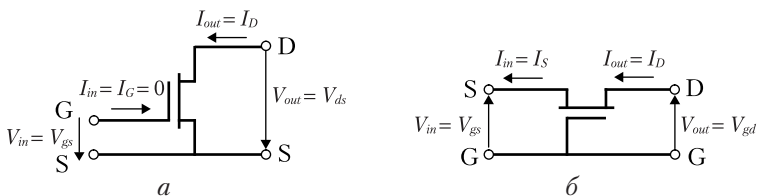


Рис. III.3.2. Включение МДПТ по схеме с общим истоком (ОИ) и с общим затвором (ОЗ)

Соотношения (III.3.17б) описывают ВАХ МДПТ при включении по схеме с *общим затвором* (ОЗ), когда входным напряжением V_{in} является напряжение затвор–исток V_{gs} , входным током I_{in} — ток истока I_S , выходным напряжением V_{out} — напряжение затвор–сток V_{gd} и выходным током I_{out} — ток стока I_D (рис. III.3.2, б).

Соотношения (III.3.17a, б) показывают, что ВАХ идеализированного МДП-транзистора определяются *двумя параметрами*: пороговым напряжением V_t и параметром β ,

определенным в (III.3.12). Аргументами ВАХ являются напряжения V_{gs} и V_{ds} или напряжения V_{gs} и $V_{gd} = V_{gs} - V_{ds}$. Напряжение подложка—исток V_{bs} не входит в число аргументов ввиду допущения 6 параграфа III.3.1 и, в соответствии с (III.2.10), влияет только на величину порогового напряжения V_t .

На рис. III.3.3, а и III.3.3, б представлены выходные и проходные ВАХ идеализированного n -канального МДПТ-транзистора с пороговым напряжением $V_t = 1$ В, включенного по схеме ОИ. Характер соотношений (III.3.17а) соответствует качественным выводам параграфа III.2.3. При $V_{ds} \ll V_{gst}$ выходные ВАХ линейны:

$$I_D \rightarrow G_0(V_{gs})V_{ds}, \quad (III.3.18)$$

где
$$G_0 = \lim_{V_{ds} \rightarrow 0} \frac{\partial I_D}{\partial V_{ds}} = \beta V_{gst} \quad (III.3.19)$$

— начальная выходная проводимость МДПТ. В крутой области ВАХ выходная проводимость может быть получена из (III.3.17 а):

$$G = \frac{\partial I_D}{\partial V_{ds}} = \beta(V_{gst} - V_{ds}). \quad (III.3.20a)$$

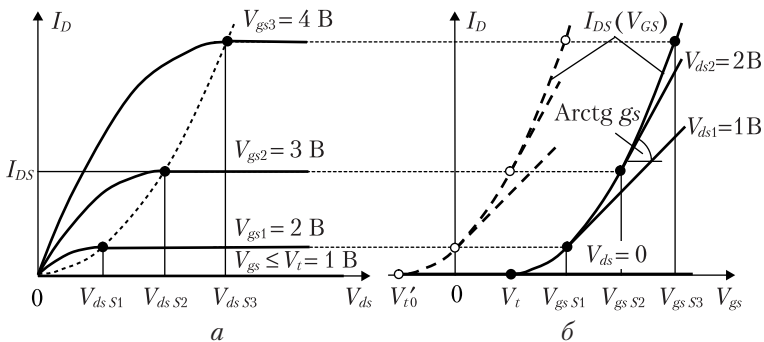


Рис. III.3.3. ВАХ идеализированного МДПТ с пороговым напряжением $V_t = 1$ В в схеме включения ОИ:

а — выходные; б — проходные. Пунктир на рис. а разделяет крутую и пологую области ВАХ. Штриховые линии на рис. б соответствуют пороговому напряжению $V'_t = -1$ В

При переходе в пологую область ВАХ выходная проводимость

$$G \rightarrow G_s = \frac{\partial I_{DS}}{\partial V_{ds}} = 0. \quad (III.3.20б)$$

На границе крутой и пологой областей ВАХ $I_D = I_{DS}$, $V_{ds} = V_{gst}$, и из (III.3.17а) получаем

$$I_D = \beta V_{ds}^2 / 2. \quad (\text{III.3.21})$$

Эта формула описывает границу раздела крутой и пологой областей ВАХ, показанную на рис. III.3.3, а пунктиром.

На рис. III.3.3, б изображены проходные ВАХ МДПТ-транзистора в схеме ОИ — зависимости $I_D(V_{gs})$ с параметром V_{ds} . При $V_{gs} < V_{ds} + V_{gst}$ транзистор работает в пологой области ВАХ, ток стока I_D равен току насыщения I_{DS} , который не зависит от напряжения и согласно (III.3.17а) составляет

$$I_{DS} = \beta V_{gst}^2 / 2. \quad (\text{III.3.22})$$

Граница крутой и пологой областей ВАХ соответствует напряжению

$$V_{gs} = V_{gsS} = V_{ds} + V_t. \quad (\text{III.3.23})$$

При этом напряжении проходные ВАХ «отслаиваются» по касательной от параболы (III.3.22), переходя согласно (III.3.17а) в прямые линии

$$I_D = \beta V_{ds} (V_{gs} - V_{ds} / 2). \quad (\text{III.3.24})$$

Усилительные свойства МДП-транзистора характеризуются *крутизной ВАХ*

$$g = \frac{\partial I_D}{\partial V_{gs}}, \quad (\text{III.3.25})$$

соответствующей тангенсу угла наклона проходных ВАХ (см. рис. III.3.3, б). Из (III.3.24) и (III.3.17а) следует, что в крутой области ВАХ

$$g = \beta V_{ds}. \quad (\text{III.3.26a})$$

Крутизна возрастает при повышении напряжения V_{ds} и достигает максимума при переходе в пологую область ВАХ, где ее значение g_S может быть получено из (III.3.22):

$$g_S = \frac{\partial I_{DS}}{\partial V_{gst}} = \beta V_{gst}. \quad (\text{III.3.26б})$$

Одним из важнейших параметров, определяющих качество МДПТ, является *удельная* (на единицу ширины канала Z) *крутизна* ВАХ $g_{sz} = g_S / Z$. С учетом (III.3.12) из (III.3.22) получаем

$$g_{S_z} = V_{gst} \mu_n C_{sG} / L. \quad (\text{III.3.27})$$

Удельная крутизна не зависит от ширины канала и определяется только напряжением затвор—исток и электрофизическими параметрами МДПТ: подвижностью электронов в канале μ_n , удельной емкостью затвора $C_{sG} = \epsilon_d \epsilon_0 / d$ и длиной канала L .

В объеме подложки подвижность носителей заряда повышается при снижении концентрации примеси (см. раздел I.4) и достигает максимального значения, соответствующего рассеянию на фононах, при концентрации примеси $N_B \approx 10^{15} \text{ см}^{-3}$ (как для электронов, так и для дырок). Канал МДПТ формируется на границе с поверхностью раздела Si—SiO₂. Вследствие рассеяния на поверхностных дефектах подвижность носителей в канале в зависимости от качества поверхности на 10—20% ниже, чем в объеме, и остается близкой к максимальному значению при $N_B < 10^{16} \text{ см}^{-3}$.

Повышение удельной емкости затвора (за счет уменьшения толщины диэлектрика и применения диэлектриков с высоким значением ϵ_d) является одной из тенденций современной МДП-микроэлектроники. Заметим, однако, что при повышении удельной емкости диэлектрика пропорционально повышается емкость затвор—канал, что ведет к снижению быстродействия.

Радикальный способ повышения удельной крутизны состоит в уменьшении длины канала L .

Еще одной из важнейших характеристик МДПТ является *собственный коэффициент усиления по напряжению*

$$K_V = \left. \frac{\partial V_{ds}}{\partial V_{gs}} \right|_{dI_D=0} = \frac{g}{G}. \quad (\text{III.3.28})$$

Его величина определяет максимальный коэффициент усиления по напряжению, который может быть получен в усилительном каскаде на МДПТ. Согласно (III.3.20а) и (III.3.26а) в крутой области ВАХ

$$K_V = \left. \frac{\partial V_{ds}}{\partial V_{gs}} \right|_{dI_D=0} = \frac{V_{ds}}{V_{gst} - V_{ds}}. \quad (\text{III.3.29a})$$

При переходе в пологую область ВАХ коэффициент усиления

$$K_V \rightarrow K_{VS} = \infty. \quad (\text{III.3.29б})$$



На рис. III.3.3, б показаны также проходные ВАХ МДПТ со встроенным каналом ($V_t' = -1$ В). При том же значении параметра β проходные ВАХ сдвигаются по оси напряжения в сторону меньших напряжений на $V_t - V_t'$, занимая частично второй квадрант.

Выходные и проходные ВАХ МДПТ, включенного по схеме ОЗ, соответствующие уравнениям (III.3.17б), показаны на рис. III.3.4. Чтобы сохранить вид выходных ВАХ, по оси напряжений отложено значение $-V_{gd}$.

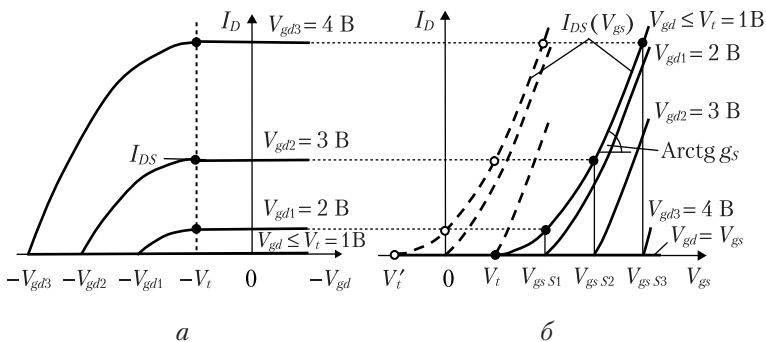


Рис. III.3.4. ВАХ идеализированного МДПТ с пороговым напряжением $V_t = 1$ В в схеме включения ОЗ:
 а — выходные; б — проходные. Пунктир на рис. а разделяет крутую и пологую области ВАХ. Штриховые линии на рис. б соответствуют пороговому напряжению $V_t' = -1$ В

По сравнению с рис. III.3.3, а выходные ВАХ (рис. III.3.4, а) сдвинуты по оси напряжения в сторону меньших напряжений на величину $V_{ds} + V_{gd} = V_{gs}$, занимая частично второй квадрант. Согласно (III.3.23) граница крутой и пологой областей ВАХ соответствует напряжению $-V_{gd} = -V_t$. Эта вертикальная прямая показана на рис. III.3.4, а пунктиром.

Сравнение соотношений (III.3.17б) и (III.3.17а) показывает, что в пологой области ($V_{gd} \leq V_t$) проходные ВАХ (рис. III.3.4, а) в схемах ОЗ и ОИ совпадают. Поэтому соотношение (III.3.26б) для крутизны ВАХ в пологой области остается справедливым и для включения ОЗ. В крутой области характеристики сдвигаются в сторону меньших токов на величину $\beta V_{gd}^2 / 2$ (см. рис. III.3.4, б).

При включении по схеме ОИ входной ток транзистора равен нулю и коэффициент передачи тока

$$K_I = \partial I_D / \partial I_G \rightarrow \infty.$$

При включении ОЗ (см. рис. III.3.2, б) входной ток транзистора $I_{in} = I_S = I_D$ и коэффициент передачи тока

$$K_I = \partial I_D / \partial I_S = 1.$$

Соотношение (III.3.17б) позволяет определить входное сопротивление МДПТ при включении по схеме ОЗ:

$$R_{in} = \partial V_{gs} / \partial I_S = \partial V_{gs} / \partial I_D.$$

В пологой области ВАХ

$$R_{inS} = 1 / g_S = 1 / \beta V_{gst}. \quad (\text{III.3.30})$$

Таким образом, при включении по схеме ОЗ транзистор обладает свойствами *повторителя тока* с низким входным и высоким выходным сопротивлением ($R_{outS} \rightarrow \infty$).

III.3.3. Инерционные свойства идеализированного МДПТ

Инерционность идеализированного МДПТ обусловлена изменением заряда подвижных носителей в канале при изменении напряжений V_{gs} и V_{gd} . В крутой области ВАХ для n -канального МДПТ распределение заряда электронов в канале $Q_{sn}(y)$ и компенсирующего его заряда на затворе $Q_{sG}(y) = -Q_{sn}(y)$ определяется соотношением (III.3.9), распределение потенциала — соотношением (III.3.14), а ВАХ — соотношением (III.3.16б):

$$Q_{sG}(y) = -Q_{sn}(y) = C_{sG}[V_g - V(y) - V_t]; \quad (\text{III.3.31})$$

$$V(y) - V_s = V_{gst} - \sqrt{V_{gst}^2 - 2yI_D / \beta L}; \quad (\text{III.3.32})$$

$$I_D = \beta(V_{gst}^2 - V_{gdt}^2) / 2. \quad (\text{III.3.33})$$

Суммарный заряд на затворе составляет

$$Q_G = -Q_n = -Z \int_0^L Q_{sn}(y) dy. \quad (\text{III.3.34})$$

Подстановка (III.3.32) в (III.3.31), а затем в (III.3.34) и интегрирование результата с учетом (III.3.33) дает



$$Q_G = C_G V_{gst} \cdot \frac{1 - V_{ds} / V_{gst} + (V_{ds} / V_{gst})^2 / 3}{1 - V_{ds} / 2V_{gst}},$$

$$\text{где} \quad C_G = C_{sG} ZL \quad (\text{III.3.35})$$

— общая емкость диэлектрика между затвором и каналом,
 $V_{ds} = V_{gs} - V_{gd} = V_{gst} - V_{gdt}$.

В режиме отсечки ($V_{gst} < 0$) $Q_G = 0$. В пологой области ВАХ ($V_{ds} / V_{gst} > 1$) канал перекрыт на границе со стоком (см. рис. III.3.1) и напряжение V_{ds} не влияет на распределение и величину заряда Q_G , который сохраняет свое значение при $V_{ds} = V_{gst}$:

$$Q_G = \frac{2C_G}{3} V_{gst}.$$

Таким образом,

$$Q_G = C_G V_{gst} \times \begin{cases} 0, & V_{gst} < 0, \\ \frac{1 - V_{ds} / V_{gst} + (V_{ds} / V_{gst})^2 / 3}{1 - V_{ds} / 2V_{gst}}, & V_{gst} \geq 0, \quad V_{ds} / V_{gst} \leq 1, \\ 2/3, & V_{gst} \geq 0, \quad V_{ds} / V_{gst} > 1. \end{cases} \quad (\text{III.3.36})$$

Отсюда следует, что в пологой области ВАХ инерционные свойства идеализированного МДПТ определяются двумя емкостями: затвор—исток C_{gs} и затвор—сток C_{gd} , причем

$$C_{gs} = \frac{\partial Q_G}{\partial V_{gs}} = \frac{2C_G}{3}; \quad (\text{III.3.37a})$$

$$C_{gd} = \frac{\partial Q_G}{\partial V_{gd}} = 0. \quad (\text{III.3.37b})$$

Частотные свойства МДПТ принято характеризовать *предельной частотой* $f_T = \omega_T / 2\pi$. Значение ω_T определяется как отношение крутизны ВАХ в пологой области g_s к суммарной емкости $C_{gs} + C_{gd}$:

$$\omega_T = \frac{g_s}{C_{gs} + C_{gd}} \bigg|_{V_{ds} \geq V_{ds s}}. \quad (\text{III.3.38})$$

Предельная частота соответствует частоте, на которой в цепочке идентичных оптимально сконструированных усили-



тельных каскадов на МДПТ коэффициент усиления каждого каскада по абсолютной величине равен 1.

Используя соотношения (III.3.38), (III.3.37) и (III.3.266), получаем

$$\omega_T = \frac{3}{2} \cdot \frac{\beta V_{gst}}{C_G} = \frac{3}{2} \cdot \frac{\mu V_{gst}}{L^2}, \quad (\text{III.3.39})$$

где μ — подвижность носителей заряда в канале (электронов для n -канального, дырок для p -канального МДПТ).

Формула (III.3.39) показывает, что предельная частота идеализированного МДПТ зависит только от режима работы (V_{gst}), подвижности носителей заряда в канале и длины канала L . Влияние длины канала особенно существенно ($\omega_T \sim 1/L^2$), так как при ее уменьшении повышается крутизна ВАХ и уменьшается емкость диэлектрика.

На границе пологой области $V_{gst} = V_{ds}$. При этом соотношение (III.3.39) можно представить в виде

$$\omega_T = \frac{3}{2} \cdot \frac{\mu V_{gst}}{L^2} = \frac{3}{2} \cdot \frac{\mu V_{ds}}{L^2} = \frac{\mu \bar{E}_y}{L} = \frac{\bar{v}}{L} = \frac{1}{T_C}, \quad (\text{III.3.40})$$

$$\text{где} \quad \bar{E}_y = 3V_{ds} / 2L \quad (\text{III.3.41})$$

— средняя напряженность продольного электрического поля в канале;

$$\bar{v} = \bar{E}_y \mu \quad (\text{III.3.42})$$

— средняя скорость дрейфа носителей заряда от истока к стоку;

$$T_C = L / \bar{v} \quad (\text{III.3.43})$$

— время пролета носителей заряда через канал. Квадратичную зависимость предельной частоты от $1/L$ можно объяснить тем, что при уменьшении длины канала повышается продольное электрическое поле в канале и, следовательно, скорость дрейфа носителей, а также уменьшается длина их пути от истока к стоку.

В реальных МДПТ зависимость времени пролета носителей через канал более слабая, чем $1/L^2$, а крутизны ВАХ — более слабая, чем $1/L$. Тем не менее, длина канала является одним из главных показателей качества технологического процесса изготовления ИС на МДПТ.



Основные выводы

1. Выходные характеристики МДП-транзистора имеют крутую и пологую области. В пологой области ток стока идеализированного МДПТ не зависит от напряжения сток—исток.

2. В идеализированной модели локальное значение порогового напряжения $V_t(y)$ совпадает с пороговым напряжением транзистора и зависит только от напряжения подложка—исток: $V_t(V_{bs}, y) = V_t(V_{bs})$. Эффективными способами регулирования порогового напряжения являются применение поликремниевого затвора n^+ - или p^+ -типа, а также подлегирование поверхности канала.

3. Идеализированная модель позволяет получить ВАХ транзистора в крутой области в простой аналитической форме (III.3.17). Вид ВАХ не зависит от потенциала подложки и полностью определяется двумя параметрами: пороговым напряжением и параметром β , определенным в (III.3.12). Граница крутой и пологой области ВАХ определяется соотношениями $V_{ds} = V_{gst}$ или $V_{gdt} = 0$.

4. Электрические свойства МДП-транзистора характеризуются крутизной ВАХ $g = \partial I_D / \partial V_{gs}$, выходной проводимостью $G = \partial I_D / \partial V_{ds}$ и собственным коэффициентом усиления по напряжению $K_v = g / G$. В пологой области ВАХ крутизна максимальна, выходная проводимость в рамках модели равна нулю, а собственный коэффициент усиления по напряжению бесконечен.

5. Крутизна ВАХ идеализированного МДПТ возрастает с уменьшением длины канала и увеличением удельной емкости подзатворного диэлектрика.

6. В пологой области ВАХ емкость затвор-исток идеализированного МДПТ составляет $2/3$ емкости затвора, а емкость затвор—сток равна нулю.

7. Предельная частота идеализированного МДПТ обратно пропорциональна квадрату длины канала.

III.4. Статические ВАХ МДП-транзисторов с длинным каналом

III.4.1. Влияние неоднородности ОПЗ под затвором

В модели идеализированного МДП-транзистора полагалось, что ширина ОПЗ l под затвором не зависит от координаты y , локальное значение порогового напряжения $V_t(y)$ совпадает с пороговым напряжением транзистора и зависит только от напряжения подложка—исток: $V_t(V_{bs}, y) = V_t(V_{bs})$. Фактически



же точки канала, находящиеся на различном расстоянии от истока, имеют различные потенциалы, и ширина ОПЗ l , а также локальное пороговое напряжение внутри прибора зависят от координаты y (см. рис. III.2.4).

При нулевом напряжении сток—исток V_{ds} потенциал канала $V(y)$ по всей длине одинаков и равен потенциалу истока: $V(y) = V_s$. В параграфе III.2.4 показано, что при этом пороговое напряжение транзистора определяется соотношением (III.2.10)

$$V_t = V_{t0} + K_B (\sqrt{2\phi_B - V_{bs}} - \sqrt{2\phi_B}) \quad (\text{III.4.1})$$

где $K_B = \sqrt{2\epsilon\epsilon_0 e N_B} / C_{sG}$ — коэффициент подложки. При $V_{ds} \neq 0$ точки канала, находящиеся на различном расстоянии y от истока (см. рис. III.2.4, III.3.1), имеют различные потенциалы и вместо напряжения подложка—исток V_{bs} в (III.4.1) следует использовать напряжение подложка—канал, заменяя

$$V_{bs} \rightarrow V_b - V(y) = V_{bs} + V_s - V(y).$$

Таким образом, внутри прибора пороговое напряжение имеет локальное значение, зависящее от напряжения V_{bs} и координаты y

$$V_t(y) = V_{t0} + K_B [\sqrt{2\phi_B - V_{bs} + V(y) - V_s} - \sqrt{2\phi_B}]. \quad (\text{III.4.2})$$

Физическая причина зависимости $V_t(y)$ состоит в том, что ширина ОПЗ под затвором l , которая определяется напряжением на ОПЗ $V_b - V(y)$, неоднородна по длине канала: $l = l(y)$ (см. рис. III.2.4, III.3.1). Вследствие этого от координаты y зависит поверхностная плотность заряда ионов в ОПЗ $Q_{sB}(y) = -eN_B l(y)$ и зависящее от нее пороговое напряжение. Соответственно изменяется дифференциальное уравнение (III.3.11)

$$I_D dy / L = \beta [V_{gs} - V_{t0} - V(y) + V_s] dV,$$

которое связывает потенциал канала $V(y)$ с током стока и используется при выводе ВАХ идеализированного КМДП в крутой области. В этом уравнении необходимо сделать замену $V_{t0} \rightarrow V_t(V_{bs}, y)$

$$I_D dy / L = \beta [V_{gs} - V_t(V_{bs}, y) - V(y) + V_s] dV. \quad (\text{III.4.3})$$

Подстановка (III.4.2) в (III.4.3) с учетом (III.4.1) и интегрирование левой части полученного уравнения по всей



длине канала от $y=0$ до $y=L$, а правой части — по $V(y)$ от $V(0)=V_s$ до $V(L)=V_d$ дает уточненную ВАХ МДПТ в крутой области, которая учитывает влияние потенциала подложки и неоднородность ширины ОПЗ под затвором:

$$I_D = \beta V_{ds} (V_{gst} - V_{ds} / 2) - \beta K_B \left[\frac{2}{3} [(2\phi_B - V_{bs} + V_{ds})^{3/2} - (2\phi_B - V_{bs})^{3/2}] - V_{ds} (2\phi_B - V_{bs})^{1/2} \right], \quad (\text{III.4.4})$$

где $V_{gst} = V_{gs} - V_t(V_{bs})$; $V_t(V_{bs})$ — пороговое напряжение транзистора, определенное в (III.4.1) (а не локальное значение $V_t(y)$).

Напряжение насыщения V_{dsS} , разграничивающее крутую и пологую области ВАХ, можно найти с помощью соотношения (III.4.2), определяющего функцию $V_t(y)$, при $y=L$, $V(L)=V_d$:

$$V_t(L) = V_{t0} + K_B \left[\sqrt{2\phi_B - V_{bs} + V_{ds}} - \sqrt{2\phi_B} \right]. \quad (\text{III.4.5})$$

Вычитая (III.4.1) из (III.4.5), получаем:

$$V_t(L) - V_t = K_B \left[\sqrt{2\phi_B - V_{bs} + V_{ds}} - \sqrt{2\phi_B - V_{bs}} \right]. \quad (\text{III.4.6})$$

При $V_{ds} = V_{dsS}$ должно выполняться равенство $V_t(L) = V_{gd} = V_{gs} - V_{dsS}$, которое означает, что канал перекрыт на границе со стоком ($y=L$). Подставляя это значение $V_t(L)$ в (III.4.6) при $V_{ds} = V_{dsS}$ и решая полученное уравнение относительно V_{dsS} , находим

$$V_{dsS} = V_{gst} - \xi K_B^2 \left[\sqrt{1 + V_{gst} / \xi^2 K_B^2} - 1 \right], \quad (\text{III.4.7})$$

где $\xi = 1/2 + \sqrt{2\phi_B - V_{bs}} / K_B$.

Нетрудно показать, что при $V_{ds} = V_{dsS}$ (граница с пологой областью ВАХ) уравнение (III.4.4) обеспечивает нулевое значение выходной проводимости

$$G(V_{dsS}) = \left. \frac{\partial I_D}{\partial V_{ds}} (V_{ds}) \right|_{V_{ds}=V_{dsS}} = 0. \quad (\text{III.4.8})$$

Соотношение (III.4.8) также может быть использовано для определения V_{dsS} .

Подстановка (III.4.7) в (III.4.4) дает значение тока насыщения



$$I_{DS} = \beta V_{ds}^2 / 2 - \beta K_B \left[\frac{2}{3} [(2\varphi_B - V_{bs} + V_{ds})^{3/2} - (2\varphi_B - V_{bs})^{3/2}] - V_{ds} (2\varphi_B - V_{bs})^{1/2} \right]. \quad (\text{III.4.9})$$

В пологой области ВАХ ток стока остается равным I_{DS} ; в режиме отсечки ток стока равен 0. Таким образом, в произвольном режиме работы ВАХ n -канального МДПТ описываются следующими уравнениями:

$$I_D = \begin{cases} 0, & V_{gst} < 0, \\ \beta V_{ds} (V_{gst} - V_{ds} / 2) - \\ - \beta K_B \left[\frac{2}{3} [(2\varphi_B - V_{bs} + V_{ds})^{3/2} - (2\varphi_B - V_{bs})^{3/2}] - \right. \\ \left. - V_{ds} (2\varphi_B - V_{bs})^{1/2} \right], & V_{gst} \geq 0, V_{ds} \leq V_{dsS}, \\ I_{DS}, & V_{gst} \geq 0, V_{ds} < V_{dsS}, \end{cases} \quad (\text{III.4.10})$$

где пороговое напряжение V_t , напряжение насыщения V_{dsS} и ток насыщения I_{DS} определяются соотношениями (III.4.1), (III.4.7) и (III.4.9).

Эффект изменения ширины ОПЗ вдоль канала приводит к явной зависимости ВАХ от напряжения подложка—исток V_{bs} , которое в идеализированной модели влияло только на величину порогового напряжения. Таким образом, число аргументов в уравнениях ВАХ увеличивается до трех: V_{gs} , V_{ds} и V_{bs} . Возрастает и число параметров статических характеристик: к пороговому напряжению V_t и параметру β добавляются напряжение φ_B и коэффициент подложки K_B .

Влияние неоднородности ОПЗ под затвором на ВАХ МДПТ иллюстрируется рис. III.4.1. Сплошные линии представляют уравнения (III.4.10), штриховые — уравнения идеализированной модели (III.3.17а). Пунктиром обозначены границы раздела крутых и пологих областей ВАХ. Учет изменения ширины ОПЗ вдоль канала приводит к уменьшению напряжения насыщения V_{dsS} и тока насыщения I_{DS} . Физической причиной этого является расширение ОПЗ по направлению от истока к стоку (см. рис. III.3.1), которое приводит к увеличению в этом направлении локального порогового напряжения $V_t(y)$.

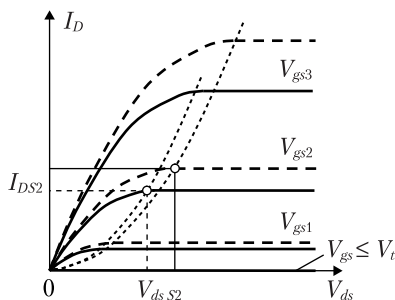


Рис. III.4.1. Влияние неоднородности ОПЗ под затвором на ВАХ МДПТ.

Сплошные линии — уравнения (III.4.10),
штриховые — уравнения идеализированной модели (III.3.17а).

Пунктир — границы раздела крутых
и пологих областей ВАХ

Разумеется, напряжение насыщения V_{dsS} и ток насыщения I_{DS} , как и в идеализированной модели, уменьшаются при увеличении запирающего напряжения подложка—исток V_{bs} вследствие возрастания порогового напряжения транзистора V_t (III.4.1).

III.4.2. Подпороговый ток

Идеализированная модель МДП-транзистора предполагает, что при напряжении затвор—исток ниже порогового (т.е. в режиме слабой инверсии, см. параграф III.1.3) проводимость канала равна нулю и ток в канале отсутствует (допущение 7 в параграфе III.3.1). Реально электроны существуют в канале и при напряжениях $V_{gs} < V_t$, хотя их поверхностная концентрация при этом меньше концентрации основных носителей в подложке. Поэтому ток стока при $V_{ds} > 0$ и $V_{gs} < V_t$ не прекращается полностью, хотя и быстро убывает при снижении напряжения V_{gs} . Ток стока в таком режиме называют *подпороговым*.

Поскольку концентрация электронов в канале в режиме слабой инверсии весьма мала, в подпороговом режиме основную роль играет диффузионный механизм переноса носителей, который не учитывался в идеализированной модели (допущение 4 в параграфе III.3.1).

ком $n_{st}(0)$ равна концентрации примеси в подложке N_B . При этом из (III.4.11в) получаем

$$n_s(0) = N_B \exp \frac{\varphi_s(0) - \varphi_{st}(0)}{\varphi_T}, \quad (\text{III.4.12})$$

где $\varphi_{st}(0)$ — поверхностный потенциал на границе с истоком при $V_{gs} = V_t$.

3. В режиме слабой инверсии не происходит экранирования каналом электрического поля E_x в полупроводнике, как это имеет место в режиме сильной инверсии. Поэтому при изменении напряжения V_{gs} относительно порогового напряжения V_t поверхностный потенциал на границе с истоком $\varphi_s(0)$ пропорционально изменяется относительно порогового значения $\varphi_{st}(0)$. Коэффициент пропорциональности определяется коэффициентом передачи емкостного делителя напряжения (рис. III.4.2, а)

$$k_C = \frac{C_{sG}}{C_{sG} + C_{sB}(0)} = \frac{1}{1 + C_{sB}(0) / C_{sG}} < 1, \quad (\text{III.4.13})$$

где C_{sG} — удельная емкости диэлектрика;

$$C_{sB}(0) = \varepsilon \varepsilon_0 / l(0) \quad (\text{III.4.14})$$

— удельная емкость ОПЗ при $y = 0$. Таким образом,

$$\varphi_s(0) - \varphi_{st}(0) = k_C V_{gst}. \quad (\text{III.4.15})$$

Подставляя (III.4.15) в (III.4.12), получаем

$$n_s(0) = N_B e^{k_C V_{gst} / \varphi_T}. \quad (\text{III.4.16a})$$

4. Разность поверхностных потенциалов на границах с истоком и стоком равна напряжению сток-исток:

$$\varphi_s(0) - \varphi_s(L) = V_{ds}.$$

При этом из (III.4.11в) получаем

$$n_s(L) = n_s(0) e^{-V_{ds} / \varphi_T}. \quad (\text{III.4.16б})$$

Соотношения (III.4.16а, б) и (III.4.13) позволяют произвести оценочный расчет ВАХ подпорогового тока. Полагая, как отмечено выше, что подпороговый ток имеет чисто диффузионный механизм, запишем

$$I_{D \text{ dif}} = e Z h_{ef} D_n \frac{dn_s}{dy} \approx e Z h_{ef} D_n \frac{n_s(0) - n_s(L)}{L}, \quad (\text{III.4.17})$$

где h_{ef} — эффективная толщина области слабой инверсии, а в качестве производной dn_s / dy принято ее среднее значение на



длине канала. Подставляя в (III.4.17) соотношения (III.4.16а, б), получаем ВАХ подпорогового тока

$$I_{D\text{ dif}} \approx I_{D\text{ dif}0} e^{k_C V_{gs} / \Phi_T} (1 - e^{-V_{ds} / \Phi_T}), \quad (\text{III.4.18})$$

где $I_{D\text{ dif}0} = e Z h_{ef} D_n N_B / L$. Эффективная толщина области слабой инверсии h_{ef} при пороговом напряжении близка к дебаевской длине экранирования в подложке

$$h_{ef} \approx \sqrt{\epsilon \epsilon_0 \Phi_T / e N_B}.$$

Таким образом, масштабный ток $I_{D\text{ dif}0}$ в (III.4.18) можно оценить соотношением

$$I_{D\text{ dif}0} = Z D_n \sqrt{e N_B \epsilon \epsilon_0 \Phi_T} / L. \quad (\text{III.4.19})$$

Отметим, что коэффициент k_C , определенный в (III.4.13), зависит от напряжения подложка—исток. На границе с истоком ширина ОПЗ $l(0)$ близка к ширине p - n перехода подложка—исток. Таким образом, удельная емкость ОПЗ в (III.4.13) определяется соотношением

$$C_{sB}(0) = C_{sB0} \sqrt{\Phi_{BS} / (\Phi_{BS} - V_{bs})},$$

где Φ_{BS} и C_{sB0} — контактная разность потенциалов и удельная емкость p - n перехода подложка—исток при $V_{bs} = 0$ и

$$k_C = \frac{1}{1 + C_{sB0} \sqrt{\Phi_{BS} / (\Phi_{BS} - V_{bs})} / C_{sG}}.$$

С учетом (III.2.12) получаем

$$k_C = \frac{1}{1 + K_B \sqrt{\Phi_{BS} / 8 \Phi_B (\Phi_{BS} - V_{bs})}}. \quad (\text{III.4.20})$$

Соотношение (III.4.18) показывает, что выходные ВАХ диффузионного подпорогового тока насыщаются при напряжении $V_{ds} > 3\Phi_T$. Проходные ВАХ представлены на рис. III.4.3. Их наклон в полулогарифмическом масштабе $1/m = k_C$ возрастает с увеличением запирающего напряжения исток—подложка в соответствии с (III.4.13) и (III.4.20). Это напряжение также сдвигает ВАХ в сторону больших напряжений V_{gs} вследствие увеличения порогового напряжения, как это следует из (III.4.1).

Соотношение (III.4.18) справедливо для области слабой инверсии в канале ($V_i < V_{gs} \leq V_t$). При $V_{gs} < V_i$ в приповерхностной области полупроводника реализуется режим обогащения поверхности основными носителями — дырками (см. параграф III.2.1). В этом режиме концентрация основных носителей в ОПЗ больше, чем неосновных, и области стока и истока



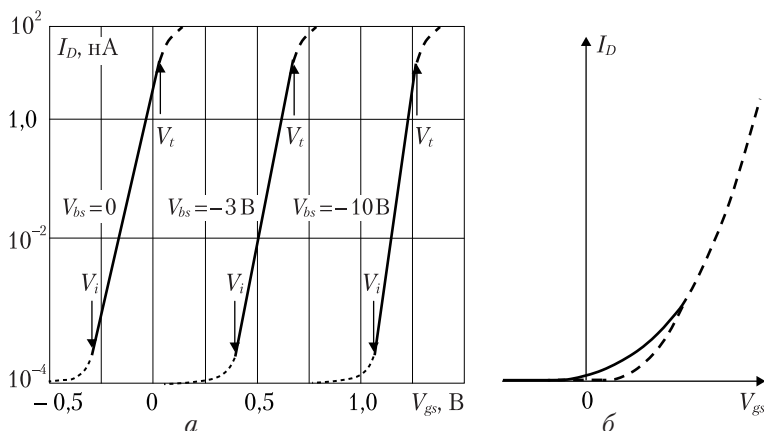


Рис. III.4.3. Прходные ВАХ МДПТ в режиме малых токов:
 а — в полулогарифмическом масштабе для трех значений V_{bs} ;
 б — в линейном масштабе. Сплошные линии — диффузионный ток, штриховые — дрейфовый, пунктир — ток утечки p - n перехода сток—подложка

разделены двумя встречно включенными p - n переходами. Ток стока при этом определяется токами утечки закрытого перехода сток—подложка (пунктирные линии на рис. III.4.3) и не равен току истока.

В режиме сильной инверсии ($V_{gs} > V_t$) поверхностная плотность заряда электронов в канале $Q_{sn}(y)$, пропорциональная их поверхностной концентрации $n_s(y)$, зависит от напряжения затвор—исток не экспоненциально, а линейно (соотношение (III.3.9)). При этом рост диффузионного тока с напряжением V_{gs} значительно замедляется и превалирует дрейфовый ток (штриховые линии на рис. III.4.3).

Для оценки диффузионного тока в режиме сильной инверсии положим, что распределение поверхностной плотности заряда неосновных носителей в канале линейно по координате y . Тогда в соответствии с (III.3.9) и с учетом (III.3.12) получаем

$$I_{Ddif} \approx ZD_n [Q_{sn}(0) - Q_{sn}(L)] / L \approx ZD_n C_{sG} (V_{gs} - V_{gd}) / L = \beta \phi_T V_{ds}.$$

Сравнение этого соотношения с (III.3.17а) показывает, что в режиме сильной инверсии диффузионный ток мал по сравнению с дрейфовым при $V_{gst} \gg \phi_T$. Этот результат оправдывает допущение 4 в параграфе III.3.1 о малости диффузионного тока.

В заключение отметим, что несмотря на малость диффузионного тока, в отдельных случаях он имеет существенное значение. В качестве примера можно указать на транзисторы, связывающие элементы хранения информации в накопителях запоминающих устройств с шинами считывания. При считывании информации во всем столбце накопителя закрыты все эти транзисторы, кроме связанного с выбранной словарной шиной. Таким образом, среди большого количества объединенных по электродам стока транзисторов открыт только один. Ток стока этого транзистора должен быть много больше суммы токов всех остальных транзисторов, что предъявляет жесткие требования к малости подпороговых токов.

III.4.3. Подвижность носителей заряда в канале

Как очевидно из энергетической диаграммы на рис. III.2.6, электроны в канале находятся под действием сильного поперечного электрического поля E_x . Это поле искривляет траектории их движения на длине свободного пробега, что приводит к снижению подвижности μ_n в направлении исток—сток. Другая причина снижения подвижности носителей в канале состоит в том, что поперечное поле E_x отклоняет их к поверхности полупроводника, где всегда повышена концентрация центров рассеяния.

Зависимость $\mu_n(E_x)$, измеренная для кремния экспериментально, приведена в работе [10, § 8.2.4]. Приближенно она описывается соотношением

$$\mu_n(E_x) = \mu_{n0} / (1 + E_x / E_0),$$

где $E_0 \approx 3 \cdot 10^5$ В/см.

Аналитический учет зависимости подвижности носителей в канале от внешних напряжений представляется чрезмерно сложным.

III.4.4. Влияние температуры

Влияние температуры на ВАХ МДП-транзистора проявляется в изменении параметра β , порогового напряжения, а также подпорогового тока.

Температурная зависимость параметра β связана главным образом с температурной зависимостью подвижности носителей в канале

$$\beta(T) = ZC_{sg}\mu(T) / L.$$



Поскольку подложка обычно легирована достаточно слабо, основным механизмом рассеяния носителей заряда в канале является рассеяние на акустических фононах (см. раздел I.4). Поэтому с ростом температуры подвижность снижается по закону, близкому к степенному (I.4.7a)

$$\mu(T) = \mu(T_0)(T / T_0)^{-\kappa},$$

где $\kappa = 1 - 1.5$. Пропорционально параметру β уменьшается ток стока при заданных напряжениях.

В формулах (III.2.56), (III.2.6), (III.1.21) и (III.4.1) для порогового напряжения от температуры наиболее сильно зависит параметр ϕ_B . С учетом этих соотношений и определения коэффициента подложки $K_B = \sqrt{2\epsilon\epsilon_0 e N_B} / C_{sG}$ температурный коэффициент порогового напряжения составляет

$$\frac{dV_t}{dT} = \frac{dV_t}{d\phi_B} \cdot \frac{d\phi_B}{dT}, \quad (\text{III.4.21})$$

$$\text{где } \frac{dV_t}{d\phi_B} = \pm 1 \pm \frac{\sqrt{2\epsilon\epsilon_0 e N_B}}{C_{sG} \sqrt{2\phi_B - V_{bs}}} = \pm 1 \pm \frac{K_B}{\sqrt{2\phi_B - V_{bs}}}. \quad (\text{III.4.22})$$

Производная $d\phi_B / dT$ может быть найдена из (III.1.21):

$$\frac{d\phi_B}{dT} = \frac{\phi_T}{T} \ln \frac{N_B}{n_i} - \frac{\phi_T}{n_i} \frac{dn_i}{dT} = \frac{\phi_B}{T} - \frac{\phi_T}{n_i} \frac{dn_i}{dT}.$$

Учитывая, что $n_i = \sqrt{N_c N_v} e^{-E_g/2kT}$ и $dn_i / dT = n_i E_g / 2kT^2$, получаем

$$d\phi_B / dT = (\phi_B - V_g / 2) / T, \quad (\text{III.4.23})$$

где $V_g = E_g / e$ — ширина запрещенной зоны в вольтах.

Подстановка (III.4.22) и (III.4.23) в (III.4.21) дает

$$\frac{dV_t}{dT} = \mp \frac{\phi_g / 2 - \phi_B}{T} \cdot \left(1 + \frac{K_B}{\sqrt{2\phi_B - V_{bs}}} \right). \quad (\text{III.4.24})$$

Здесь, как и выше, в двойных знаках « $\pm$, $\mp$ » верхнее значение относится к n -канальному МДПТ.

Отметим, что пороговое напряжение n -канальных транзисторов с ростом температуры снижается, а p -канальных — возрастает. Это означает, что для любого типа канала ток I_D при заданных напряжениях возрастает с температурой.

Абсолютная величина температурного коэффициента порогового напряжения лежит обычно в диапазоне 1,2–3 мВ/К



и снижается при уменьшении коэффициента подложки, т.е. при уменьшении концентрации примеси в подложке и толщины подзатворного диэлектрика.

Температурная зависимость подпорогового тока связана в основном с последним экспоненциальным множителем в (III.4.18). Максимальная величина подпорогового тока (при $V_{gs} = V_t$, $V_{ds} > 3\phi_T$) слабо изменяется с температурой. При повышении температуры возрастает параметр $m\phi_T$, характеризующий наклон ВАХ подпорогового тока в полулогарифмическом масштабе, а также уменьшается (для n -канальных МДПТ) пороговое напряжение. Поэтому для снижения подзатворного тока до требуемого уровня необходимо более глубокое запираение транзистора по затвору.

III.4.5. Умножение носителей в канале

При достаточно высоком напряжении сток—исток в области отсечки канала происходит слабый лавинный пробой. Генерированные при этом электроны уходят в сток, а генерированные дырки собираются подложкой, создавая ток I_B . На рис. III.4.4 в полулогарифмическом масштабе представлены

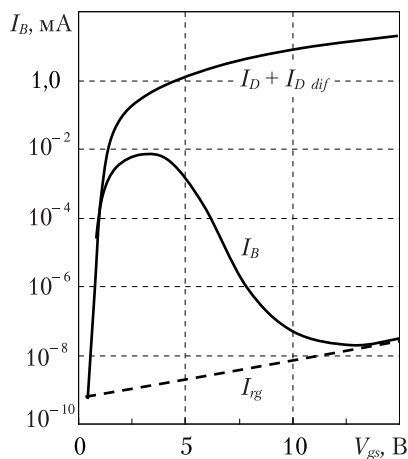


Рис. III.4.4. Проходная ВАХ МДПТ при наличии ударной ионизации электронно-дырочных пар в области сильного поля вблизи стока ($V_{ds} = 10$ В, $V_{bs} = 5$ В).

Штриховая линия — ток генерации—рекомбинации в ОПЗ вблизи стока без учета лавинного умножения

типичная проходная ВАХ, а также зависимость тока подложки I_B от напряжения V_{gs} [10]. Штриховая линия соответствует току генерации—рекомбинации в ОПЗ вблизи стока без учета лавинного умножения. Как очевидно из рисунка, ток подложки вначале возрастает при увеличении напряжения V_{gs} , а затем снижается. Объяснение этого эффекта состоит в том, что при наличии ударной ионизации ток I_B пропорционален исходному току стока

$$I_B = M(\bar{E}_M)I_D,$$

где $M(\bar{E}_M)$ — коэффициент размножения носителей; $\bar{E}_M$ — средняя напряженность поля в области ударной ионизации.

При заданном напряжении V_{ds} увеличение напряжения V_{gs} , с одной стороны, ведет к увеличению тока I_D , а с другой — к увеличению напряжения насыщения $V_{dsS} = V_{gs} - V_t$. Последний фактор приводит к снижению продольного поля $\bar{E}_M$ на участке ударной ионизации и уменьшению коэффициента размножения $M(\bar{E}_M)$, что и объясняет уменьшение тока I_B .

Основные выводы

1. Изменение ширины ОПЗ вдоль канала МДП-транзистора приводит к некоторому уменьшению тока стока и напряжения насыщения по сравнению с идеализированной моделью. ВАХ транзистора с учетом этого эффекта описывается соотношением (III.4.10). К числу аргументов ВАХ добавляется напряжение подложка—исток, а к числу параметров — напряжение ϕ_B и коэффициент подложки K_B .

2. При $V_i < V_{gs} \leq V_t$ в подзатворной области МДПТ реализуется режим слабой инверсии. Подпороговый ток в этом режиме имеет диффузионную природу, экспоненциально зависит от напряжения затвор—исток и насыщается при напряжении сток—исток более $3\phi_T$. Запирающее напряжение подложка—исток сдвигает проходные ВАХ в сторону больших напряжений затвор—исток из-за увеличения порогового напряжения. ВАХ диффузионного подпорогового тока определяется соотношениями (III.4.18) и (III.4.19).

3. При напряжении затвор—исток, меньшем напряжения инверсии V_i , ток стока определяется токами утечки закрытого p - n перехода сток—подложка и не равен току истока. При $V_{gs} > V_t$ диффузионная составляющая тока стока значительно меньше дрейфовой.

4. Подвижность носителей заряда в канале снижается с увеличением электрического поля, перпендикулярного поверхности полупроводника, что приводит к снижению тока стока.



5. Влияние температуры на ВАХ МДП-транзистора в режиме сильной инверсии проявляется главным образом в изменении подвижности носителей в канале, а также порогового напряжения. Температурный коэффициент порогового напряжения n -канальных транзисторов отрицателен, а p -канальных — положителен. Его величина определяется соотношением (III.4.24). В режиме слабой инверсии с ростом температуры уменьшается наклон ВАХ подпорогового тока в полулогарифмическом масштабе.

III.5. Эффекты короткого канала

III.5.1. Природа эффектов короткого канала

Соотношения (III.3.17), (III.4.10) и (III.4.18), определяющие ВАХ МДПТ, показывают, что во всех режимах работы транзистора ток стока не изменяется при пропорциональном изменении длины и ширины канала. Ширина Z канала определяет масштаб ВАХ по току, не изменяя частотных свойств транзистора (см. параграф III.3.3). Уменьшение длины канала L способствует повышению удельной крутизны ВАХ ($g_{sz} \sim 1/L$, (III.3.27)) и предельной частоты ($\omega_T \sim 1/L^2$, (III.3.40)). Поэтому в течение длительного периода времени основной тенденцией развития техники МДПТ было и остается уменьшение длины канала L .

Отмеченные закономерности оставались справедливыми до тех пор, пока топологические размеры транзисторов оставались достаточно большими. Когда продольные размеры МДПТ достигли значений $\approx 1\text{--}2$ мкм, эти закономерности перестали выполняться. Кроме того, оказалось, что при уменьшении топологических размеров (особенно длины канала) изменяется и сама форма ВАХ. Причины проявления *эффектов короткого канала* можно условно разделить на три группы.

К 1-й группе относятся эффекты, связанные с влиянием *краевых областей* ОПЗ под контактами стока и истока, на распределение электрического поля под затвором. Влияние этих эффектов проявляется в изменении порогового напряжения V_t и одного из основных параметров транзистора β , определяющего масштаб ВАХ.

Ко 2-й группе относятся эффекты, связанные с разогревом носителей заряда в продольном электрическом поле, напряженность которого возрастает с уменьшением длины канала. При этом изменяются их кинетические характеристики и, как



следствие, вид ВАХ в пологой области. Кроме того, появляется возможность проникновения носителей в диэлектрик.

Наконец, эффекты 3-й группы возникают вследствие появления новых механизмов протекания тока в цепи сток—исток, связанных с близким расположением электродов стока и истока. Эти эффекты приводят к резкому увеличению тока стока при повышении напряжения сток—исток и по своему проявлению близки к явлениям пробоя.

Несмотря на кажущуюся простоту конструкции и физики работы, достаточно точные математические модели МДП-транзистора весьма сложны. Главная причина состоит в том, что МДПТ является в принципе неоднородным прибором — управляющее поле затвора и тянущее поле стока направлены взаимно перпендикулярно. В простых рассмотренных выше моделях задача сводилась к квазиодномерной — взаимодействие продольного и поперечного полей не учитывалось. Этот подход по возможности сохраняется и при аналитическом описании более тонких эффектов.

III.5.2. Пороговое напряжение

Влияние длины канала на пороговое напряжение МДП-транзистора проявляется в искажении одномерного характера электрического поля в ОПЗ под затвором (рис. III.5.1, а) вследствие краевых эффектов. Этот эффект относится к 1-й группе эффектов короткого канала (см. параграф III.5.1). В первом приближении эффект учитывается заменой реальной поверхностной плотности заряда примесями ионов в ОПЗ под затвором Q_{sB} ее эффективным значением Q_{sBef} :

$$Q_{sB} = Q_B / ZL, \quad Q_{sBef} = \frac{Q_{Bef}}{ZL} = Q_{sB} \frac{S_y}{S_{y0}},$$

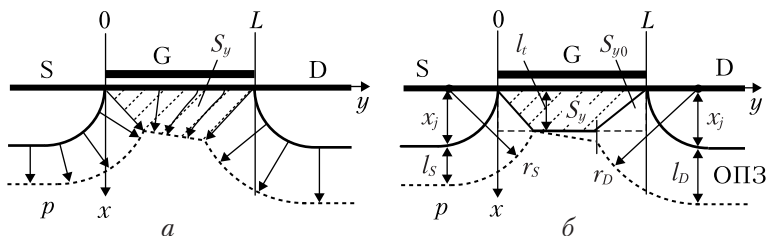


Рис. III.5.1. Силовые линии электрического поля в ОПЗ (а) и эффективный объем ОПЗ (б)

где Q_B — полный заряд примесных ионов; S_y — площадь продольного (в направлении оси y) сечения ОПЗ, в которой заряд ионов влияет на величину порогового напряжения; S_{y0} — площадь сечения ОПЗ в транзисторе с длинным каналом. Строго говоря, площадь S_y ограничена со стороны стока и истока силовыми линиями, проходящими через границы канала ($x=0, y=L$ и $x=0, y=0$ на рис. III.5.1, а), а снизу — границей ОПЗ. Практически принято считать, что границы ОПЗ у краев канала представляют собой дуги окружностей с радиусами $r_s = x_j + l_s$ и $r_D = x_j + l_D$, где x_j — толщина контактных n^+ -областей стока и истока, а l_s и l_D — ширина p - n переходов подложка—исток и подложка—сток (рис. III.5.1, б). Площадь S_y рассчитывается в предположении, что она ограничена трапецией, высота которой равна пороговой ширине ОПЗ под истоком l_t , а боковые стороны проведены через границы канала со стоком, истоком и точки пересечения упомянутых выше дуг с нижней стороной (см. рис. III.5.1, б).

Для транзистора с длинным каналом краевыми эффектами, искажающими форму ОПЗ, можно пренебречь и при напряжении $V_{ds} = 0$ (а именно так определено пороговое напряжение транзистора в параграфе III.2.4) $S_{y0} = Ll_t$.

Из соотношений (III.2.8) следует, что сдвиг порогового напряжения, обусловленный краевыми искажениями ОПЗ, составляет

$$\Delta V_t = -\frac{Q_{sBef} - Q_{sB}}{C_{sG}} = \frac{Q_{sB}}{C_{sG}} \left(1 - \frac{S_y}{S_{y0}} \right).$$

Учитывая, что $Q_{sB} = \mp e N_B S_{y0} Z$, из простых геометрических соображений получаем

$$\Delta V_t = \mp \frac{e N_B l_t}{2 L C_{sG}} (\sqrt{(x_j + l_s)^2 - l_t^2} + \sqrt{(x_j + l_D)^2 - l_t^2} - 2x_j) = \quad (\text{III.5.1a})$$

$$= \mp \frac{l_t K_B \sqrt{\Phi_B}}{l_{t0} L \sqrt{2}} (\sqrt{(x_j + l_s)^2 - l_t^2} + \sqrt{(x_j + l_D)^2 - l_t^2} - 2x_j), \quad (\text{III.5.1б})$$

где пороговая ширина ОПЗ под истоком l_t определяется соотношениями (III.2.7) и (III.2.4), а ширина ОПЗ под стоком и истоком — обычными формулами для ступенчатого p - n перехода:

$$l_s = l_{s0} \sqrt{(1 - V_{bs} / \Phi_{BS})}, \quad (\text{III.5.2a})$$

$$l_D = l_{D0} \sqrt{1 - (V_{bs} - V_{ds}) / \Phi_{BD}}, \quad (\text{III.5.2б})$$

$$l_t = l_{t0} \sqrt{1 - V_{bs} / 2\Phi_B}, \quad (\text{III.5.2в})$$

где $l_{s0} = l_{D0} = \sqrt{2\epsilon\epsilon_0\phi_{BS} / eN_B}$; $l_{t0} = \sqrt{2\epsilon\epsilon_0 \cdot 2\phi_B / eN_B}$; $\phi_{BS} = \phi_{BD}$ — контактные разности потенциалов переходов подложка—исток и подложка—сток. Соотношение (III.5.16) получено с учетом (III.2.11). Практически значения ϕ_{BS} и $2\phi_B$ очень близки, и в (III.5.1), (III.5.2) можно полагать $\phi_{BS} = 2\phi_B$, $l_{s0} = l_{D0} = l_{t0}$, $l_s = l_t$.

Анализ соотношения (III.5.1) позволяет сделать следующие выводы.

1. При уменьшении длины канала пороговое напряжение *n*-канальных МДПТ *уменьшается*, *p*-канальных — *увеличивается*. В обоих случаях это приводит к увеличению тока стока в заданном электрическом режиме. Сдвиг порогового напряжения обусловлен уменьшением абсолютной величины полного заряда примесных ионов под затвором.

2. Зависимость порогового напряжения от длины канала усиливается при *увеличении толщины x_j контактных n^+ -областей стока и истока*.

3. *Пороговое напряжение короткоканальных транзисторов зависит от напряжения сток—исток* (в соответствии с (III.5.2б) от V_{ds} зависит ширина ОПЗ под стоком l_D).

При малых значениях ширины канала Z на величину порогового напряжения оказывают влияние краевые эффекты на боковой границе затвора. На рис. III.5.2 показана форма ОПЗ под затвором в поперечном сечении МДПТ, перпендикулярном направлению исток—сток, при $V_{gs} = V_t$. Из рисунка очевидно, что краевые эффекты увеличивают абсолютную величину полного заряда примесных ионов под затвором Q_B , пропорциональную заштрихованной площади S_z . Эффективная поверхностная плотность заряда ионов составляет

$$Q_{sBef} = \mp eN_B l_t S_z / S_{z0}.$$

Полагая, что боковые границы ОПЗ являются секторами окружностей с радиусом l_t , находим

$$S_z / S_{z0} = 1 + \pi l_t / 2Z.$$

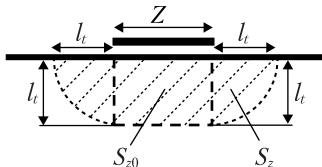


Рис. III.5.2. Форма ОПЗ под затвором МДПТ в поперечном сечении при $V_{gs} = V_t$

Таким образом, сдвиг порогового напряжения, обусловленный краевыми искажениями ОПЗ, составляет

$$\Delta V_t = -\frac{Q_{sB\text{ef}} - Q_{sB}}{C_{sG}} = \frac{Q_{sB}}{C_{sG}} \left(1 - \frac{S_z}{S_{z0}} \right) = \pm \frac{eN_{B^*} J_t^2 \pi}{2C_{sd} Z}.$$

С учетом (III.2.7) получаем

$$\Delta V_t = \pm \frac{\pi \epsilon \epsilon_0 (2\phi_B - V_{bs})}{C_{sG} Z}. \quad (\text{III.5.3})$$

Отметим, что в МДПТ с малыми топологическими размерами $S_y < S_{y0}$ (рис. III.5.1, б), но $S_z > S_{z0}$ (см. рис. III.5.2). Поэтому уменьшение длины и ширины канала изменяют пороговое напряжение в противоположные стороны.

Соотношение (III.5.3) показывает, что сдвиг порогового напряжения в МДПТ с малой шириной канала очень слабо зависит от концентрации примеси в подложке (через параметр ϕ_B) и определяется главным образом свойствами подзатворного диэлектрика (C_{sG}), а также напряжением V_{bs} .

III.5.3. Эффект смыкания канала

Эффект смыкания канала относится одновременно к 1-й и 3-й группам эффектов короткого канала (см. параграф III.5.1) и состоит в смыкании под затвором ОПЗ стока и истока при повышении напряжения сток—исток до некоторого значения $V_{ds} = V_{\text{pinch}}$. Смыкание ОПЗ происходит за счет расширения ОПЗ под стоком и ее распространения под затвор в сторону истока (рис. III.5.3, а).

На рис. III.5.3, б приведены энергетические диаграммы поверхности полупроводника для фиксированного напряжения затвор—исток и нескольких значений напряжения сток—исток V_{ds} (показано только расположение уровней E_c). Направление возрастания V_{ds} указано стрелкой. Отметим, что при отсутствии тока ($V_{ds} = 0$) в области канала уровень E_c расположен выше, чем в n^+ -областях стока и истока, что обусловлено значительно большей концентрацией электронов в этих областях (в вырожденных n^+ -областях канала уровень Ферми F находится несколько выше E_c). Таким образом, при $V_{ds} = 0$ область канала представляет собой потенциальный барьер между стоком и истоком (см. рис. III.5.3, б). Высота этого барьера ΔE_s , как и концентрация электронов в канале, является функцией напряжения затвор—исток.



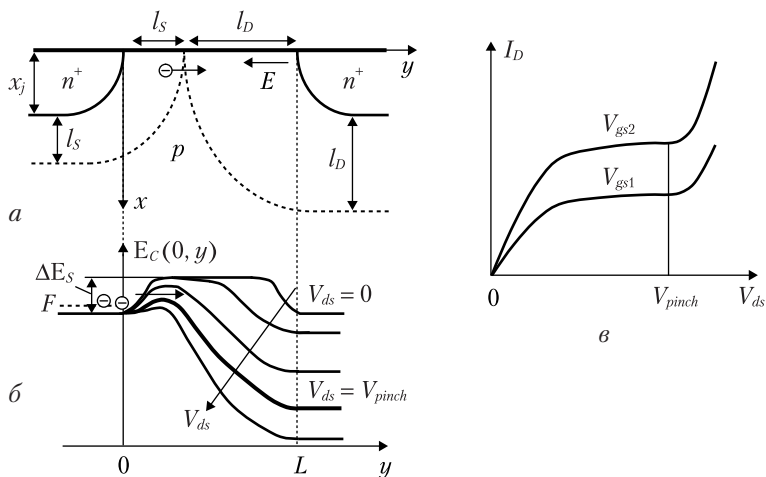


Рис. III.5.3. Эффект смыкания канала:

a — границы ОПЗ при $V_{ds} = V_{pinch}$; b — энергетические диаграммы поверхности полупроводника (стрелка указывает направление возрастания V_{ds}); $в$ — проявление эффекта смыкания канала на ВАХ МДПТ

При повышении напряжения сток—исток энергетический уровень E_c в канале снижается по направлению к области стока, имеющей максимальный потенциал. Если длина канала L достаточно велика, это никак не сказывается на высоте барьера ΔE_s .

Ширина ОПЗ под стоком l_D , которая определяется соотношением (III.5.2б), увеличивается с ростом напряжения V_{ds} вследствие расширения p - n перехода подложка—сток. Примерно на такое же расстояние ОПЗ под стоком расширяется в сторону истока под каналом. В случае короткого канала ближайшая к истоку граница этой ОПЗ оказывается в непосредственной близости к границе ОПЗ под истоком. При этом взаимодействие ОПЗ под стоком и под истоком приводит к снижению потенциального барьера ΔE_s между истоком и каналом. Фактически это эквивалентно повышению напряжения затвор—исток, которое в действительности остается постоянным. В результате повышается концентрация электронов в канале на границе с истоком (автоэмиссия электронов из истока в канал) и, как следствие, ток стока. Такое достаточно резкое повышение тока стока с ростом напряжения V_{ds} напоминает картину пробоя (рис. III.5.3, в).

Оно происходит при напряжении $V_{ds} = V_{pinch}$, соответствующем смыканию ОПЗ стока и истока на поверхности полупроводника. Это напряжение определяется условием смыкания (см. рис. III.5.3, а)

$$l_s + l_D(V_{pinch}) = L. \quad (III.5.4)$$

Рисунок III.5.3, б показывает, что при $V_{ds} = V_{pinch}$ высота потенциального барьера ΔE_s между истоком и каналом значительно снижается (жирная кривая $E_c(0, y)$ на рис. III.5.3, б).

С учетом (III.5.2б) напряжение смыкания составляет

$$V_{pinch} = (\phi_{BS} - V_{bs})[(L/l_s - 1)^2 - 1] \approx \quad (III.5.5a)$$

$$\approx (\phi_{BS} - V_{bs})(L/l_t - 1)^2 - 1 \approx \quad (III.5.5б)$$

$$\approx (\phi_{BS} - V_{bs}) \left[\left(\frac{L}{l_{t0}} \sqrt{1 - V_{bs}/2\phi_{BD}} - 1 \right)^2 - 1 \right]. \quad (III.5.5в)$$

Заметим, что при $V_{ds} > V_{pinch}$ эмиссия электронов из истока может происходить не только в поверхностную область канала, но и охватывать существенную часть ширины ОПЗ и даже проникать в подложку. Если напряжение V_{ds} существенно превышает величину V_{pinch} , ток, связанный с эмиссией электронов в ОПЗ, ограничивается пространственным зарядом избыточных электронов в ОПЗ (поскольку дрейфовая скорость электронов в сильных полях ограничена, увеличение тока возможно только за счет повышения концентрации электронов). В этих условиях ВАХ тока инжекции имеет вид [9, § 8.4.3]

$$I_D \approx 9\epsilon\epsilon_0\mu_n S_s V_{ds}^2 / 8L^3, \quad (III.5.6)$$

где S_s — площадь поперечного сечения ОПЗ на границе с истоком.

В режиме смыкания канала на выходных характеристиках пологая область вырождается в крутую (см. рис. III.5.3, в), что приводит к снижению собственного коэффициента усиления МДПТ по напряжению K_v (см. параграф III.3.2) и резкому ухудшению качества прибора. По существу эффект смыкания канала может считаться видом пробоя.

Напряжение смыкания чрезвычайно сильно зависит от длины канала, резко снижаясь при ее уменьшении. При условии $L \leq 2l_s$ смыкание канала происходит уже при нулевом напряжении сток—исток. Для повышения напряжения смыкания следует увеличивать концентрацию примеси в подложке N_B , в результате чего снижаются значения l_s и l_D .

III.5.4. Зарядка оксида

При высоких продольных электрических полях E_y , которые имеют место в транзисторах с коротким каналом, часть электронов в канале разогревается до энергии, достаточной для преодоления энергетического барьера на границе диэлектрик—полупроводник (для системы $\text{SiO}_2\text{—Si}$ высота этого барьера ≈ 3 эВ). Такие горячие электроны, а также электроны из прилегающей к стоку ОПЗ могут инжектироваться в оксид. Кроме того, в оксид могут инжектироваться электроны, генерированные в объеме подложки и разогретые большим поперечным полем на пути к диэлектрику.

Главным следствием эффекта зарядки оксида является сдвиг порогового напряжения в сторону положительных значений благодаря изменению заряда $Q_{ss\,ef}$ в (III.2.5) и (III.2.8). Зарядка оксида существенно снижает стабильность характеристик прибора и в конечном счете срок его службы, так как приводит к непрерывной деградации рабочих характеристик.

Этот вредный эффект можно отнести ко 2-й группе эффектов короткого канала (см. параграф III.5.1).

III.5.5. Ограничение дрейфовой скорости носителей в канале

Этот эффект относится ко 2-й группе эффектов короткого канала и связан с разогревом носителей заряда в продольном электрическом поле. При уменьшении длины канала возрастает напряженность продольного электрического поля в канале E_y , определяющего дрейфовый ток стока. При этом дрейфовая скорость носителей насыщается с ростом поля (см. раздел I.4, рис. I.4.1). Наиболее простые аналитические результаты получаются при аппроксимации зависимости $v(E)$ кусочно-линейной функцией (I.1.4.9)

$$v(E) = \begin{cases} \mu E, & E \leq E_s, \\ v_s = \mu E_s, & E > E_s, \end{cases} \quad (\text{III.5.7})$$

где μ — подвижность носителей в слабом поле. Начальные подвижности зависят от концентрации примеси в полупроводнике. Предельные скорости дрейфа электронов и дырок в кремнии составляют

$$v_{sn} \approx 1,1 \cdot 10^7 \left(\frac{T_0}{T} \right)^{0,87} \text{ см/с}; \quad v_{sp} \approx 0,95 \cdot 10^7 \left(\frac{T_0}{T} \right)^{0,52} \text{ см/с, где } T_0 = 300 \text{ К.}$$



Эти значения практически не зависят от концентрации примеси, так как скорость ограничивается при достижении носителями энергии оптического фонона. Напряженность поля E_s , соответствующая достижению предельной скорости, в зависимости от степени легирования составляет

$$E_s = \frac{v_s}{\mu} \approx \begin{cases} 15 - 25 \text{ кВ/см} & \text{— для } n\text{-Si}, \\ 30 - 40 \text{ кВ/см} & \text{— для } p\text{-Si}. \end{cases}$$

Анализируя влияние зависимости (III.5.7) на ВАХ МДП-транзистора, следует принять во внимание следующие соображения.

1. Вывод ВАХ проводился ранее (см. параграфы III.3.2, III.4.1) в предположении, что поверхностная плотность заряда электронов в канале *локально* зависит от потенциала поверхности канала (допущение 8 в параграфе III.3.1 — приближение плавного канала). Такое допущение не может выполняться на участке канала, где скорость электронов достигает насыщения. Действительно, потенциал поверхности этого участка $V(y)$ (см. рис. III.3.1) изменяется вдоль оси y , а скорость электронов постоянна $v(y) = v_{sn}$. Поскольку ток в любом сечении канала равен току стока I_D , поверхностная плотность заряда электронов $Q_{sn} = I_D / v_{sn} Z$ также должна быть постоянной. Таким образом, *допущение о локальной зависимости $Q_{sn}[V(y)]$ нельзя считать справедливым на участке канала, где $E_y > E_s$.*

2. При малых напряжениях V_{ds} , когда во всем канале $E_y < E_s$ и локальный характер зависимости $Q_{sn}(V)$ сохраняется, максимальная напряженность поля достигается на стоковой границе канала (в плоскости $y = L$ на рис. III.3.1), где потенциал поверхности максимален, а величина Q_{sn} минимальна.

3. При повышении напряжения V_{ds} до некоторого значения V_{dsS} значение $E_y = E_s$ достигается в точке $y = L$, где поверхностная плотность заряда электронов в канале составляет $Q_{sn} = I_D / v_{sn} Z$. Дальнейшее повышение напряжения V_{ds} не приводит к повышению дрейфовой скорости электронов в плоскости $y = L$. Поэтому возрастание тока стока возможно только за счет уменьшения *эффективной длины канала* L_{ef} , на которой напряженность поля $E_y < E_s$.

4. В рассмотренных выше моделях использовано допущение о постоянной подвижности носителей заряда в канале (допущение 5 в параграфе III.3.1). В пологой области ВАХ ($V_{ds} > V_{dsS}$) канал считался полностью перекрытым на границе со стоком. При этом $Q_{sn}(L) = 0$ и протекание тока



$I_D = I_{DS}$ может быть обеспечено только при *бесконечной скорости электронов* в плоскости $y = L$. При постоянной подвижности электронов бесконечная скорость соответствует *бесконечной напряженности поля* $E_y(L)$, что и показано на рис. III.3.1, б. Заметим, что при $V_{ds} > V_{dsS}$ напряженность поля, согласно модели, бесконечна только в одной точке, так что напряжение $V_{ds} = \int_0^L E_y dy$ остается конечным. Разумеется,

реально поля всегда конечны, а использованное допущение — всего лишь абстракция, дающая правдоподобный результат. Реально *приближение плавного канала* (или эквивалентное ему допущение $E_y \ll E_x$ параграфа III.3.1) *не может выполняться вблизи стока в пологой области ВАХ*.

5. Ограничение дрейфовой скорости носителей в канале приводит к тому, что канал на границе со стоком не перекрывается полностью. Таким образом, напряжение насыщения V_{dsS} , соответствующее достижению на границе со стоком конечной напряженности поля E_S , должно быть меньше значения $V_{dsS} = V_{gst}$ в идеализированной модели КМДП или даже в модели, учитывающей неоднородность ОПЗ под затвором (III.4.7).

Таким образом, учет ограничения дрейфовой скорости носителей в канале сводится к нахождению напряжения V_{dsS} , ограничивающего круглую область ВАХ. При $V_{ds} \leq V_{dsS}$ ВАХ описываются моделью с постоянной подвижностью, а при $V_{ds} > V_{dsS}$ ток стока изменяется только вследствие изменения эффективной длины канала, рассмотренного ниже.

III.5.6. Влияние ограничения дрейфовой скорости носителей в канале на основные свойства МДПТ

Анализ ВАХ МДПТ с учетом ограничения дрейфовой скорости носителей в канале мы проведем, пренебрегая, как и в идеализированной модели, неоднородностью ОПЗ под затвором (допущение 6 в параграфе III.3.1). При этом распределение поля в канале для пологой области ВАХ определяется соотношением (III.3.15). Подставляя в него значение тока стока из (III.3.16а) или из (III.3.17а), при $y = L$ получаем

$$E_y(L)L = \frac{V_{ds}(V_{gst} - V_{ds}/2)}{V_{gst} - V_{ds}}.$$



При $V_{ds} = V_{dsS}$ должно выполняться условие $E_y(L) = E_s$.
Отсюда

$$V_{dsS} = V_{gst} + V_L - \sqrt{V_{gst}^2 + V_L^2} \rightarrow \quad (\text{III.5.8a})$$

$$\rightarrow \begin{cases} V_{gst}, & V_{gst} \ll V_L, \\ V_L, & V_{gst} \gg V_L, \end{cases} \quad (\text{III.5.8б})$$

где $V_L = E_s L. \quad (\text{III.5.9})$

Заметим, что всегда выполнено неравенство

$$V_{dsS} < V_L. \quad (\text{III.5.10})$$

Подстановка (III.5.8a) в (III.3.16a) или (III.3.17a) дает значение тока насыщения

$$I_{DS} = \beta V_L (\sqrt{V_{gst}^2 + V_L^2} - V_L) \rightarrow \quad (\text{III.5.11a})$$

$$\rightarrow \begin{cases} \beta V_{gst}^2 / 2, & V_{gst} \ll V_L, \\ \beta V_L V_{gst}, & V_{gst} \gg V_L. \end{cases} \quad (\text{III.5.11б})$$

Исключая из (III.5.8a) и (III.5.11a) напряжение V_{gst} , получим уравнение $I_{DS}(V_{dsS})$, соответствующее границе крутой и пологой областей ВАХ

$$I_{DS} = \beta V_L^2 \left[\sqrt{1 + \left(\frac{V_{ds}}{2V_L} \right)^2 \left(\frac{V_{ds} - 2V_L}{V_{ds} - V_L} \right)^2} - 1 \right]. \quad (\text{III.5.12})$$

Таким образом, ограничение дрейфовой скорости носителей в канале не изменяет вида ВАХ (III.3.16a) в крутой области, однако крутая область заканчивается при меньшем напряжении насыщения и, соответственно при меньшем токе насыщения.

Уравнения (III.5.8) и (III.5.11) показывают, что ограничение дрейфовой скорости носителей учитывается введением одного *дополнительного параметра* — напряжения $V_L = E_s L$, а *степень проявления этого эффекта определяется отношением* V_{gst} / V_L .

На рис. III.5.4 показан вид ВАХ МДПТ с учетом (жирные линии) и без учета (тонкие линии) ограничения дрейфовой скорости носителей в канале, а также (на рис. III.5.4, а) границы крутой и пологой областей ВАХ. Штриховыми линиями на рис. III.5.4, а обозначены значения тока насыщения. При расчетах приняты следующие значения параметров:

$\beta = 1 \text{ мА/В}^2$; $V_t = 1 \text{ В}$; $V_L = 1 \text{ В}$ (длина канала $L \approx 0,5 \text{ мкм}$). Изломы выходных ВАХ связаны с использованием кусочно-линейной аппроксимации $v(E)$ (III.5.7).

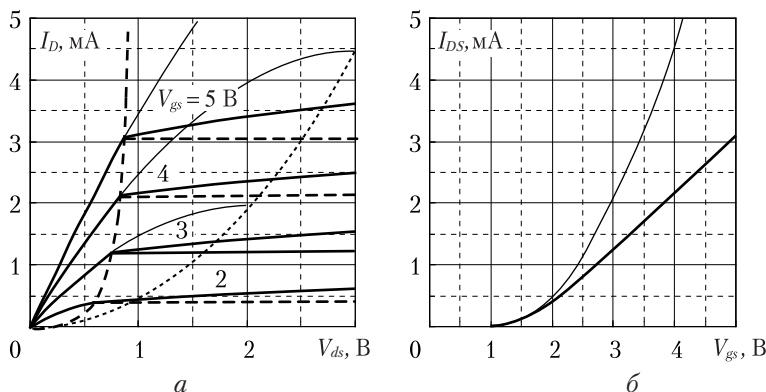


Рис. III.5.4. ВАХ МДПТ с учетом (жирные линии) и без учета (тонкие линии) ограничения дрейфовой скорости носителей в канале:

a — выходные; b — зависимости $I_{DS}(V_{gs})$. Пунктирные линии на рис. a — границы крутой и пологой областей ВАХ. Параметры МДПТ: $V_t = 1 \text{ В}$; $V_L = 1 \text{ В}$; $\beta = 1 \text{ мА/В}^2$. Штриховые линии на рис. a не учитывают эффект модуляции длины канала

Как очевидно из рис. III.5.4, b и соотношений (III.5.11), при напряжении $V_{gst} \gg V_L$ проходная ВАХ тока насыщения $I_{DS}(V_{gs})$ вырождается в линейную. Из (III.5.11а) следует, что крутизна ВАХ на границе крутой и пологой областей составляет

$$g_s = \frac{\partial I_{DS}}{\partial V_{gs}} = \frac{\beta V_L V_{gst}}{\sqrt{V_{gst}^2 + V_L^2}}.$$

Учитывая, что $\beta = Z\mu C_{sg}/L$ (III.3.12), $V_L = E_s L$ (III.5.9) и $\mu E_s = v_s$ (III.5.7), крутизну удобно представить в следующем виде:

$$g_s = g_{\max} \frac{V_{gst}/V_L}{\sqrt{1 + V_{gst}^2/V_L^2}} \rightarrow \quad (III.5.13a)$$

$$\rightarrow \begin{cases} g_{\max} V_{gst}/V_L, & V_{gst} \ll V_L, \\ g_{\max}, & V_{gst} \gg V_L, \end{cases} \quad (III.5.13б)$$

$$\text{где} \quad g_{\max} = \beta V_L = Zv_s C_{sg}. \quad (III.5.14)$$

На рис. III.5.5 представлена зависимость $g_s / g_{\max}(V_{gst} / V_L)$. Из рисунка очевидно, что крутизна нарастает с ростом напряжения затвор—исток только до тех пор, пока это напряжение остается достаточно малым. При $V_{gst} \gg V_L$ крутизна достигает максимального значения $g_{\max}$, определенного в (III.5.14). Таким образом, параметр

$$g_{\max z} = g_{\max} / Z = v_s C_{sG} \quad (\text{III.5.15})$$

имеет смысл *предельно достижимой удельной крутизны ВАХ* МДПТ. Отметим, что *ее значение не зависит от длины канала* и определяется только *удельной емкостью диэлектрика* C_{sG} и *предельной дрейфовой скоростью носителей заряда* в канале v_s .

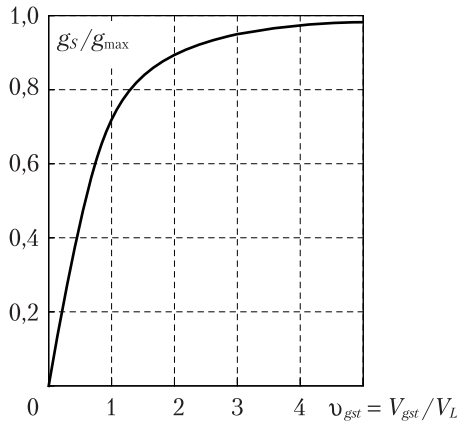


Рис. III.5.5. Зависимость крутизны на границе пологой области ВАХ от напряжения затвор—исток

Рисунок III.5.5 показывает, что ограничение скорости носителей в канале проявляется заметно уже при $V_{gst} \approx V_L$. Это означает, что при типичных значениях $V_{gst} = 2\text{ В}$ и $E_s = 20\text{ кВ/см}$ длина канала, при которой этот эффект следует учитывать, составляет $L = V_L / E_s \approx V_{gst} / E_s = 1\text{ мкм}$. В p -канальных МДПТ эффект проявляется при $L \approx 2\text{ мкм}$ вследствие меньшей подвижности дырок и большего значения E_s .

Естественно, что ограничение скорости носителей в канале влияет на предельную частоту МДПТ, определенную в (III.3.38):

$$\omega_T = \frac{g_s}{C_{gs} + C_{gd}} \Big|_{V_{ds} \geq V_{dsS}}. \quad (\text{III.5.16})$$

При $V_{gst} = V_{gs} - V_t \gg V_L$ с учетом (III.5.86), (III.5.136) и (III.5.14) получаем

$$g_s \approx g_{\max} = Zv_s C_{sG}; \quad (\text{III.5.17})$$

$$V_{dsS} \approx V_L \ll V_{gst};$$

$$V_{gdt} = V_{gst} - V_{dsS} \approx V_{gst}$$

и согласно (III.3.37a, б)

$$C_{gs} + C_{ds} \approx C_G = ZLC_{sG}. \quad (\text{III.5.18})$$

Подстановка (III.5.17) и (III.5.18) в (III.5.16) дает $\omega_T = g_{\max} / ZLC_{sG} = v_s / L$. Таким образом, при $V_{gs} - V_t \gg V_L$ максимально достижимая предельная частота МДПТ составляет

$$\omega_{T\max} = v_s / L = 1 / T_{C\min}, \quad (\text{III.5.19})$$

$$\text{где} \quad T_{C\min} = L / v_s \quad (\text{III.5.20})$$

— время пролета носителей заряда через канал с предельной скоростью v_s .

Сравнение (III.5.19) и (III.3.40) показывает, что с увеличением напряжения V_{gs} предельная частота МДПТ достигает максимального значения $\omega_{T\max} = v_s / L$, не зависящего от напряжения затвор—исток.

При этом предельная частота слабее возрастает с уменьшением длины канала ($\omega_T \sim 1/L$ вместо $\omega_T \sim 1/L^2$). Физическая причина этой закономерности состоит в том, что вследствие ограничения скорости носителей в канале уменьшение длины канала снижает емкость затвор—канал, но не повышает крутизну ВАХ МДПТ.

III.5.7. Модуляция длины канала

С насыщением дрейфовой скорости носителей заряда в канале связан еще один эффект короткого канала — *модуляция длины канала напряжением сток—исток*. Этот эффект поясняется рис. III.5.6, где схематически представлена «форма» канала (напомним, что на рисунках изображается не геометрическая форма канала, а поверхностная концентрация электронов) и распределение продольного электрического поля в канале.



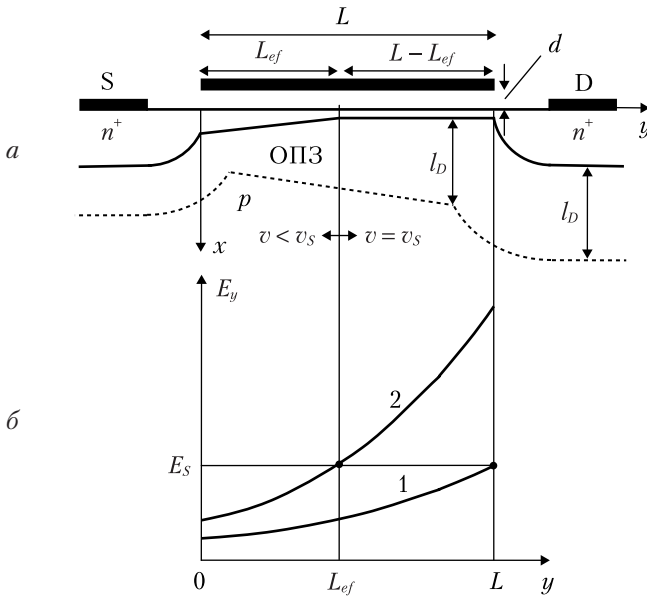


Рис. III.5.6. Эффективная длина канала (а) и распределение продольного электрического поля в канале (б):

$$\begin{aligned} 1 - V_{ds} &= V_{dss}, \\ 2 - V_{ds} &> V_{dss} \end{aligned}$$

В *крутой* области ВАХ ($V_{ds} \leq V_{dsS}$) напряженность продольного поля во всем канале $E_y < E_s$ (кривая 1 на рис. III.5.6, б), подвижность носителей постоянна и ВАХ описывается уравнениями (III.3.17) или (III.4.10). На *границе* крутой и пологой областей ($V_{ds} = V_{dsS}$) $E_y = E_s$ только на границе со стоком ($y = L$). В *пологой* области ($V_{ds} > V_{dsS}$) продольное поле достигает значения E_s в некоторой плоскости $y = L_{ef} < L$ (кривая 2 на рис. III.5.6, б). При этом транзистор можно разделить на две области: «виртуальный транзистор» с *эффективной длиной канала* L_{ef} , занимающий участок $0 \leq y \leq L_{ef}$, и область насыщения дрейфовой скорости носителей, занимающая участок $L_{ef} < y \leq L$ (рис. III.5.6, а). В *крутой* области ВАХ ($V_{ds} \leq V_{dsS}$) эффективная длина канала равна истинной: $L_{ef} = L$. В *пологой* области ($V_{ds} > V_{dsS}$) эффективная длина канала является убывающей функцией напряжения сток–исток: $L_{ef}(V_{ds})$ (рис. III.5.7).

Заметим, что «виртуальный транзистор» всегда работает на границе крутой и пологой областей, так как напряженность поля на границе с его «виртуальным стоком», расположенным

в плоскости $y = L_{ef}$, составляет $E_y(L_{ef}) = E_S$. Поэтому значение тока стока может быть найдено из уравнения (III.5.11а) как ток насыщения «виртуального транзистора» с заменами

$$L \rightarrow L_{ef}; \quad \beta \rightarrow \beta_{ef} = \beta L / L_{ef}; \quad V_{ds} \rightarrow V_{ds\,ef}; \quad V_L \rightarrow V_L L_{ef} / L:$$

$$I_D = \beta V_L \left[\sqrt{V_{gst}^2 + (V_L L_{ef} / L)^2} - V_L L_{ef} / L \right]. \quad (\text{III.5.21})$$

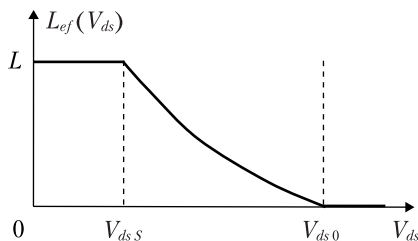


Рис. III.5.7. Зависимость эффективной длины канала от напряжения сток–исток

Напряжение насыщения $V_{ds\,ef}$ «виртуального транзистора» определяется уравнением (III.5.8а) с теми же заменами

$$V_{ds\,ef} = V_{gst} + V_L L_{ef} / L - \sqrt{V_{gst}^2 + (V_L L_{ef} / L)^2}. \quad (\text{III.5.22})$$

Зависимость $L_{ef}(V_{ds})$ можно найти путем приближенно-го решения двухмерного уравнения Пуассона в плоскости $x = 0$ области насыщения, занимающей участок $L_{ef} < y \leq L$ (рис. III.5.6, а), с граничными условиями

$$\begin{cases} V(L) - V(L_{ef}) = V_{ds} - V_{ds\,ef}, \\ \frac{dV}{dy}(L_{ef}) = -E_S. \end{cases}$$

Приближенное решение этого уравнения получено в приложении (в конце раздела III.5) (П.9):

$$V_{ds} - V_{ds\,ef} = V_L (a / L) \cdot \text{sh}[(L - L_{ef}) / a], \quad (\text{III.5.23})$$

$$\text{где} \quad a = \sqrt{d_{ef} l_D}; \quad (\text{III.5.24})$$

$d_{ef} = d\epsilon / \epsilon_d$ — эффективная (приведенная к значению диэлектрической проницаемости полупроводника) толщина диэлектрика; l_D — ширина ОПЗ вблизи стока (III.5.26).

При некотором напряжении сток—исток V_{ds0} эффективная длина канала становится равной нулю (см. рис. III.5.7) и во всем канале носители заряда движутся с предельной дрейфовой скоростью v_s . Напряжение V_{ds0} можно найти из уравнения (III.5.23), полагая $L_{ef} = 0$ и $V_{ds\,sef} = 0$,

$$V_{ds0} = V_L(a/L) \cdot \text{sh}(L/a). \quad (\text{III.5.25})$$

Из (III.5.21) с учетом (III.5.14) получаем ток стока I_{D0} и крутизну $g_0 = \partial I_{D0} / \partial V_{gs}$, соответствующие этому напряжению,

$$I_{D0}(V_{gs}) = \beta V_L V_{gst} = g_{\max} V_{gst}; \quad (\text{III.5.26a})$$

$$g_0 = g_{\max}. \quad (\text{III.5.26b})$$

Крутизна на границе с крутой областью ($V_{ds} = V_{ds\,s}$) определяется соотношением (III.5.13a), а при $V_{ds} = V_{ds0}$ ($V_{ds\,sef} = 0$, $L_{ef} = 0$) согласно (III.5.26b) достигает предельного значения $g_{\max}$. Выходная проводимость в пологой области ВАХ определяется уравнениями (III.5.21) и (III.5.23):

$$G = \frac{\partial I_D}{\partial V_{ds}} = \frac{\partial I_D}{\partial L_{ef}} \cdot \frac{\partial L_{ef}}{\partial V_{ds}};$$

$$\frac{\partial I_D}{\partial L_{ef}} = -\frac{\beta V_L^2}{L} \left[1 - \frac{V_L L_{ef} / L}{\sqrt{V_{gst}^2 + (V_L L_{ef} / L)^2}} \right];$$

$$\frac{\partial V_{ds}}{\partial L_{ef}} = -E_s \cdot \text{ch}[(L - L_{ef}) / a].$$

На границе с крутой областью ($V_{ds} = V_{ds\,s} = V_{ds\,sef}$, $L_{ef} = L$)

$$G_s(V_{gs}) = g_{\max} (1 - V_L / \sqrt{V_{gst}^2 + V_L^2}). \quad (\text{III.5.27})$$

При $V_{ds} = V_{ds0}$ ($V_{ds\,sef} = 0$, $L_{ef} = 0$):

$$G_0 = g_{\max} / \text{ch}(L/a). \quad (\text{III.5.28})$$

Параметры $V_L = E_s L$ и $g_{\max} V_L$ удобно использовать в качестве естественных единиц измерения напряжения и тока. При этом безразмерные значения напряжений и токов составляют

$$v_{gst} = V_{gst} / V_L; \quad (\text{III.5.29a})$$

$$v_{ds} = V_{ds} / V_L; \quad (\text{III.5.29b})$$

$$v_{ds0} = V_{ds0} / V_L = \text{sh}(L/a) / (L/a); \quad (\text{III.5.29в})$$

$$i_D = I_D / g_{\max} V_L. \quad (\text{III.5.29г})$$



III.5.8. Пологая область ВАХ МДПТ и коэффициент усиления

Уравнение пологой области ВАХ МДПТ находится путем исключения переменных $V_{ds\ s\ ef}$ и L_{ef} из уравнений (III.5.21)–(III.5.23). Чтобы обойти необходимость решения трансцендентного уравнения (III.5.23), представим приближенно выходную ВАХ уравнением

$$I_D = I_{DS} + G_S(V_{ds} - V_{dsS})[1 + \phi(V_{ds} - V_{dsS})/V_L + \chi(V_{ds} - V_{dsS})^2/V_L^2]. \quad (\text{III.5.30})$$

На границе с крутой областью ($V_{ds} = V_{dsS}$) аппроксимация (III.5.30) дает правильные значения тока ($I_D = I_{DS}$) и выходной проводимости ($G = G_S$), а коэффициенты ϕ и χ определяются из уравнений

$$\begin{cases} I_{D0} = I_{DS} + G_S(V_{ds0} - V_{dsS})[1 + \phi(V_{ds0} - V_{dsS})/V_L + \chi(V_{ds0} - V_{dsS})^2/V_L^2], \\ G_0 = G_S[1 + 2\phi(V_{ds0} - V_{dsS})/V_L + 3\chi(V_{ds0} - V_{dsS})^2/V_L^2], \end{cases}$$

обеспечивающих правильные значения тока ($I_D = I_{D0}$) и выходной проводимости ($G = G_0$) при $V_{ds} = V_{ds0}$.

В безразмерных обозначениях (III.5.29) с учетом (III.5.8а), (III.5.11а), (III.5.14) и (III.5.25)–(III.5.28) ВАХ МДПТ для произвольного режима работы имеет вид

$$i_D = \begin{cases} 0, & v_{gst} < 0, \\ v_{ds}(v_{gst} - v_{ds}/2), & v_{gst} \geq 0, v_{ds} \leq v_{dsS}, \\ i_{DS} + (G_S/g_{\max})(v_{ds} - v_{dsS})[1 + \\ + \phi(v_{ds} - v_{dsS}) + \chi(v_{ds} - v_{dsS})^2], & v_{gst} \geq 0, v_{ds0} \geq v_{ds} > v_{dsS}. \end{cases} \quad (\text{III.5.31})$$

На рис. III.5.8 представлены зависимости $i_D(v_{ds})$ при $v_{gst} = V_{gst}/V_L = 1$ для случаев $L/a = 0,3$ (кривая 1), $L/a = 1$ (кривая 2) и $L/a = 3$ (кривая 3). Для сравнения на рисунке приведена ВАХ идеализированного МДПТ (III.3.17а) (штриховая линия 4).

Как очевидно из рис. III.5.8, ограничение скорости носителей в канале и эффект модуляции длины канала приводят к тому, что в пологой области ВАХ выходная проводимость не равна нулю и собственный коэффициент усиления напряжения K_V оказывается конечным. Соотношения (III.5.13а) и (III.5.26)–(III.5.28) позволяют оценить его значения K_{VS} на границе с крутой областью ($V_{DS} = V_{DSs}$)



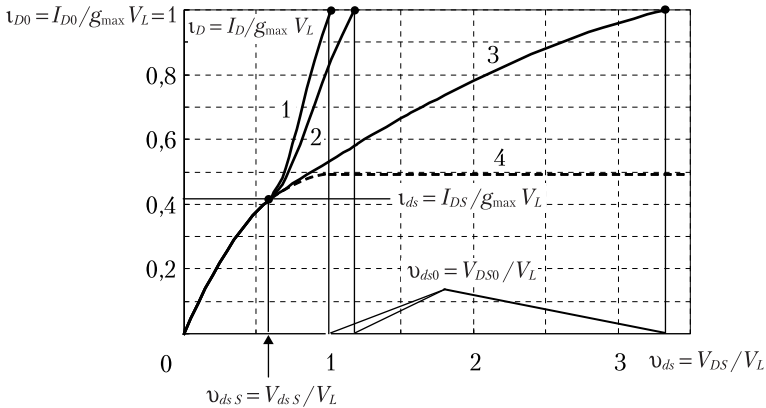


Рис. III.5.8. Аппроксимация (III.5.31) выходной ВАХ в безразмерных обозначениях (III.5.29) для фиксированного напряжения $v_{gst} = (V_{gs} - V_t)/V_L = 1$: 1 — $L/a = 0,3$; 2 — $L/a = 1$; 3 — $L/a = 3$. Штриховая линия 4 — ВАХ идеализированного МДПТ (III.3.17а)

и K_{V0} при напряжении $V_{ds} = V_{ds0}$. В безразмерных обозначениях (III.5.29)

$$K_{VS}(v_{gst}) = \frac{g_s}{G_s} = \frac{v_{gst}}{(2 + v_{gst}^2) \sqrt{1 + v_{gst}^2} - 2}; \quad (\text{III.5.32a})$$

$$K_{V0}(L/a) = g_{\max} / G_0 = \text{ch}(L/a). \quad (\text{III.5.32б})$$

Зависимости $K_{VS}(v_{gst})$ и $K_{V0}(L/a)$ приведены на рис. III.5.9, а. Из рисунка очевидно, что на границе с крутой областью приемлемое значения коэффициента усиления $K_{VS} > 10$ достигается только при $v_{gst} = (V_{gs} - V_t)/V_L < 0,9$. Для кремниевого n -канального транзистора с пороговым напряжением $V_t = 1$ В и длиной канала $L = 0,3$ мкм ($V_L \approx 0,6$ В) это соответствует напряжению затвор—исток $V_{gs} < 1,5$ В. При $V_{ds} = V_{ds0}$ приемлемое значения коэффициента усиления $K_{V0} > 10$ достигается только при $L/a > 3$, т.е. $a = \sqrt{d_{ef} I_D} < 0,1$ мкм. Это означает, что для типичного значения $d = 0,03$ мкм ($d_{ef} = d\epsilon / \epsilon_d \approx 0,1$ мкм) ширина ОПЗ l_d не должна превышать 0,1 мкм, и, следовательно, концентрация примеси в подложке должна быть не менее $(1-2) \cdot 10^{17} \text{ см}^{-3}$. На рис. III.5.9, а приведены также зависимости $v_{ds S} = V_{ds S}(v_{gst})/V_L$ (III.5.8а) и $v_{ds0} = V_{ds0}(L/a)/V_L$ (III.5.25).

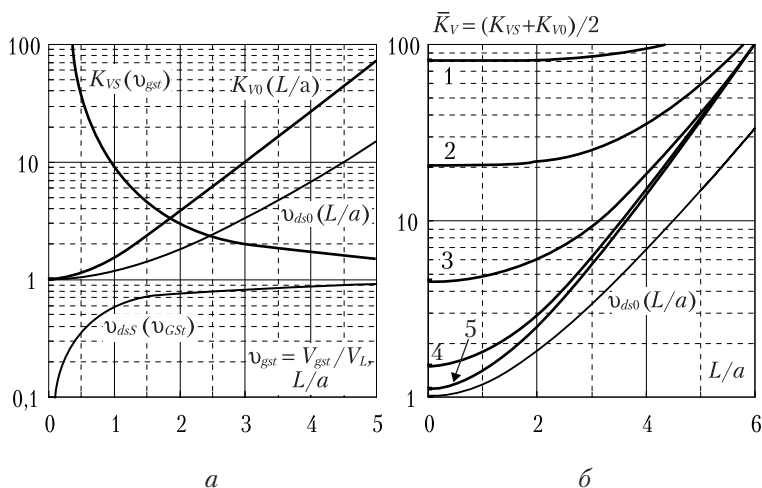


Рис. III.5.9. Зависимости $K_{VS}(v_{gst})$ по формуле (III.5.32а), $K_{V0}(L/a)$ по формуле (III.5.32б) и $v_{ds0}(L/a)$ по формуле (III.5.25)[а]; зависимости $\bar{K}_V(v_{gst}, L/a)$ по формуле (III.5.32в) и $v_{gst}(L/a)$ по формуле (III.5.25)[б]:

1 — $v_{gst} = 0,3$; $v_{ds} = 0,26$; 2 — $v_{gst} = 0,5$; $v_{ds} = 0,38$; 3 — $v_{gst} = 1$; $v_{ds} = 0,59$; 4 — $v_{gst} = 3$; $v_{ds} = 0,84$; 5 — $v_{gst} = 10$; $v_{ds} = 0,95$

Для принятого выше значения $V_L \approx 0,6$ В при $L/a = 3$ напряжение V_{ds0} составляет около 2 В.

В качестве среднего значения коэффициента усиления напряжения $\bar{K}_V$ в пологой области ВАХ можно принять

$$\bar{K}_V(v_{gst}, L/a) = [K_{VS}(v_{gst}) + K_{V0}(L/a)] / 2. \quad (\text{III.5.32в})$$

На рис. III.5.9, б приведены зависимости $\bar{K}_V(v_{gst}, L/a)$ (III.5.32в) для пяти значений параметра v_{gst} от 0,3 до 10 и $v_{ds0}(L/a)$ (III.5.25). Для каждого значения параметра v_{gst} указаны значения напряжений v_{ds} . На рис. III.5.10 в зависимости от параметра L/a приведено максимальное напряжение затвор—исток, при котором средний коэффициент усиления в пологой области ВАХ $\bar{K}_V = 10$. Как очевидно из рисунка, при $L/a < 3,6$ значение $\bar{K}_V = 10$ может быть обеспечено только при $V_{gst} \gg V_L$.

Полученные результаты показывают, что снижение коэффициента усиления $\bar{K}_V$ при уменьшении длины канала вследствие эффекта модуляции длины канала создает серьезные ограничения при создании высококачественных короткока-

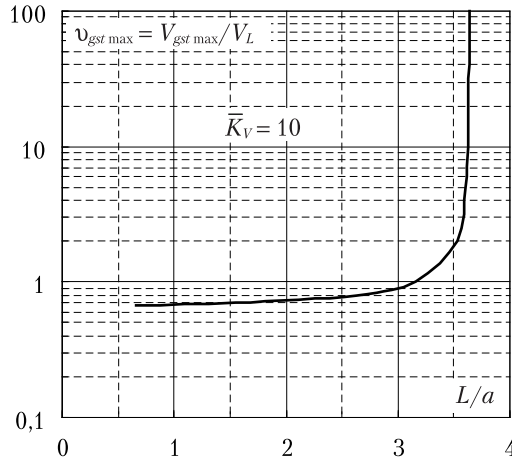


Рис. III.5.10. Максимальное напряжение затвор—исток, при котором средний коэффициент усиления в пологий области ВАХ $\bar{K}_v = 10$, в зависимости от параметра L/a

нальных транзисторов. Единственный способ ослабить этот эффект состоит в *увеличении отношения $L/a = L/\sqrt{d_{ef}l_D}$* . Это означает, что в короткоканальных транзисторах необходимо *уменьшать эффективную толщину диэлектрика $d_{ef} = d\epsilon/\epsilon_d$ и ширину ОПЗ l_D* . Уменьшение ширины ОПЗ требует *повышения степени легирования подложки*, что приводит к повышению емкостей *p-n* переходов подложка—сток и подложка—исток и снижает быстродействие транзистора. Возможности добиться в этом отношении некоторого компромисса обсуждаются в разделе III.8.

Физическая причина влияния параметра L/a на коэффициент усиления $K_v = g/G$ поясняется рис. III.5.11, где показаны силовые линии избыточного электрического поля, возникающего при повышении напряжения сток—исток на величину ΔV_{ds} . В случае $L/a \gg 1$ (рис. III.5.11, *a*) эти силовые линии замыкаются на электродах затвора и подложки, не оказывая существенного влияния на ток стока (выходная проводимость G остается низкой). При $L/a \approx 1$ значительная часть силовых линий замыкается на избыточных электронах, вытягиваемых полем из электрода истока, что ведет к возрастанию тока стока. Таким образом, выходная проводимость G заметно возрастает, что ведет за собой снижение коэффициента усиления.

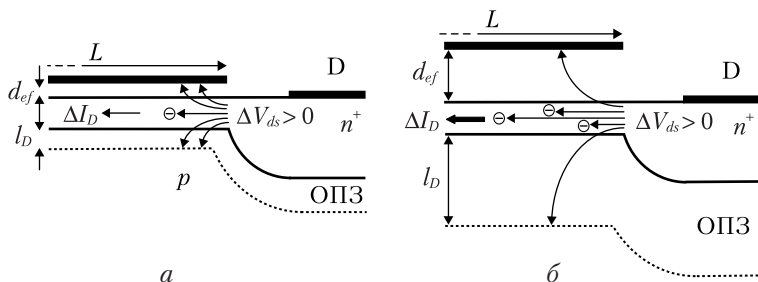


Рис. III.5.11. Влияние параметра $L/a = L/\sqrt{d_{ef}l_D}$ на коэффициент усиления МДПТ по напряжению:
 $a - L/a \gg 1$; $б - L/a \approx 1$

III.5.9. Емкости затвор—исток и затвор—сток

В пологой области ВАХ эффект модуляции длины канала оказывает влияние и на емкости затвор—исток, затвор—сток. В идеализированной модели зависимости этих емкостей от напряжений $C_{gs}, C_{gd}(V_{gs}, V_{gd})$ описываются соотношениями (III.3.37а, б). В параграфе III.3.3 отмечалось, что в пологой области ВАХ емкость затвор—исток составляет $\frac{2}{3}$ емкости диэлектрика ($C_{gs} = 2C_G/3$), а емкость затвор—сток $C_{gd} = 0$. Последнее обстоятельство объясняется тем, что в пологой области ВАХ канал перекрыт на границе со стоком и напряжение V_{gd} не влияет на распределение и величину заряда электронов в канале Q_n и компенсирующего его заряда на затворе $Q_G = -Q_n$. Эффект модуляции длины канала приводит к перераспределению заряда Q_G при изменении напряжения V_{gd} , поэтому в пологой области ВАХ емкость C_{gd} остается отличной от нуля. Ее значение можно оценить, представив в виде

$$C_{gd} = \frac{\partial Q_G}{\partial V_{gd}} = \frac{\partial Q_G}{\partial V_{ds}} = \frac{\partial Q_G}{\partial I_D} \cdot \frac{\partial I_D}{\partial V_{ds}}. \quad (\text{III.5.33})$$

Частная производная $\partial Q_G / \partial I_D = \partial |Q_n| / \partial I_D$ имеет смысл времени пролета носителей заряда через участок канала длиной $L - L_{ef}$ (см. рис. III.5.6, а). Поскольку на этом участке их дрейфовая скорость имеет постоянное значение v_s , $\partial Q_G / \partial I_D = (L - L_{ef}) / v_s$. Частная производная $\partial I_D / \partial V_{ds}$ представляет собой выходную проводимость G . Таким образом,

$$C_{gd} = G(L - L_{ef}) / v_s. \quad (\text{III.5.34})$$

Значения $L - L_{ef}$ и G могут быть получены из уравнений (III.5.21)–(III.5.23).

Емкость C_{gs} можно представить в виде

$$C_{gs} = \partial Q_G / \partial V_{gs} = G_{gs1} + C_{gs2},$$

где $C_{gs1,2} = \partial |Q_{n1,2}| / \partial V_{gs}$; (III.5.35a)

Q_{n1} — заряд электронов в канале «виртуального транзистора» длиной L_{ef} ; $Q_{n2} = I_D (L - L_{ef}) / v_s$ — заряд электронов на участке насыщения их скорости длиной $L - L_{ef}$. Составляющая C_{gs1} вычисляется из соотношения (III.3.37a, б) с заменой $V_{gd} = V_{gs} - V_{ds} \rightarrow V_{gd\,ef} = V_{gs} - V_{ds\,ef}$, а составляющую C_{gs2} можно представить в виде

$$C_{gs2} = \frac{\partial |Q_{n2}|}{\partial V_{gs}} = \frac{\partial |Q_{n2}|}{\partial I_D} \cdot \frac{\partial I_D}{\partial V_{gs}} = g(L - L_{ef}) / v_s. \quad (III.5.35б)$$

Процедура вычисления емкостей C_{gd} и C_{gs} достаточно громоздка. Отметим только, что при напряжении $V_{ds} = V_{ds0}$, когда $L_{ef} = L$, $g = g_{\max}$ и $G = G_0 = g_{\max} / K_{V0}$, $C_{gs1} = 0$ и емкости $C_{gs, gd}$ составляют:

$$C_{gs} = g_{\max} L / v_s, \quad (III.5.36)$$

$$C_{gd} = g_{\max} L / v_s K_{V0}. \quad (III.5.37)$$

При этом предельная частота МДПТ согласно (III.3.38) выражается соотношением

$$\omega_{T\max} = v_s (1 + 1 / K_{V0}) / L \approx v_s / L. \quad (III.5.38)$$

Таким образом, *максимально достижимая предельная частота определяется временем пролета носителей заряда через канал с предельной скоростью дрейфа v_s .*

III.5.10. Особенности ВАХ короткоканального МДПТ

ВАХ n -канального МДПТ с коротким каналом приведены на рис. III.5.12, где для сравнения показаны характеристики, соответствующие идеализированной модели (см. раздел III.3). На рисунке отражены следующие эффекты короткого канала: 1) изменение порогового напряжения и его зависимость от напряжения сток—исток (см. параграф III.5.2); 2) ограничение скорости носителей в канале (см. параграф III.5.5); 3) модуляция длины канала (см. параграф III.5.7); 4) эффект смыкания канала (см. параграф III.5.3).

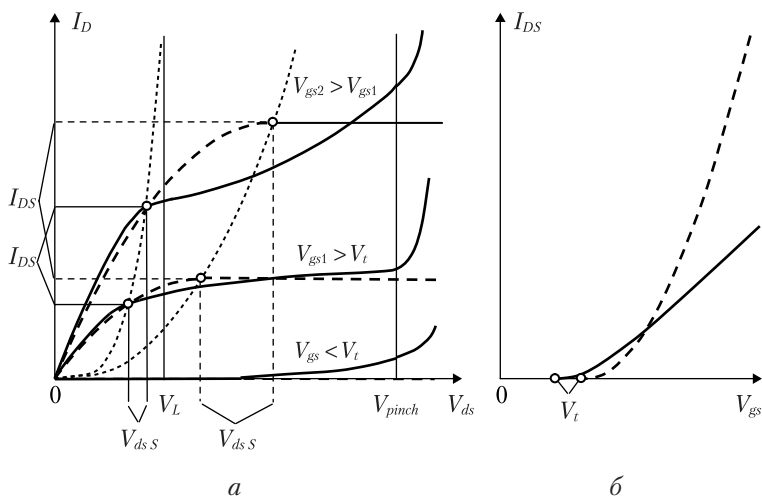


Рис. III.5.12. ВАХ n -канального МДПТ с коротким каналом: а — выходные; б — $I_{DS}(V_{gs})$. Штриховые линии — идеализированная модель; пунктир — границы крутой и пологой областей

При малой длине канала пороговое напряжение V_t сдвигается в сторону меньших напряжений затвор—исток, что отражено на рис. III.5.12, б. На выходных характеристиках это приводит к некоторому увеличению тока стока при малом напряжении сток—исток (рис. III.5.12, а). Кроме того, пороговое напряжение снижается при увеличении напряжения сток—исток V_{ds} (см. параграф III.5.2). Поэтому для напряжения затвор—исток, несколько меньшего V_t , повышение V_{ds} может привести к отпиранию транзистора, что показано на рис. III.5.12, а.

Ограничение дрейфовой скорости электронов в канале приводит к снижению напряжения насыщения V_{dsS} , которое не превышает величины $V_L = E_s L$, и к соответствующему уменьшению тока насыщения I_{DS} (см. рис. III.5.12, а, б).

В полой области ВАХ ток стока возрастает с ростом напряжения сток—исток (см. рис. III.5.12, а), что является следствием эффекта модуляции эффективной длины канала. Если напряжение затвор—исток не слишком велико (V_{gs1} на рис. III.5.12, а), выходная ВАХ в полой области имеет выпуклый характер; при больших напряжениях затвор—исток (V_{gs2} на рис. III.5.12, а) ВАХ становится вогнутой.

Наконец, эффект смыкания канала приводит к резкому возрастанию тока стока при напряжении сток—исток, большем напряжения смыкания V_{pinch} (см. рис. III.5.12, а, параграф III.5.3).

III.5.11. Подпороговый ток

Эффекты короткого канала проявляются на подпороговых ВАХ двояким образом.

Во-первых, модуляция длины канала напряжением стока увеличивает в соотношении (III.4.18) масштабный ток $I_{D\text{dif}0}$, который, как и дрейфовый ток в режиме сильной инверсии, обратно пропорционален эффективной длине канала.

Другой механизм состоит в изменении эффективного значения удельной емкости ОПЗ C_{sb} (III.4.14) вследствие неоднородных эффектов. Каждая точка поверхности полупроводника под затвором как бы «чувствует» близко расположенные n^+ -области стока и истока, что приводит к увеличению эффективной емкости C_{sb} с уменьшением длины канала. Этот эффект аналогичен увеличению емкости плоского конденсатора за счет краевых эффектов, когда размеры обкладок становятся сравнимыми с толщиной диэлектрика. Он проявляется, когда длина канала становится сравнимой с толщиной ОПЗ и, следовательно, снижается при повышении концентрации примеси в подложке.

Увеличение эффективной емкости C_{sb} приводит к уменьшению коэффициента передачи емкостного делителя напряжения $C_{sg} - C_{sb}$ (см. рис. III.4.2, *a*), т.е. к увеличению коэффициента $m = 1/k_c$ в ВАХ подпорогового тока (III.4.18). При снижении напряжения V_{gs} его изменения слабее влияют на потенциал поверхности полупроводника и, следовательно, на поверхностную концентрацию электронов. В результате подпороговый ток слабее уменьшается с уменьшением V_{gs} и наклон ВАХ в полулогарифмическом масштабе становится меньше. Аналитический расчет ВАХ подпорогового тока в короткоканальных транзисторах провести весьма сложно; некоторые экспериментальные результаты приведены в работе [10, § 8.4.1].

III.5.12. Критерий короткого канала

Как было показано, все рассмотренные выше эффекты короткого канала являются нежелательными. Степень их проявления зависит от многих факторов. Мы показали, что наиболее важными параметрами структуры МДПТ в этом смысле являются эффективная толщина диэлектрика d_{ef} , толщина n^+ -областей стока и истока x_j , а также ширина ОПЗ под стоком l_D и истоком l_S , значения которой зависят от сте-



пени легирования подложки и определяются соотношениями (III.5.2а, б).

В результате большого количества экспериментальных исследований, выполненных на образцах МДПТ с различными значениями параметров структуры, предложено [11] обобщить критерий короткого канала простым эмпирическим соотношением

$$L_{\min} = 0,4[x_j d_{ef}(l_s + l_D)^2]^{1/3}, \quad (\text{III.5.39})$$

где параметры L , x_j , l_s и l_D измеряются в микрометрах, а d_{ef} — в ангстремах. По данным [11] при $L > L_{\min}$ зависимость $I_{ds}(L)$ отличается от обратно пропорциональной не более чем на 10%, а также исключается существенная зависимость порогового напряжения от длины канала. Заметим, однако, что чрезмерно доверять критерию (III.5.39), по-видимому, не следует, так как он не содержит информации об эффектах ограничения скорости носителей в канале и модуляции длины канала и, следовательно, о пологой области ВАХ и собственном коэффициенте усиления напряжения K_V . В частности, критерий (III.5.39) не содержит напряжения V_{gst} , которое, как показано в параграфе III.5.6, определяет степень проявления эффекта ограничения дрейфовой скорости носителей в канале.

Основные выводы

1. Пороговое напряжение МДП-транзисторов при малой длине канала изменяется вследствие уменьшения по абсолютной величине суммарного заряда ионов подложки в ОПЗ. Для n -канальных транзисторов уменьшение длины канала приводит к уменьшению порогового напряжения, для p -канальных — к увеличению. К тому же эффекту приводит увеличение напряжения сток—исток. Уменьшение ширины канала приводит к обратному эффекту.

2. Сближение контактных областей стока и истока в коротко-канальных МДПТ приводит к смыканию ОПЗ стока и истока под затвором. Этот эффект проявляется на выходных характеристиках подобно пробую.

3. Разогрев носителей заряда в сильных электрических полях делает возможным их проникновение в подзатворный диэлектрик. Этот эффект увеличивает пороговое напряжение для n -канальных транзисторов и уменьшает — для p -канальных, а также ухудшает стабильность характеристик транзисторов.

4. Ограничение предельной дрейфовой скорости носителей заряда в канале учитывается введением дополнительного параметра



ВАХ МДПТ — напряжения V_L , которое равно произведению длины канала на критическую напряженность электрического поля: $V_L = E_s L$. Ограничение дрейфовой скорости носителей снижает напряжение насыщения ВАХ МДПТ и ток насыщения. Степень проявления этого эффекта определяется отношением $v_{gst} = V_{gst} / V_L$. При $v_{gst} \gg 1$ крутизна ВАХ МДПТ достигает предельного значения $g_{max} = Z v_s C_{sg}$, не зависящего ни от длины канала, ни от напряжения затвор—исток, а предельная частота определяется временем пролета носителей заряда через канал с предельной скоростью дрейфа v_s .

5. Ограничение предельной дрейфовой скорости носителей заряда в канале является причиной модуляции длины канала напряжением сток—исток. Этот эффект приводит к нарастанию тока стока при увеличении напряжения сток—исток в пологой области ВАХ и ограничивает собственный коэффициент усиления МДПТ по напряжению. Для повышения коэффициента усиления по напряжению в короткоканальных транзисторах необходимо уменьшать эффективную толщину диэлектрика и ширину ОПЗ.

6. При уменьшении длины канала подпороговый ток слабее уменьшается при запираании транзистора по затвору. Причиной этого является увеличение эффективной удельной емкости ОПЗ вследствие краевых эффектов.

Приложение к параграфу III.5.7

Определение зависимости эффективной длины канала от напряжения сток—исток

Для определения зависимости $L_{ef}(V_{ds})$ используем уравнение Пуассона для плоскости поверхности полупроводника ($x = 0$) на участке $L_{ef} \leq y \leq L$, где напряженность продольного электрического поля $E_y > E_s$ и скорость электронов в канале достигает максимального значения v_s (рис. П.1):

$$\frac{\partial E_x}{\partial x}(y) + \frac{\partial E_y}{\partial y}(y) = \frac{\rho(0, y)}{\epsilon \epsilon_0}. \quad (\text{П.1})$$

При решении уравнения (П.1) примем следующие допущения:

- а) толщина ОПЗ l не зависит от координаты y ;
- б) эффективная толщина канала δ не зависит от координаты y ;
- в) в плоскости $y = L_{ef}$: $\partial E_y / \partial y \ll \partial E_x / \partial x$.

Плотность объемного заряда $\rho(0, y)$ в (П.1) обусловлена ионами примеси в ОПЗ и электронами в канале

$$\rho(0, y) = \rho_N - en(0, y).$$



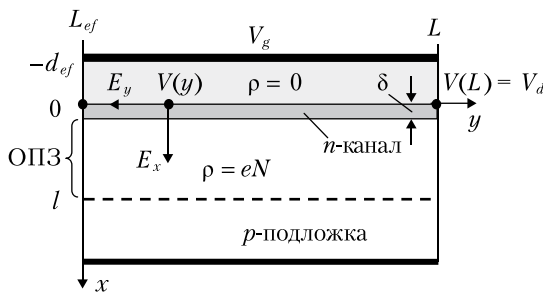


Рис. П.1. Участок насыщения скорости в канале

На участке насыщения скорости ток в канале $I = ev_s n(0, y) Z \delta$ не зависит от координаты y . Поэтому при допущении б) поверхностная концентрация электронов в канале $n(0, y)$ и плотность объемного заряда $\rho(0, y)$ в (П.1) также не зависят от координаты y : $\rho(0, y) = \rho(0)$. Таким образом, при допущении в) в плоскости $y = L_{ef}$ уравнение (П.1) имеет вид

$$\frac{\partial E_x}{\partial x}(L_{ef}) = \frac{\rho(0)}{\epsilon \epsilon_0}. \quad (\text{П.2})$$

Вычитая (П.2) из (П.1), получаем

$$\frac{\partial E_x}{\partial x}(y) - \frac{\partial E_x}{\partial x}(L_{ef}) + \frac{\partial E_y}{\partial y}(y) = \frac{\rho(0)}{\epsilon \epsilon_0}. \quad (\text{П.3})$$

На нижней границе ОПЗ ($x = l$, рис. П.1) $E_x(l, y) = 0$, а на поверхности канала ($x = 0$) поле E_x определяется разностью потенциалов затвор-канал: $E_x(0, y) = [V_g - V(y)] / d_{ef}$. Поэтому среднее по x значение производной $\partial E_x / \partial x$ составляет

$$\overline{\frac{\partial E_x}{\partial x}}(y) = \frac{E_x(0, y)}{l} = \frac{V_g - V(y)}{d_{ef} l} = \frac{V_g - V(y)}{a^2},$$

где

$$a = \sqrt{d_{ef} l}. \quad (\text{П.4})$$

Полагая в (П.3) $\partial E_x / \partial x \approx \overline{\partial E_x / \partial x}$ и учитывая, что $E_y = -\partial V / \partial y$, получаем

$$d^2 V(y) / dy^2 - [V(y) - V(L_{ef})] / a^2 = 0. \quad (\text{П.5})$$

Уравнение (П.5) должно быть решено с граничными условиями

$$dV / dy(L_{ef}) = -E_s; \quad (\text{П.6a})$$

$$V(L) - V(L_{ef}) = V_{ds} - V_{ds \text{эф}}, \quad (\text{П.6б})$$

где $V_{ds\,sef}$ — напряжение сток—исток насыщения для виртуального транзистора с эффективной длиной канала L_{ef} . Это напряжение определяется соотношением (III.5.8a) с заменами $L \rightarrow L_{ef}$, $V_L = E_S L \rightarrow V_{L_{ef}} = E_S L_{ef}$

$$V_{ds\,sef} = V_{gst} + V_{L_{ef}} - \sqrt{V_{gst}^2 + V_{L_{ef}}^2}. \quad (\text{П.7})$$

Решение уравнения (П.5) с граничными условиями (П.6a, б) имеет вид

$$V(y) - V(L_{ef}) = \frac{(V_{ds} - V_{ds\,sef}) \operatorname{ch} \frac{y - L_{ef}}{a} - E_S a \left(\operatorname{sh} \frac{L - y}{a} + \operatorname{sh} \frac{y - L_{ef}}{a} \right)}{\operatorname{ch} [(L - L_{ef}) / a] - 1}. \quad (\text{П.8})$$

На границе со стоком ($y = L$) из (П.8) получаем

$$\begin{aligned} V_{ds} - V_{ds\,sef} &= E_S a \cdot \operatorname{sh} [(L - L_{ef}) / a] = \\ &= V_L (a / L) \cdot \operatorname{sh} [(L - L_{ef}) / a]. \end{aligned} \quad (\text{П.9})$$

III.6. Моделирование МДП-транзисторов

III.6.1. Классификация моделей полупроводниковых приборов

Для детального анализа свойств полупроводниковых приборов, оптимального проектирования их структуры и симуляции их поведения в ИС применяются различные методы моделирования. Несколько упрощенная классификация моделей полупроводниковых приборов приведена в таблице III.6.1.

Наиболее достоверные *физические* модели основаны на физическом описании движения каждой частицы (носителя заряда) либо макрочастиц (меньшего количества частиц с увеличенными зарядами) и последующем расчете распределения их концентраций и электрического поля.

Достаточно точные результаты во многих случаях дают «квазигидродинамические» *физические* модели, основанные на решении уравнений переноса частиц, импульса, энергии. Эти уравнения являются дифференциальными уравнениями в частных производных и получаются путем интегрирования по всем импульсам кинетического уравнения Больцмана, умноженного на 1, импульс и квадрат импульса, соответственно. Полученные уравнения содержат кинетические коэффициенты (коэффициент диффузии, подвижность и т.п.), зависящие от состояния ансамбля частиц. При таком подходе



Таблица III.6.1

Классификация моделей полупроводниковых приборов

Класс модели	Основа модели	Математический аппарат	Форма представления	Разновидности
Физические	Уравнения движения частиц (метод Монте-Карло)	Обыкновенные дифференциальные уравнения, дифференциальные уравнения в частных производных	Математические уравнения	Полевые, температурные
	Уравнения переноса частиц, импульса, энергии		Математические уравнения	
	Уравнения ВАХ и ВФХ элементов прибора	Алгебраические уравнения		Для большого сигнала (нелинейные), для малого сигнала (линейные)
	Уравнения Кирхгофа	Обыкновенные дифференциальные уравнения	Эквивалентная схема	
Формализованные	Уравнения ВАХ и ВФХ элементов прибора	Обыкновенные дифференциальные уравнения		Для большого сигнала (нелинейные)
	Уравнения Кирхгофа	Обыкновенные дифференциальные уравнения	Эквивалентная схема	
Табличные	Таблицы значений ВАХ и ВФХ элементов прибора	Алгебраические уравнения	Эквивалентная схема	Для большого сигнала (нелинейные)
Формальные	Формальный четырехполюсник (многополюсник)	Обыкновенные дифференциальные уравнения	Матрица параметров	Для малого сигнала (линейные)



основная трудность состоит в «проблеме локальности» — выборе подходящего аргумента, от которого локально зависят кинетические коэффициенты. Строго говоря, такой единый аргумент отсутствует. Обычно в качестве этого аргумента выбирается напряженность неравновесного электрического поля, средняя энергия частиц, либо их температура (которая может быть различной для частиц различного рода).

Такие физические модели высокого уровня могут быть одно-, двух- или трехмерными. Анализ двух- или трехмерных объектов производится исключительно численными методами, причем требует даже для одного прибора значительных затрат вычислительных ресурсов. Применение этих моделей ограничивается в основном исследованием неизвестных свойств приборов и оптимизацией их параметров.

Физические модели, предназначенные для схемотехнических расчетов, представляются в виде эквивалентных схем с конечным числом сосредоточенных параметров. Электрические характеристики элементов эквивалентных схем описываются алгебраическими уравнениями, а взаимодействие элементов — уравнениями Кирхгофа (обыкновенными дифференциальными уравнениями). Такой подход используется в широко распространенных SPICE-подобных системах автоматизированного проектирования ИС, где в зависимости от требуемой точности применяются модели различного уровня. Модели высокого уровня могут содержать до 100 и более параметров; модели низкого уровня не обладают высокой точностью и могут использоваться даже для аналитических расчетов. Для анализа нелинейных цепей применяются «модели для большого сигнала», пригодные для описания произвольного режима работы приборов. Для линейных цепей могут использоваться линеаризованные или «малосигнальные» модели, в которых параметры элементов рассчитываются для заданного электрического режима и далее считаются постоянными.

Безусловным преимуществом таких физических моделей является возможность прогнозировать изменения параметров при совершенствовании технологических процессов и ужесточении проектных норм, а также возможность теоретического предсказания температурных зависимостей параметров. На практике, к сожалению, для описания температурных зависимостей часто приходится использовать подгоночные коэффициенты.

Значительная экономия вычислительных ресурсов достигается при использовании *формализованных* моделей, в кото-



рых характеристики элементов описываются не реальными физическими уравнениями, а некими формальными функциями, дающими сходные результаты. Серьезной проблемой при использовании таких моделей является верификация параметров моделей, а также их температурной зависимости. Эта задача требует проведения большого количества статистически достоверных экспериментальных исследований.

Некоторое распространение получили *табличные* модели, в которых ВАХ и/или ВФХ заданы в виде табличных массивов данных. Их недостатками являются использование больших объемов памяти и необходимость создания новых массивов данных при изменении температуры.

Последним является класс *формальных* моделей, предназначенных для анализа исключительно линейных электрических цепей. В таких моделях прибор описывается как линейный четырехполюсник (или многополюсник) неизвестной структуры. Токи и напряжения, определяющие электрическое состояние этого «черного ящика», связаны линейными дифференциальными уравнениями, коэффициенты которых (для данного режима работы) зависят только от частоты. Таким образом, модель описывается матрицей частотно-зависимых параметров. Главное преимущество такого подхода состоит в простоте измерения параметров.

Ниже мы рассмотрим только два типа моделей, пригодных для схемотехнического анализа и проектирования ИС: модели на основе эквивалентных схем (для большого и малого сигнала), а также формальные модели (для малого сигнала).

III.6.2. Особенности моделирования МДПТ

Модель прибора должна правильно воспроизводить статические ВАХ, а также учитывать факторы, определяющие его инерционность. Обычно это достигается введением в эквивалентную схему реактивных элементов. Принципиальным механизмом инерционности МДП-транзистора является перезаряд емкости C_G между затвором и каналом. Эта емкость не является паразитной, так как именно она определяет степень проявления эффекта поля, на котором основан принцип действия прибора. Так, например, крутизна ВАХ, определяющая усилительные свойства транзистора, согласно (III.3.27), (III.5.13) и (III.5.14) пропорциональна удельной емкости затвора C_{sG} .



Помимо этого, работа МДП-транзистора неизбежно связана с перезарядом целого ряда паразитных емкостей. Их минимизация не ухудшает усилительных свойств прибора и является одной из главных задач при его конструировании.

Для различных приборов используется различный подход к синтезу эквивалентной схемы. Эквивалентная схема МДПТ отражает особенности его структуры. На рис. III.6.1 схематично представлена структура МДПТ и выделена область «внутреннего» транзистора, ограниченная подзатворной областью канала и находящейся под ней областью ОПЗ и подложки. «Внутреннему» транзистору приписаны «внутренние» электроды затвора g , истока s , стока d и подложки b .

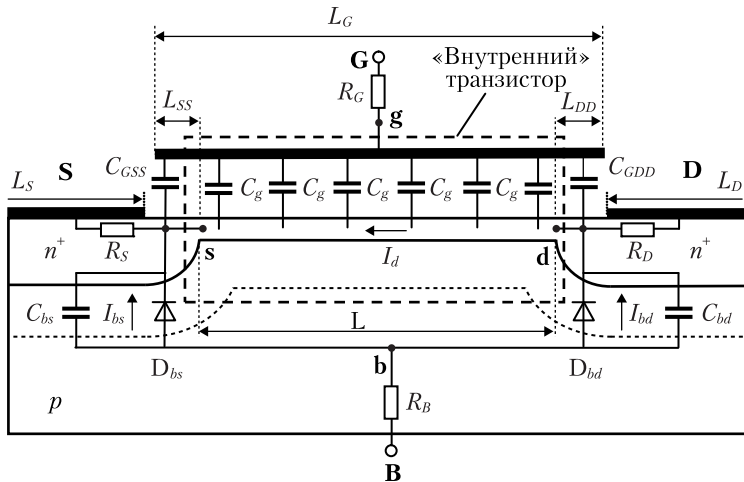


Рис. III.6.1. «Внутренний» транзистор и паразитные элементы МДПТ

С «внешними» элементами конструкции связаны паразитные реактивные и резистивные элементы эквивалентной схемы. Емкости C_{bs} и C_{bd} представляют собой барьерные емкости p - n переходов исток—подложка, сток—подложка и зависят от соответствующих напряжений, а через диоды D_{bs} и D_{bd} протекают токи этих переходов (обычно — обратные токи). Длина затвора МДПТ L_G несколько превышает длину канала L . Выступающим за пределы «внутреннего» транзистора областям затвора длиной L_{DD} и L_{SS} соответствуют емкости C_{GDD} и C_{GSS} между этими областями и n^+ -контактными областями стока и истока (емкости перекрытия). Рези-

стивные элементы R_D и R_S соответствуют сопротивлениям n^+ -областей стока и истока и включают также переходные сопротивления между этими областями и металлическими контактами стока и истока. Наконец, резистивный элемент R_B соответствует сопротивлению объема подложки между «внутренним» электродом b и внешним выводом B .

Основная трудность при моделировании МДПТ связана с распределенным характером емкости затвор—канал C_G . Очевидно, что адекватная модель должна содержать бесконечное количество элементов, чтобы отразить распределенный характер этой емкости. Практически стремятся уменьшить число внутренних узлов эквивалентной схемы, так как это резко упрощает аналитические расчеты, а при машинном моделировании — повышает производительность САПР. Такой подход, разумеется, снижает точность модели, однако и он встречает существенные трудности. В общем случае оказывается принципиально невозможным адекватно моделировать распределенную емкость C_G двумя емкостями C_{gs} и C_{gd} , включенными между «внутренними» электродами затвора, истока и стока. Этот вопрос обсуждается в следующем параграфе. Те же соображения относятся и к емкости C_{gb} , а также к сопротивлению подложки R_B , хотя их распределенный характер в меньшей степени сказывается на точности модели.

III.6.3. Нелинейная эквивалентная схема МДПТ для большого сигнала

Несколько упрощенная эквивалентная схема МДПТ для большого сигнала представлена на рис. III.6.2. Схема содержит все электрические элементы, выделенные на рис. III.6.1.

Штриховым контуром ограничена часть эквивалентной схемы, соответствующая «внутреннему» транзистору. Главным элементом эквивалентной схемы является *управляемый генератор тока* I_d , моделирующий ВАХ «внутреннего» транзистора. Аргументами тока I_d являются напряжения V_{gs} , V_{ds} и V_{bs} между «внутренними» электродами затвора g , истока s , стока d и подложки b . Эти напряжения не совпадают с внешними напряжениями V_{GS} , V_{DS} и V_{BS} вследствие падения напряжений на резистивных элементах R_S , R_D , R_G и R_B . В стационарном состоянии ток затвора $I_G = 0$, однако в общем случае через сопротивление R_G может протекать ток смещения в емкости C_{gs} , C_{gd} и C_{bs} .



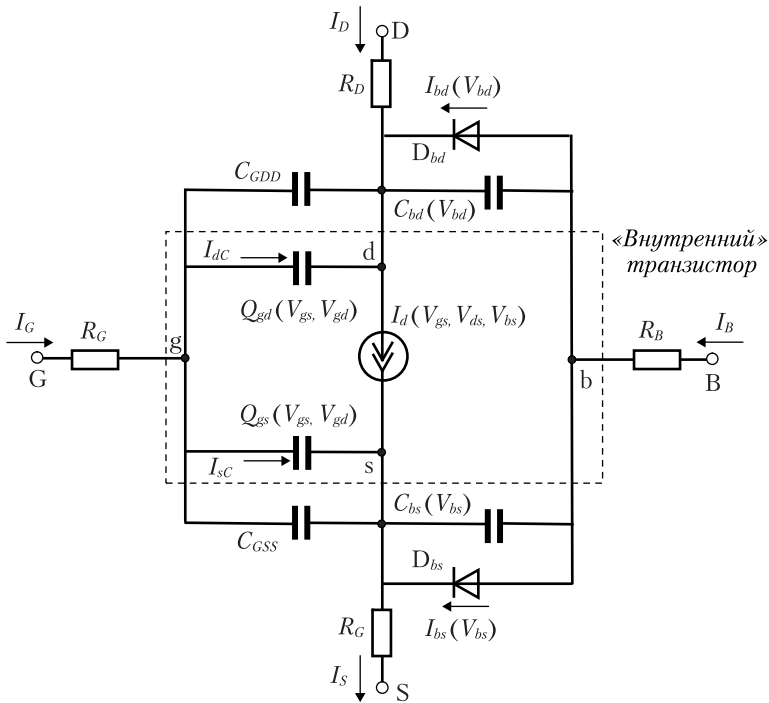


Рис. III.6.2. Эквивалентная схема МДПТ для большого сигнала

Конкретный вид функции $I_d(V_{gs}, V_{ds}, V_{bs})$ зависит от требований к точности модели и может, например, соответствовать уравнениям (III.3.17) идеализированной модели МДПТ, уравнениям (III.4.10), учитывающим неоднородность ОПЗ под затвором, уравнениям (III.5.31), учитывающим модуляцию эффективной длины канала или их модификациям. В моделях низкого уровня (даже компьютерных) влияние потенциала подложки не учитывается. Конечная выходная проводимость в пологой области ВАХ часто учитывается формализованной модификацией функции $I_d(V_{gs}, V_{ds}, V_{bs})$. Широко распространена следующая форма ее представления:

$$I_d(V_{gs}, V_{ds}, V_{bs}) = I_{dS}(V_{gs}, V_{bs}) \cdot F(V_{ds}, V_{gs}) \cdot [1 + \lambda(V_{gs})V_{ds}],$$

где функция $I_{dS}(V_{gs}, V_{bs})$ описывает ток стока насыщения (и не содержит аргумента V_{ds}), функция $F(V_{ds}, V_{gs})$ описывает выходную ВАХ без учета модуляции длины канала, а

функция $\lambda(V_{gs})V_{ds}$ учитывает конечную выходную проводимость в пологой области ВАХ. Функция $I_d(V_{gs}, V_{ds}, V_{bs})$ может также учитывать нарастание тока стока вследствие эффекта смыкания ОПЗ стока и истока. При необходимости ток $I_d(V_{gs}, V_{ds}, V_{bs})$ включает диффузионный подпороговый ток, описываемый, например, уравнением (III.4.18).

В предыдущем параграфе отмечалась принципиальная невозможность точного моделирования распределенной нелинейной емкости затвора C_G двумя емкостями $C_{gs}(V_{gs}, V_{gd})$ и $C_{gd}(V_{gs}, V_{gd})$, включенными между «внутренними» электродами затвора, истока и стока. Такой подход не является адекватным и может привести к принципиальным ошибкам. Путем физического или численного анализа может быть установлена функция $Q_G(V_{gs}, V_{gd})$, определяющая зависимость заряда электронов в канале Q_n и компенсирующего его заряда на затворе $Q_G = -Q_n$ от напряжений затвор–исток и затвор–сток (например, функция (III.3.36) для идеализированной модели). Модель транзистора должна правильно описывать токи смещения I_{sC} и I_{dC} в цепях стока и истока, связанные с изменением заряда Q_G во времени. Эти токи определяются соотношениями

$$I_{sC}, dC = dQ_{gs, gd} / dt, \quad (\text{III.6.1})$$

где Q_{gs} и Q_{gd} — составляющие заряда Q_G , проходящие соответственно через цепи истока и стока. Таким образом, функция $Q_G(V_{gs}, V_{gd})$ должна быть представлена в виде суммы

$$Q_G(V_{gs}, V_{gd}) = Q_{gs}(V_{gs}, V_{gd}) + Q_{gd}(V_{gs}, V_{gd}). \quad (\text{III.6.2})$$

При этом

$$\begin{aligned} \left| \frac{dQ_{gs}}{dQ_{gd}} \right| &= \left| \frac{\partial Q_{gs}}{\partial V_{gs}} \right| dV_{gs} + \left| \frac{\partial Q_{gs}}{\partial V_{ds}} \right| dV_{ds} = \\ &= \left| \frac{C_{sgs}}{C_{dgs}} \right| dV_{gs} + \left| \frac{C_{sds}}{C_{dds}} \right| dV_{ds}, \end{aligned} \quad (\text{III.6.3})$$

где $C_{sgs, sds} = \partial Q_{gs} / \partial V_{gs, ds}$; $C_{dgs, gds} = \partial Q_{gd} / \partial V_{gs, ds}$, а токи смещения в цепях истока, стока описываются соотношениями (III.6.1) в виде

$$\left| \frac{I_{sC}}{I_{dC}} \right| = \left| \frac{C_{sgs}}{C_{dgs}} \right| \frac{dV_{gs}}{dt} + \left| \frac{C_{sds}}{C_{dds}} \right| \frac{dV_{ds}}{dt}. \quad (\text{III.6.4})$$



Таким образом, заряд $Q_G(V_{gs}, V_{ds})$ в емкости затвора должен моделироваться двумя *накопителями заряда* $Q_{gs}(V_{gs}, V_{ds})$ и $Q_{gd}(V_{gs}, V_{ds})$, включенными соответственно в цепи истока и стока (см. рис. III.6.2). Это соответствует представлению емкости затвор—канал *тензором* 2-го ранга

$$C_G = \begin{vmatrix} C_{sgs}(V_{gs}, V_{ds}) & C_{sds}(V_{gs}, V_{ds}) \\ C_{dgs}(V_{gs}, V_{ds}) & C_{dds}(V_{gs}, V_{ds}) \end{vmatrix},$$

компоненты которого C_{sgs} и C_{sds} включены в цепь истока, а C_{dgs} и C_{dds} — в цепь стока.

Заметим, что моделирование заряда $Q_G(V_{gs}, V_{gd})$ двумя емкостями $C_{gs,gd}(V_{gs}, V_{gd}) = \partial Q_G(V_{gs}, V_{gd}) / \partial V_{gs,gd}$ с представлением токов смещения в виде

$$I_{s,c,dC} = \frac{\partial Q_G}{\partial V_{gs,gd}} \cdot \frac{dV_{gs,gd}}{dt} = C_{gs,gd} \frac{dV_{gs,gd}}{dt}$$

не только ошибочно, но и приводит к физически нелепому результату. Выражение $C_{gs} dV_{gs} + C_{gd} dV_{gd}$ *не является полным дифференциалом* заряда Q_G , поэтому контурные интегралы $\oint (C_{gs} dV_{gs} + C_{gd} dV_{gd})$ на плоскости (V_{gs}, V_{gd}) не равны нулю и при многократном обходе по контуру в одном направлении заряды в этих емкостях будут неограниченно накапливаться.

Правильное разбиение заряда Q_G на составляющие $Q_{gs,gd}$ встречает существенные трудности. Можно показать, что *в рамках идеализированной модели* МДПТ, когда функция $Q_G(V_{gs}, V_{ds})$ описывается соотношением (III.3.36), функции $Q_{gs,gd}(V_{gs}, V_{gd})$ имеют вид

$$\frac{Q_{gs}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ 2(V_{gst}^2 + V_{gst} V_{gdt} - V_{gdt}^2 / 2) / 3(V_{gst} + V_{gdt}), & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 2V_{gst} / 3, & V_{gst} \geq 0, V_{gdt} < 0; \end{cases} \quad (\text{III.6.5a})$$

$$\frac{Q_{gd}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ V_{gdt}^2 / (V_{gst} + V_{gdt}), & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 0, & V_{gst} \geq 0, V_{gdt} < 0. \end{cases} \quad (\text{III.6.5б})$$

Отсюда могут быть вычислены компоненты тензора емкости затвор—канал



$$\frac{C_{sgs}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ 2/3 + V_{gdt}^2 / 3(V_{gst} + V_{gdt})^2, & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 2/3, & V_{gst} \geq 0, V_{gdt} < 0, \end{cases} \quad (\text{III.6.6a})$$

$$\frac{C_{dgd}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ V_{gdt}(2V_{gst} + V_{gdt}) / (V_{gst} + V_{gdt})^2, & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 0, & V_{gst} \geq 0, V_{gdt} < 0, \end{cases} \quad (\text{III.6.6б})$$

$$\frac{C_{sgd}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ -V_{gdt}(2V_{gst} + V_{gdt}) / 3(V_{gst} + V_{gdt})^2, & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 0, & V_{gst} \geq 0, V_{gdt} < 0; \end{cases} \quad (\text{III.6.6в})$$

$$\frac{C_{dgs}}{C_G} = \begin{cases} 0, & V_{gst} < 0, \\ -V_{gdt}^2 / (V_{gst} + V_{gdt})^2, & V_{gst} \geq 0, V_{gdt} \geq 0, \\ 0, & V_{gst} \geq 0, V_{gdt} < 0. \end{cases} \quad (\text{III.6.6г})$$

Заряды $Q_{gs, gd}$ и компоненты тензора емкости затвор—канал как функции отношения V_{ds} / V_{gst} для идеализированной модели МДПТ представлены на рис. III.6.3.

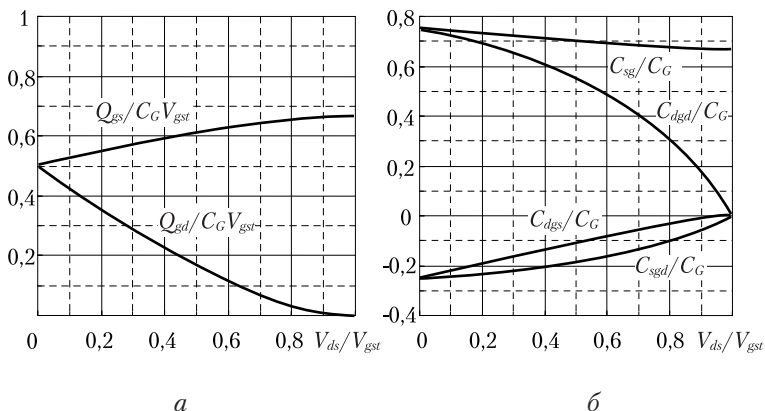


Рис. III.6.3. Заряды $Q_{gs, gd}$ и компоненты тензора емкости затвор—канал как функции отношения V_{ds} / V_{gst} для идеализированной модели МДПТ

Отметим, что в компьютерных моделях функции $Q_{gs, gd}(V_{gs}, V_{gd})$ должны быть непрерывными и дифференцируемыми, в то время как функции (III.6.5) терпят разрыв производных $\partial Q_{gs, gd} / \partial V_{gs, gd}$ при $V_{gst} = 0$. Этот недостаток устраняется формальной коррекцией функций $Q_{gs, gd}(V_{gs}, V_{gd})$ в небольшом диапазоне напряжений $V_{gst} < 0$.

Остальные элементы эквивалентной схемы моделируют паразитные емкости и сопротивления «внешней» части транзистора. Диоды $D_{bs, bd}$, как правило, не работают при прямом включении, однако для выявления схемотехнических ошибок их ВАХ обычно задается в форме (II.5.34)

$$I_{bs, bd} = I_{bs0, bd0} (e^{V_{bs, bd} / m\varphi_T} - 1).$$

При необходимости к токам $I_{bs, bd}$ добавляются токи, описывающие предпробойный режим, например, в форме (II.6.6):

$$I_{break\ bs, bd} = I_{break0} / [1 - (V_{bs, bd} / V_{br})^v].$$

Эти токи могут быть существенными в диоде D_{bd} , который работает при обратном напряжении $V_{bd} = -(V_{bs} + V_{ds})$.

Вольтфарадные характеристики емкостей $C_{bs, bd}(V_{bs, bd})$ соответствуют ВФХ барьерных емкостей p - n переходов подложка—исток, подложка—сток и в компьютерных моделях описываются обычно соотношением

$$C_{bs, bd}(V_{bs, bd}) = C_{bs0, bd0} \times \begin{cases} \sqrt{k\varphi_{BS} / (\varphi_{BS} - V_{bs, bd})}, & V_{bs} \leq \kappa\varphi_{BS}, \quad (\text{III.6.7a}) \\ [1 - \kappa(1+k) + kV_{bs, bd} / \varphi_{BS}] / (1 - \kappa)^{1+k}, & V_{bs} > \kappa\varphi_{BS}. \quad (\text{III.6.7б}) \end{cases}$$

Формула (III.6.7а) соответствует соотношению (II.7.2). Формула (III.6.7б), где $\kappa \approx 1/2$, формально корректирует погрешность приближения полного обеднения ОПЗ, принятого при выводе (II.7.2), и предотвращает расходимость функций $C_{bs, bd}(V_{bs, bd})$ при $V_{bs, bd} = \varphi_{BS}$. Вид функций $C_{bs, bd}(V_{bs, bd})$ представлен на рис. III.6.4. При необходимости емкости $C_{bs, bd}$ могут включать диффузионные составляющие, например, в форме (II.7.11):

$$C_{Dbs, bd} = \tau_B (I_{bs, bd} - I_{bs0, bd0}) / m\varphi_T.$$

Начальные значения $C_{bs0, bd0}$ этих емкостей определяются площадью контактных n^+ -областей $L_{S,D}Z$, где $L_{S,D}$ — их длина (см. рис. III.6.1), а также равновесной шириной p - n переходов $l_{S0,D0} \approx l_{t0}$.



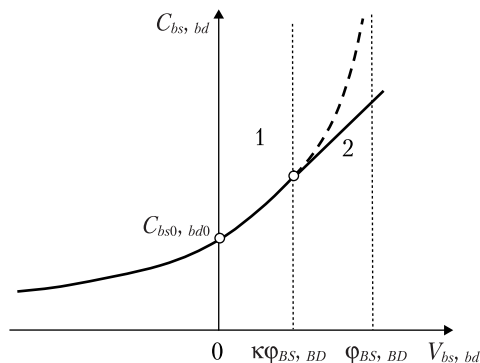


Рис. III.6.4. Вольфарадные характеристики барьерных емкостей $C_{bs, bd}$:
1 — формула (III.6.7a); 2 — формула (III.6.7б)

Емкости перекрытия $C_{GSS, GDD}$ не зависят от напряжений и определяются площадью перекрытия электрода затвора с контактными n^+ -областями истока и стока $L_{SS}Z$ и $L_{DD}Z$.

Из резистивных элементов наибольшее влияние на ВАХ МДПТ оказывает сопротивление R_S . Его действие проявляется в отрицательной обратной связи по напряжению последовательного типа, так как падение напряжения на нем $I_S R_S$ снижает внешнее напряжение затвор—исток: $V_{gs} = V_{GS} - I_S R_S$. Влияние R_S проявляется при значительных токах истока, когда сопротивление становится сравнимым с обратной крутизной ВАХ «внутреннего транзистора» ($R_S \sim 1/g$) и проявляется в линеаризации проходных ВАХ.

Сопротивление затвора R_G не влияет на статические ВАХ МДПТ, поскольку через него протекают только токи смещения. Его действие проявляется в снижении предельной частоты транзистора. Заметим, что ток через него протекает в направлении, перпендикулярном оси сток—исток, и сопротивление R_G пропорционально ширине канала Z (в отличие от сопротивлений $R_{S, D} \sim 1/Z$). Поэтому его влияние проявляется в транзисторах с широким и коротким каналом. В таких случаях для снижения сопротивления R_G применяют *секционирование транзистора* — разбиение его на параллельно включенные секции малой ширины. Этот прием снижает сопротивление R_G в N^2 раз, где N — число секций.

III.6.4. Линейные эквивалентные схемы МДПТ для малого сигнала

Для анализа и проектирования линейных электронных устройств используются *малосигнальные* (линеаризованные) эквивалентные схемы МДПТ. В малосигнальных эквивалентных схемах режим работы транзистора (постоянные составляющие напряжений V_{gs} , V_{ds} , V_{bs} и постоянная составляющая тока I_d) считается заданным и исследуются только малые переменные составляющие напряжений и токов (сигналы). Малосигнальная эквивалентная схема, представленная на рис. III.6.5, *a*, получена из эквивалентной схемы для большого сигнала (см. рис. III.6.2) путем линеаризации ее элементов и исключения постоянных составляющих токов. Поскольку ток I_d является функцией трех напряжений V_{gs} , V_{ds} , V_{bs} , получаем

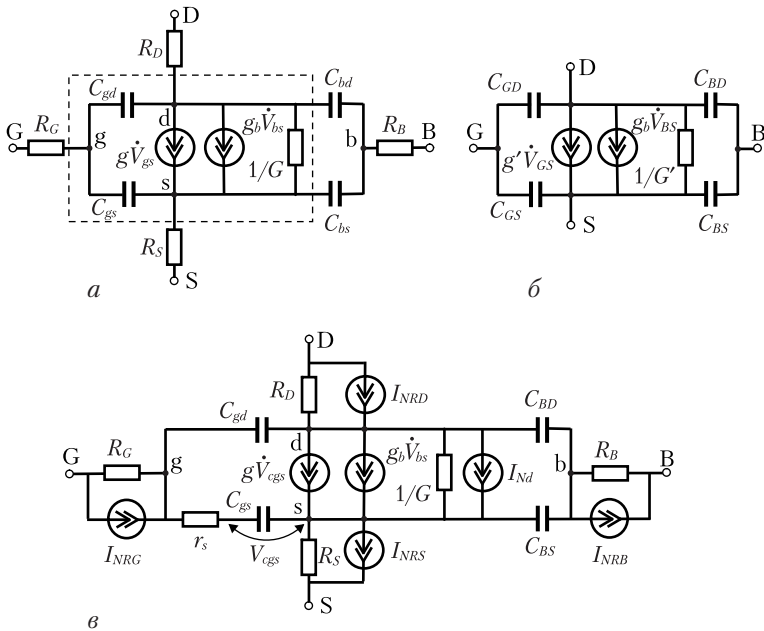


Рис. III.6.5. Малосигнальные эквивалентные схемы МДПТ:

- a* — линеаризация схемы для большого сигнала;
- б* — упрощенная эквивалентная схема без внутренних узлов;
- в* — эквивалентная схема, учитывающая внутреннюю инерционность транзистора и источники шума

$$\begin{aligned} \dot{I}_d &= \frac{\partial I_d}{\partial V_{gs}} \dot{V}_{gs} + \frac{\partial I_d}{\partial V_{ds}} \dot{V}_{ds} + \frac{\partial I_d}{\partial V_{bs}} \dot{V}_{bs} = \\ &= g \dot{V}_{gs} + G \dot{V}_{ds} + g_b \dot{V}_{bs}, \end{aligned} \quad (\text{III.6.8})$$

где точками помечены малые переменные составляющие тока и напряжений; $g = \partial I_d / \partial V_{gs}$ — крутизна; $G = \partial I_d / \partial V_{ds}$ — выходная проводимость; $g_b = \partial I_d / \partial V_{bs}$ — *крутизна по подложке*.

Составляющие $g \dot{V}_{gs}$ и $g_b \dot{V}_{bs}$ тока $\dot{I}_d$ моделируются специальными управляемыми генераторами тока (см. рис. III.6.5, а). Для составляющей $G \dot{V}_{ds}$ нет необходимости вводить отдельный генератор тока, так как она может быть учтена простым резистивным элементом с сопротивлением $1/G$ (см. рис. III.6.5, а).

В малосигнальной эквивалентной схеме предполагается, что p - n переходы подложка—исток и подложка—сток никогда не находятся под прямым напряжением, поэтому переменные составляющие токов $\dot{I}_{bs, bd}$ ничтожно малы и в схеме не учитываются. Кроме того, обычно считается, что транзистор работает в пологой области ВАХ, где он обладает с наилучшими усилительными свойствами.

Все емкости в линейной модели являются константами, величина которых определяется режимом работы транзистора по постоянному току. Рисунок III.6.3 показывает, что диагональные компоненты C_{sgd} и C_{dgs} тензора емкости затвор—канал практически одинаковы. При этом емкость затвор—канал может быть отображена тремя элементами C_{gs} , C_{gd} и C_{ds} , включенными между электродами истока, стока и затвора (см. рис. III.6.5, а). Значения емкостей C_{gs} , C_{gd} и C_{ds} находятся из условий равенства токов смещения

$$\begin{aligned} I_{sC} &= \frac{C_{gs} dV_{gs}}{dt} + \frac{C_{ds} dV_{ds}}{dt} = \frac{(C_{gs} + C_{ds}) dV_{gs}}{dt} - \frac{C_{ds} dV_{gd}}{dt}, \\ I_{dC} &= \frac{C_{gd} dV_{gd}}{dt} - \frac{C_{ds} dV_{ds}}{dt} = \frac{(C_{gd} + C_{ds}) dV_{gd}}{dt} - \frac{C_{ds} dV_{gs}}{dt} \end{aligned}$$

в схеме на рис. III.6.5, а токам смещения, определенным в (III.6.4):

$$C_{gs} = C_{sgs} + C_{sgd}; \quad (\text{III.6.9a})$$

$$C_{gd} = C_{dgd} + C_{sgd}; \quad (\text{III.6.9б})$$



$$C_{ds} = -C_{sgd} = -C_{dgs}. \quad (\text{III.6.9в})$$

Из соотношений (III.6.6) следует, что в рамках идеализированной модели МДПТ в пологой области ВАХ $C_{dgd} = C_{sgd} = C_{dgs} = 0$. При этом из (III.6.9а–в) получаем $C_{gd} = C_{ds} = 0$.

На рис. III.6.5, б представлена упрощенная эквивалентная схема, не содержащая внутренних узлов. В этой схеме исключены сопротивления затвора R_G , n^+ -областей стока/истока и омических контактов к этим областям R_D и R_S , также подложки R_B . Наиболее существенно влияние сопротивления истока R_S . Как отмечалось в предыдущем параграфе, это сопротивление создает отрицательную обратную связь по напряжению последовательного типа. В соответствии с общей теорией обратной связи это приводит к снижению крутизны и выходной проводимости МДПТ по сравнению с крутизной g и выходной проводимостью G «внутреннего» транзистора. В эквивалентной схеме на рис. III.6.5, б это учтено использованием эффективных значений крутизны и выходной проводимости:

$$g' = g / (1 + gR_S). \quad (\text{III.6.10а})$$

$$G' = G / (1 + gR_S). \quad (\text{III.6.10а})$$

Упрощенные эквивалентные схемы используются для предварительных оценочных расчетов линейных ИС высокой степени интеграции с целью экономии времени и вычислительных ресурсов.

Во многих случаях малосигнальные модели должны обладать более высокой точностью, чем нелинейные модели для большого сигнала. Это связано с тем, что линейные аналоговые устройства обычно включают цепи обратной связи, которые могут привести к потере устойчивости (паразитной генерации). В связи с этим следует отметить, что эквивалентные схемы на рис. III.6.5, а и особенно на рис. III.6.5, б не отражают внутренней инерционности транзистора, так как при малом сопротивлении истока R_S управляющее напряжение затвор–исток $\dot{V}_{gs}$ подается непосредственно на емкость C_{gs} . Если источник сигнала имеет малое выходное сопротивление и близок по свойствам к источнику ЭДС, то он заряжает емкость C_{gs} практически мгновенно и мгновенно вводит в действие генератор тока $g\dot{V}_{gs}$. Этот недостаток связан с учетом распределенного характера емкости затвора, о котором говорилось выше. Частично его можно нейтрализовать

включением последовательно с емкостью C_{gs} сопротивления r_s , как это показано на рис. III.6.5, в. Можно показать, что сопротивление r_s обратно пропорционально крутизне ВАХ

$$r_s \approx 1 / 3g. \quad (\text{III.6.11})$$

В модели, предназначенные для анализа и проектирования линейных устройств СВЧ-диапазона, вводятся дополнительные реактивные элементы, учитывающие конструктивные реактивности — емкости между контактными площадками выводов и индуктивности цепей истока, затвора и стока.

При моделировании линейных устройств коммуникационной и измерительной техники большое значение имеет учет внутренних шумов транзистора. Шумовые свойства полупроводниковых приборов детально проанализированы в работах [12, 13]. В эквивалентной схеме на рис. III.6.5 в основные источники шума учтены шумовыми генераторами тока I_{NRG} , I_{NRS} , I_{NRD} , I_{NRB} и I_{Nd} . Генераторы тока I_{NRG} , I_{NRS} , I_{NRD} , I_{NRB} моделируют тепловой шум соответствующих сопротивлений R_G , R_S , R_D , R_B . Спектральная плотность теплового шума не зависит от частоты (белый шум) и определяется только сопротивлением

$$\overline{I_{NR}^2} = 4kT / R. \quad (\text{III.6.12})$$

Генератор тока I_{Nd} моделирует шумовой ток канала, который включает тепловой шум выходного сопротивления транзистора $1/G$ и фликкер-шум, связанный со случайными процессами генерации—рекомбинации носителей заряда, захватом их на мелкие ловушечные уровни, взаимодействием с поверхностными и объемными дефектами, а также другими процессами, многие из которых остаются неисследованными. Спектральная плотность фликкер-шума обратно пропорциональна частоте («шум вида $1/f$ ») и зависит от постоянной составляющей тока стока I_d . С учетом распределенного характера сопротивления канала спектральная плотность тока I_{Nd} составляет

$$\overline{I_{Nd}^2} = \frac{8kTG}{3} + \frac{K_F I_d^{A_F}}{f}, \quad (\text{III.6.13})$$

где коэффициенты K_F и $A_F \approx 1$ могут быть определены только экспериментально.

Отметим, что учет внутренних источников шума необходим при анализе и проектировании не только приемных усилительных устройств, предназначенных для обработки



весьма малых сигналов, но и генераторных устройств, так как шумы искажают чистоту спектра генерированного сигнала («частотно-модулирующий шум»).

III.6.5. Формальные линейные модели

Формальный метод моделирования линейного четырехполюсника состоит в представлении его в виде «черного ящика» неизвестного содержания, имеющего два входных и два выходных электрода (рис. III.6.6). Электрическое состояние четырехполюсника определяется четырьмя электрическими переменными: входным и выходным напряжениями $\dot{V}_1$, $\dot{V}_2$, а также входным и выходным токами $\dot{I}_1$, $\dot{I}_2$. Из этих четырех переменных любые две могут быть выбраны в качестве независимых (аргументов), а две другие (функции) выражаются через аргументы линейными уравнениями. Всего возможны четыре способа выбора аргументов и функций, определяющих систему параметров модели; они приведены в табл. III.6.2.



Рис. III.6.6. Формальное моделирование линейного четырехполюсника как «черного ящика»

Индекс «1» у напряжений и токов соответствует входу, индекс «2» – выходу

Таблица III.6.2

Системы параметров линейного четырехполюсника

Система параметров	z-параметры	y-параметры	h-параметры	a-параметры
Уравнения связи	$\dot{V}_1 = z_{11}\dot{I}_1 + z_{12}\dot{I}_2$	$\dot{I}_1 = y_{11}\dot{V}_1 + y_{12}\dot{V}_2$	$\dot{V}_1 = h_{11}\dot{I}_1 + h_{12}\dot{V}_2$	$\dot{I}_1 = a_{11}\dot{V}_1 + a_{12}\dot{I}_2$
аргументов и функций	$\dot{V}_2 = z_{21}\dot{I}_1 + z_{22}\dot{I}_2$	$\dot{I}_2 = y_{21}\dot{V}_1 + y_{22}\dot{V}_2$	$\dot{I}_2 = h_{21}\dot{I}_1 + h_{22}\dot{V}_2$	$\dot{V}_2 = a_{21}\dot{V}_1 + a_{22}\dot{I}_2$

Малые переменные могут быть гармоническими сигналами, или зависеть от времени. В первом случае они задаются комплексными функциями частоты $\dot{V}_j(\omega) = \dot{V}_j e^{j\omega t}$, $\dot{I}_k(\omega) = \dot{I}_k e^{j\omega t}$,

где $\dot{V}_j$ и $\dot{I}_k$ — комплексные амплитуды напряжения и тока, во втором — операторными изображениями $\dot{V}_j(s)$, $\dot{I}_k(s)$, как это принято в линейной электротехнике.

Выбор системы параметров определяется удобством задания аргументов при измерениях параметров. Если четырехполюсник обладает высоким входным (выходным) импедансом $|\dot{Z}_{in(out)}|$, то удобно задавать на входе (выходе) напряжение (невозможно задать напряжение на нулевом сопротивлении). Если четырехполюсник имеет низкий входной (выходной) импеданс, то удобно задавать на входе (выходе) ток (нельзя задать ток в бесконечное сопротивление). Для МДП-транзистора в области не слишком высоких частот (значительно меньших предельной частоты транзистора) наиболее удобна система y -параметров, так как в наиболее важной схеме включения (общий затвор) МДПТ имеет высокий входной импеданс (ток затвора мал) и высокий выходной импеданс в полгой области ВАХ (выходная проводимость мала).¹

Основное преимущество формального моделирования четырехполюсника как «черного ящика» состоит в удобстве измерения параметров модели, так как параметры модели имеют простой физический смысл. В частности, из уравнений связи аргументов и функций в системе y -параметров следует, что

$\dot{y}_{11} = \dot{I}_1 / \dot{V}_1|_{\dot{V}_2=0}$ — входной адмиттанс² при коротком замы-

кании (КЗ) на выходе;

$\dot{y}_{12} = \dot{I}_1 / \dot{V}_2|_{\dot{V}_1=0}$ — обратная комплексная крутизна при КЗ

на входе;

$\dot{y}_{21} = \dot{I}_2 / \dot{V}_1|_{\dot{V}_2=0}$ — комплексная крутизна при КЗ на выходе;

$\dot{y}_{22} = \dot{I}_2 / \dot{V}_2|_{\dot{V}_1=0}$ — выходной адмиттанс при КЗ на входе.

В системе y -параметров все измерения осуществляются при коротком замыкании, что легко осуществить при малых значениях входного и выходного адмиттанса. Разумеется, КЗ

¹ В диапазоне СВЧ для всех четырехполюсников используется специальная система S -параметров, в которой входной и выходной импедансы источников сигнала имеют значение 50 Ом, аргументами являются мощности входного и выходного сигналов, а функциями — мощности сигналов, отраженных от входа и выхода.

² Адмиттанс (комплексная проводимость) — величина, обратная импедансу.



осуществляется по переменному току (например, подключением к входу или выходу конденсатора большой емкости или LC-фильтра низких частот) и сохраняет возможность задания требуемых постоянных напряжений на входе и выходе (например, напряжений затвор-исток и сток-исток).

Нетрудно видеть, что формальной модели в системе y -параметров соответствует формальная эквивалентная схема, приведенная на рис. III.6.7.

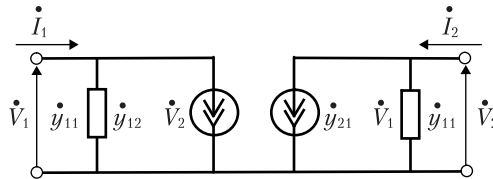


Рис. III.6.7. Формальная эквивалентная схема линейного четырехполюсника в системе y -параметров

Если формальная модель составляется для электронного прибора, ее параметры зависят от способа его включения. Так для МДПТ y -параметры имеют различные значения для схем включения с общим истоком ($y_{jk\text{ОИ}}$) и общим затвором ($y_{jk\text{ОЗ}}$).

Как отмечено выше, формальная модель четырехполюсника весьма удобна для измерения ее параметров, но она не содержит информации об устройстве четырехполюсника. Зная физическую или формализованную эквивалентную схему и ее параметры, можно определить параметры формальной модели, но не наоборот. Так, например, если в упрощенной эквивалентной схеме МДПТ на рис. III.6.5, б исключить емкость затвор—сток C_{GD} (которая в пологой области ВАХ в рамках идеализированной модели равна 0) и пренебречь емкостями C_{BS} и C_{BD} , находящимися за пределами «внутреннего транзистора», получаем:

$$\dot{y}_{11} = i\omega C_{GS}; \quad \dot{y}_{12} = 0; \quad \dot{y}_{21} = g'; \quad \dot{y}_{22} = G.$$

Для активного (способного усилить входной сигнал) электронного прибора в любой системе параметров важнейшим является параметр $\dot{a}_{21}$, где a — обобщенное обозначение параметра. В МДПТ важнейшим является параметр $\dot{y}_{21}$, определяющий его усилительные свойства.

Основные выводы

1. Физические и формализованные модели МДПТ составляются на основе эквивалентных схем, отражающих структуру и конструкцию транзистора.

2. Основная трудность при моделировании МДПТ связана с распределенным характером емкости затвор—канал C_G , которая не является паразитным элементом, а лежит в основе принципа действия прибора. Заряд в этой емкости должен моделироваться двумя накопителями заряда, включенными в цепи истока и стока, что эквивалентно представлению емкости C_G в виде тензора второго ранга.

3. Нелинейная эквивалентная схема МДПТ для большого сигнала соответствует рис. III.6.3.

4. Малосигнальная (линеаризованная) эквивалентная схема МДПТ может быть получена из эквивалентной схемы для большого сигнала путем линеаризации ее элементов и исключения постоянных составляющих токов. Точность малосигнальной эквивалентной схемы может быть повышена за счет учета внутренней инерционности транзистора, внутренних источников шума и паразитных конструктивных реактивностей.

5. С точки зрения удобства измерения параметров оптимальными являются формальные малосигнальные модели, основанные на представлении прибора в виде четырехполюсника как «черного ящика», входные и выходные сигналы которого связаны формальными линейными уравнениями. В области не слишком высоких частот (значительно меньших предельной частоты транзистора) для МДПТ наиболее удобна система y -параметров, в которой аргументами являются входные и выходные напряжения, а функциями — входные и выходные токи.

III.7. Масштабирование МДП-транзисторов

Соотношения (III.3.17) и (III.2.5)—(III.2.6), определяющие ВАХ МДПТ в рамках простейшей идеализированной модели, показывают, что ВАХ содержат два параметра: параметр $\beta = \mu_n C_{sd} (Z / L)$ (III.3.12) и пороговое напряжение V_t . Таким образом, уравнения (III.3.17) подчиняются простому правилу масштабирования, согласно которому **уравнения ВАХ МДПТ не изменяются при сохранении отношения Z/L** . В частности, крутизна ВАХ, максимальное значение которой достигается в пологой области ВАХ, составляет (III.3.27)



$$g_s = V_{Gst} \mu_n C_{sG} (Z / L). \quad (\text{III.7.1})$$

Этому правилу подчиняется и ВАХ диффузионного подпорогового тока (III.4.18).

Пороговое напряжение в соответствии с (III.2.8)–(III.2.9) зависит от электрофизических параметров транзисторной структуры: контактной разности потенциалов затвор–подложка ϕ_{GB} , параметра ϕ_B , а также от отношений поверхностной плотности заряда поверхностных состояний $Q_{ss} = eN_{ss}$ и поверхностной плотности заряда в ОПЗ $Q_{sB0} \sim l_{t0} \sim N_B^{1/2}$ к удельной емкости затвора $C_{sG} = \epsilon \epsilon_0 / d_{ef}$.

Параметр $\phi_B = \phi_T \ln(N_B / n_i)$ очень слабо зависит от свойств подложки – при изменении концентрации примеси N_B на два порядка в пределах $10^{15}–10^{17} \text{ см}^{-3}$ его значение изменяется всего на $\approx 100 \text{ мВ}$. При $N_{ss} < 2 \cdot 10^{10} \text{ см}^{-2}$ и $d_{ef} < 0,1 \text{ мкм}$ $Q_{ss} / C_{sG} < 30 \text{ мВ}$. Поэтому можно считать, что при нулевом напряжении подложка–исток V_{BS} единственным геометрическим размером конструкции МДПТ, существенно влияющим на пороговое напряжение, является пороговая ширина ОПЗ $l_{t0} \sim N_B^{1/2}$. Фактором, определяющим зависимость порогового напряжения от напряжения V_{BS} , является коэффициент подложки $K_B = \sqrt{2\epsilon \epsilon_0 e N_B} / C_{sG}$ (III.2.11). Соотношение (III.2.12) показывает, что коэффициент подложки определяется отношением эффективной толщины диэлектрика (III.1.8) к пороговой ширине ОПЗ d_{ef} / l_{t0} .

Таким образом, **ВАХ МДПТ не изменяются при сохранении отношений**

$$Z / L \text{ и } d_{ef} / l_{t0}.$$

Учет неоднородности ширины ОПЗ под затвором (см. параграф III.4.1) изменяет уравнения ВАХ, вводит в них добавочный аргумент (напряжение подложка–исток) и добавочные параметры – параметр ϕ_B и коэффициент подложки K_B . Однако эти изменения оставляют в силе сформулированное правило.

Соотношение (III.3.40) показывает, что предельная частота «внутреннего» МДПТ пропорциональна напряжению V_{gst} и обратно пропорциональна квадрату длины канала

$$\omega_T = \frac{3}{2} \cdot \frac{\mu V_{gst}}{L^2}. \quad (\text{III.7.2})$$

Структура и эквивалентная схема МДПТ содержит паразитные конструктивные емкости: емкости p – n переходов под



контактными областями $C_{bs,bd}$ и емкости перекрытия затвора $C_{GSS,GDD}$ (см. рис. III.6.1, III.6.2, III.6.4). Удельные значения емкостей $C_{bs,bd}$ определяются толщиной ОПЗ, которая при нулевом напряжении на переходе практически совпадает с l_{t0} . Поэтому отношение этих удельных емкостей к удельной емкости затвора определяется отношением d_{ef}/l_{t0} . Удельные емкости перекрытия равны удельной емкости затвора. Если все топологические размеры МДПТ изменяются пропорционально длине канала, то и абсолютные значения емкостей $C_{bs,bd}$ и $C_{GSS,GDD}$ изменяются пропорционально длине канала. Таким образом, можно утверждать, что **предельная частота МДПТ с длинным каналом подчиняется закономерности $\omega_T \sim V_{gst} / L^2$ при учете всех паразитных реактивностей.**

Эффекты короткого канала обобщают и корректируют сформулированные правила масштабирования. *Краевые эффекты*, искажающие форму краевых областей ОПЗ под контактами стока и истока, приводят к появлению нового масштабного фактора — отношения x_j/l_{t0} толщины контактных областей стока и истока x_j к пороговой ширине ОПЗ l_{t0} . Этот фактор влияет на пороговое напряжение, и **его сохранение дополняет правила масштабирования.**

Опасность *смыкания канала* требует также **сохранения отношения l_{t0}/L** пороговой ширины ОПЗ к длине канала для сохранения напряжения смыкания V_{pinch} (III.5.5в).

Ограничение дрейфовой скорости носителей заряда в канале обобщает соотношение, определяющее крутизну ВАХ в пологой области. Согласно (III.5.13а), (III.5.14)

$$g_S = g_{\max} \frac{V_{gst}/V_L}{\sqrt{1 + V_{gst}^2/V_L^2}}, \quad (\text{III.7.3a})$$

$$\text{где} \quad g_{\max} = Zv_s C_{sG}. \quad (\text{III.7.3б})$$

Таким образом, описание этого эффекта требует введения нового параметра МДПТ

$$V_L = E_s L, \quad (\text{III.7.4})$$

а правила масштабного сохранения ВАХ дополняются новым параметром V_L и требуют **сохранения отношения V_{gst}/V_L .**

При $V_{gst}/V_L \gg 1$ достижимая предельная частота МДПТ составляет (III.5.19)

$$\omega_{T\max} = v_s / L. \quad (\text{III.7.5a})$$

Здесь уместно заметить, что соотношение (III.7.2) для предельной частоты МДПТ с длинным каналом ($V_{gst}/V_L \ll 1$) можно формально записать в виде

$$\omega_T = (3V_{gst}/2V_L) \cdot (v_s/L). \quad (\text{III.7.5б})$$

Для произвольной длины канала формулы (III.7.5а, б) обобщаются аппроксимацией

$$\omega_T \approx \frac{v_s/L}{1 + 2V_L/3V_{gst}} = \frac{\omega_{T\max}}{1 + 2V_L/3V_{gst}}, \quad (\text{III.7.6})$$

асимптотически точной в предельных случаях $V_{gst}/V_L \ll 1$ и $V_{gst}/V_L \gg 1$. Таким образом, следует считать, что **сохранение отношения $\omega_T/\omega_{T\max}$ обеспечивается сохранением отношения V_{gst}/V_L .**

Дальнейшие коррективы в правила масштабирования вносит *эффект модуляции эффективной длины канала*. Его принципиально важным следствием является *ограничение собственного коэффициента усиления МДПТ по напряжению $K_v = g/G$ за счет возрастания выходной проводимости G в пологой области ВАХ при уменьшении длины канала*. В разделе III.5 показано, что в пологой области ВАХ коэффициент усиления K_v является функцией безразмерного напряжения $v_{gst} = V_{gst}/V_L$ и отношения длины канала к параметру a (III.5.24), который определяется как среднее геометрическое между эффективной толщиной диэлектрика и шириной ОПЗ вблизи стока: $a = \sqrt{d_{ef}l_D}$ (III.5.24). Таким образом, правила масштабирования дополняются **сохранением отношения L/a .**

Рассмотренные правила масштабирования относятся к «внутреннему» транзистору. Эквивалентные схемы МДПТ включают также паразитные емкости и сопротивления «внешней» части транзистора. Сохранение отношения паразитных емкостей перекрытия затвора $C_{gss, gdd}$, а также емкостей подложка—исток и подложка—сток $C_{bs, bd}$ к емкости затвора C_g требует пропорционального уменьшения размеров $L_{ss, dd}$ областей затвора, а также длины $L_{s, d}$ контактных областей истока и стока (рис. III.6.1). Это требование дополняет правила масштабирования **сохранением отношений $L_{ss, dd}/L$ и $L_{s, d}/L$.**

Заметим, что пропорциональное уменьшение *торцевых* составляющих емкостей $C_{bs, bd}$ обеспечивается рассмотренным выше правилом сохранения отношений x_j/l_{t0} и l_{t0}/L .



Сформулированные выше правила масштабирования сведены в табл. III.7.1.

Таблица III.7.1

Правила масштабирования МДПТ

№	Эффект	Характеристика МДПТ	Масштабный фактор
1		ВАХ	Z/L
2	Модуляция ширины ОПЗ	Пороговое напряжение V_p , коэффициент подложки K_B	d_{ef}/l_{t0}
3	Краевые эффекты в ОПЗ	Пороговое напряжение V_t	x_j/l_{t0}
4	Смыкание канала	Напряжение смыкания V_{pinch}	l_{t0}/L
5	Ограничение дрейфовой скорости носителей	Напряжение насыщения V_{ds^*} , ток насыщения I_{DS} , крутизна g_s , предельная частота ω_T	V_{gs^*}/V_L
6	Модуляция длины канала	Коэффициент усиления напряжения K_V	$L/(d_{ef}/l_{t0})^{1/2}$
7	Паразитные емкости	Предельная частота ω_T	$L_{SS,DD}/L;$ $L_{S,D}/L$

Общий закон масштабирования МДПТ состоит в том, что *все геометрические размеры его конструкции должны изменяться пропорционально при сохранении электрофизических параметров полупроводника, диэлектрика и затвора.*

Этот закон вытекает из симметрии уравнений, описывающих характеристики МДПТ *в рамках рассмотренных теорий*, однако его строгое соблюдение невозможно. Принципиальной причиной является *несоблюдение принципа подобия* фундаментальных физических уравнений относительно изменения геометрических размеров объекта.

Законы квантовой физики сохраняют неизменными размеры микрочастиц и связанные с ними параметры кристаллической решетки твердых тел. Представляется принципиально возможным пропорциональное уменьшение топологических и вертикальных размеров (толщины полупроводниковых и диэлектрических слоев) в структуре МДПТ. Однако пропорциональное уменьшение ширины ОПЗ уже в настоящее время становится невозможным. Ширина ОПЗ l уменьша-



ется при повышении концентрации примеси в подложке N_B пропорционально $N_B^{-1/2}$. За последние 10–15 лет длина канала МДПТ уменьшилась примерно на порядок. Таким образом, пропорциональное уменьшение ширины ОПЗ требует повышения степени легирования в 100 раз. Тенденция повышения степени легирования полупроводниковых слоев в микроэлектронике действительно существует, но ее развитие в принципе ограничено. Не говоря о технических трудностях, естественной границей является предел растворимости примеси в полупроводнике.

При уменьшении линейных размеров до значений, сравнимых с длиной свободного пробега, движение носителей заряда приобретает баллистический характер, теряется смысл их кинетических коэффициентов (подвижности, коэффициента диффузии) и становятся некорректными уравнения переноса в использованных нами моделях. Кроме того, при достаточно малых размерах становятся статистически недостоверными такие понятия, как концентрации носителей заряда и примесных ионов.

Принцип геометрического подобия нарушается и в макрообъектах. Площади поверхностей и объемы изменяются пропорционально квадрату и кубу линейных размеров, соответственно. По этой причине не могут существовать геометрически подобные объекты существенно различных размеров как в мире техники, так и в живой природе.

Одним из масштабных факторов является отношение напряжений V_{gst}/V_L . Согласно приведенной в разделе III.5 теории уменьшение размеров ($V_L \sim L$) должно приводить к снижению питающих напряжений и токов в МДПТ. Такая тенденция в микроэлектронике также существует, однако сохранение фактора V_{gst}/V_L не реализуется. За упомянутые выше 10–15 лет напряжение питания в процессорах компьютеров снизилось в 2–3 раза, т.е. в 3–5 раз меньше, чем длина канала. Одна из трудностей здесь состоит в снижении помехоустойчивости и помехозащищенности ИС при снижении питающего напряжения.

В заключение отметим, что здесь сформулированы лишь наиболее общие правила масштабирования, не касающиеся технологических тонкостей изготовления МДПТ. Детальные сведения по вопросам масштабирования МДП-ИС содержатся в работе [14].



III.8. Структуры короткоканальных МДП-транзисторов

III.8.1. Требования к МДПТ в СБИС

В настоящее время МДПТ являются основными активными элементами СБИС. Как было показано, усилительные и частотные свойства транзисторов улучшаются с уменьшением длины канала. Однако при этом проявляется ряд эффектов (раздел III.5), которые накладывают определенные ограничения (помимо технологических) на минимальную длину канала. Нейтрализация этих эффектов является одной из основных задач при проектировании МДП-транзисторов. При минимальных размерах транзисторы в СБИС должны иметь стабильное пороговое напряжение, зависимость которого от напряжений подложка—исток и сток—исток сведена к минимуму. Высокое значение крутизны может быть достигнуто только при достаточно малых паразитных сопротивлениях стока и истока; подпороговый ток должен быть достаточно мал, а напряжение смыкания достаточно велико, чтобы обеспечить нормальную работу транзистора при напряжениях питания 2–5 В. Для повышения быстродействия должны быть достаточно малы паразитные емкости исток—подложка, сток—подложка, канал—подложка, а также краевые емкости затвор—исток и затвор—сток. Высокая помехозащищенность логических элементов в БИС может быть достигнута только при достаточно большом собственном коэффициенте усиления по напряжению, однако его величина снижается при уменьшении длины затвора.

Для структуры МДПТ, изображенной на рис. III.2.1, эти требования являются взаимно противоречивыми. Так, например, малые значения паразитных емкостей на подложку могут быть обеспечены за счет снижения концентрации легирующей примеси N_B , однако при этом снижается напряжение смыкания, возрастают зависимость порогового напряжения от напряжения сток—исток и токи утечки. Для уменьшения паразитных сопротивлений стока и истока необходимо увеличивать толщину контактных областей x_j под электродами стока и истока, что усиливает нежелательные эффекты короткого канала.

В большинстве случаев для реализации СБИС необходимо сочетание МДПТ с каналами n - и p -типа (комплементарные транзисторные пары). Поэтому проектирование транзисторных структур должно предусматривать возможность

совместного изготовления комплементарных пар в едином технологическом цикле.

Ниже рассмотрены основные разновидности транзисторных структур МДП-СБИС, в той или иной степени удовлетворяющие предъявляемым требованиям. Материал этой главы не претендует на детальное описание всех элементов структуры современных МДПТ и подробностей технологических процессов; эти вопросы детально освещены в работах [14–16]. Мы здесь стараемся лишь показать методы достижения компромиссов для удовлетворения противоречивых требований к правилам масштабирования МДПТ (см. раздел III.7), а также к технологическим ограничениям.

III.8.2. Типовая структура короткоканального МДПТ

Типовая структура короткоканального n -канального МДПТ схематично показана на рис. III.8.1 [17, 18]. Особенности структуры направлены на выполнение следующих требований:

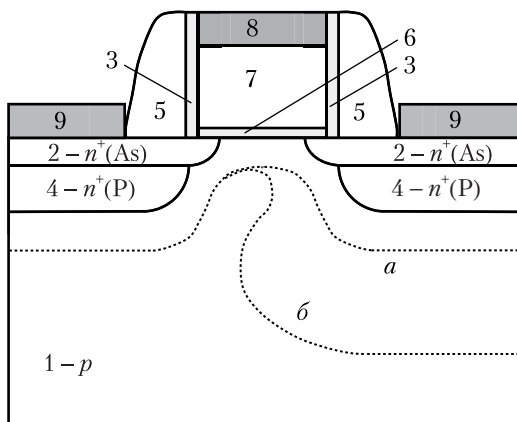


Рис. III.8.1. Типовая структура короткоканального n -канального МДПТ:

- 1 — p -подложка ($N_B \sim 10^{17} \text{ см}^{-3}$); 2 — расширенные контактные области (extended spaces, As); 3 — тонкие спейсеры (thin spacer, SiO_2); 4 — толстые контактные n^+ -области (P); 5 — толстые спейсеры (thick spacer, SiO_2); 6 — подзатворный диэлектрик SiO_2 (Si_3N_4 , N/O, ZrO_2 , HfO_2 ,...); 7 — затвор (n^+ - или p^+ - Si^*); 8 — металлическое покрытие затвора; 9 — металлические электроды истока, стока. Пунктир — границы ОПЗ при $V_{ds} = 0$ (a) и $V_{ds} \approx V_{pinch}$ (б)

- 1) малая длина канала;
- 2) малые паразитные сопротивления пассивных областей стока и истока;
- 3) малые емкости перекрытия затвора с электродами истока и стока;
- 4) малое сопротивление затвора;
- 5) приемлемое значение коэффициента усиления по напряжению;
- 6) приемлемое значение напряжения смыкания;
- 7) требуемое значение порогового напряжения.

Возможности выполнения этих требований обеспечиваются перечисленными ниже мерами.

1. **Длина канала** может быть уменьшена до 0,12–0,2 мкм в промышленных изделиях, а в лабораторных образцах — до величины порядка 15 нм. Возможность получения столь малых размеров методами оптической литографии обеспечивается применением операций *самосовмещения* при изготовлении транзистора.

Для создания n^+ -областей стока и истока применяются операции ионного легирования. Малое боковое расширение этих областей в сторону канала обеспечивается их двухслойной структурой. Физическая длина канала определяется расстоянием между тонкими сильнолегированными n^+ -слоями — расширенными контактными областями 2 (extended spaces), для легирования которых используются доноры с малым коэффициентом диффузии (обычно мышьяк — As). При операции легирования маской служит затвор, что обеспечивает высокую воспроизводимость топологических размеров канала. Для уменьшения бокового расширения тонких n^+ -слоев в сторону канала перед их формированием боковые поверхности затвора покрываются весьма тонкими защитными вертикальными слоями диоксида кремния SiO_2 — тонкими спейсерами 3 (thin spacer).

2. **Сопротивления пассивных областей стока и истока** снижаются за счет использования сравнительно толстых сильнолегированных контактных n^+ -областей 4, для легирования которых используются доноры с высоким коэффициентом диффузии (обычно фосфор — P). Внутренние границы этих областей должны быть расположены как можно ближе к затвору. Это достигается использованием толстых вертикальных слоев SiO_2 — толстых спейсеров 5 (thick spacer), которые служат ограничивающими масками при создании n^+ -областей. Толщина толстых спейсеров выбирается при-

мерно равной толщине формируемых n^+ -областей. При этом внутренние границы n^+ -областей автоматически оказываются расположенными вблизи границы между тонкими и толстыми спейсерами.

3. Эти меры обеспечивают также **минимальные значения емкостей перекрытия затвор—исток, затвор—сток**, поскольку технология самосовмещения до минимума снижает размеры областей перекрытия канала с электродами стока и истока.

4. **Коэффициент усиления по напряжению** снижается при уменьшении длины канала. В разделах III.5, III.7 показано, что для сохранения приемлемого значения коэффициента усиления необходимо уменьшать значение параметра $a = \sqrt{d_{ef} l}$ (l — ширина ОПЗ; $d_{ef} = d\epsilon / \epsilon_d$ — эффективная толщина подзатворного диэлектрика 6). Наиболее эффективным является уменьшение толщины диэлектрика d . Минимальное значение толщины диэлектрика ограничено его диэлектрической прочностью, возможностью образования сквозных дефектов (пор) и протекания туннельного тока.

До настоящего времени в качестве подзатворного диэлектрика (6 на рис. III.8.1) в основном используется диоксид кремния SiO_2 , обладающий непревзойденными технологическими характеристиками (простота получения, высокая термостабильность и надежность, малые утечки, малый встроенный заряд, низкая плотность поверхностных состояний). Основным недостатком этого диэлектрика является низкая диэлектрическая проницаемость ($\epsilon_d = 3 - 4$). В последнее время исследуется возможность использования диэлектриков HfO_2 , Al_2O_3 , $\text{HfO}_2/\text{Al}_2\text{O}_3$, ZrO_2 , $\text{AlN}_y(\text{O}_x)$ [19], у которых значение диэлектрической проницаемости может достигать на порядок большей величины. В промышленных изделиях минимальная толщина диэлектрика (SiO_2) может составлять около 10 нм [14–16]. В лабораторных образцах МДПТ с субмикронным каналом эффективная толщина диэлектрика составляет менее 1 нм. При этом рабочие напряжения не превышают 1,5–2 В.

Ширина ОПЗ l уменьшается при повышении степени легирования подложки. Для n -канальных МДПТ приемлемой является концентрация акцепторов в подложке $N_B \approx 10^{17} \text{ см}^{-3}$, позволяющая получить ширину ОПЗ $l \approx 0,1$ мкм. Для комплементарных транзисторных пар необходимость создания n -кармана (см. рис. III.2.3) вынуждает снижать концентрацию акцепторов в подложке до величины $N_B \approx (2 - 5) \cdot 10^{15} \text{ см}^{-3}$, в результате чего ширина ОПЗ в n -канальном МДПТ увели-

чивается в 5–7 раз. Этого можно избежать при использовании слаболегированной подложки p^- - или n^- -типа ($N \approx 3 \cdot 10^{15} \text{ см}^{-3}$) (рис. III.8.2), в которой для n - и p -канальных МДПТ формируются p - и n -карманы 10 и 11, соответственно, с поверхностной концентрацией примеси $N_B \approx 10^{17} \text{ см}^{-3}$. Взаимная изоляция транзисторов осуществляется глубокими разделительными областями оксида 12, под которыми поверхность подложки сильно легирована акцепторами, чтобы исключить образование паразитных каналов из-за положительного заряда в оксиде (противоканальные p^+ -области 13). Однотипные p -канальные транзисторы, расположенные в одном кармане, также изолируются друг от друга глубокими разделительными областями оксида.

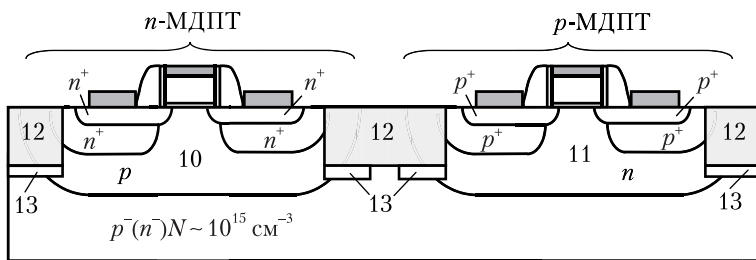


Рис. III.8.2. Схематическое изображение типовой структуры комплементарной пары МДПТ с p - и n -карманами:

10, 11 — карманы ($N_B \approx 10^{17} \text{ см}^{-3}$); 12 — разделительный оксид SiO_2 ; 13 — противоканальные p^+ -области

5. Малая величина сопротивления затвора обеспечивается использованием в качестве материала затвора сильнолегированного поликремния 7 (n^+ - или p^+ - Si^*), на который сверху нанесен металлический слой 8. Ток в электроде затвора (ток смещения через емкости затвор—исток, затвор—сток) протекает в поперечном направлении (по отношению к направлению тока в канале). Как отмечено в параграфе III.6.3, транзисторы с большой шириной канала Z выполняются в виде параллельно соединенных секций.

6. Напряжение смыкания повышается при уменьшении толщины n^+ -областей стока и истока x_j , примыкающих к каналу (см. параграф III.5.3, раздел III.7). Это достигается использованием мелких расширенных n^+ -областей стока и истока 2. На рис. III.8.1 показана форма ОПЗ при напряжении сток—исток $V_{ds} = 0$ (а) и напряжении $V_{ds} \approx V_{pinch}$, близком к напряжению смыкания (б). Наличие мелких расширенных

контактных n^+ -областей обеспечивает существенное повышение напряжения смыкания V_{pinch} .

III.8.3. Структурные модификации короткоканальных МДПТ

Основным недостатком описанной типовой структуры является повышение емкостей C_{bd} , C_{bs} p - n переходов между подложкой и контактными n^+ -областями стока и истока (см. рис. III.6.1), вызванное стремлением уменьшить ширину ОПЗ для соблюдения правил масштабирования и повышения напряжения смыкания. Компромиссное решение состоит в сильном легировании не всей подложки (или p - и n -карманов на рис. III.8.2), а только тонкого слоя под каналом, как это показано на рис. III.8.3, а (p^+ -слой 14). Этот слой соприкасается только с торцевыми областями контактных n^+ -областей, что создает возможность повысить степень его легирования до $\sim 10^{18} \text{ см}^{-3}$ при снижении емкостей C_{bd} , C_{bs} .

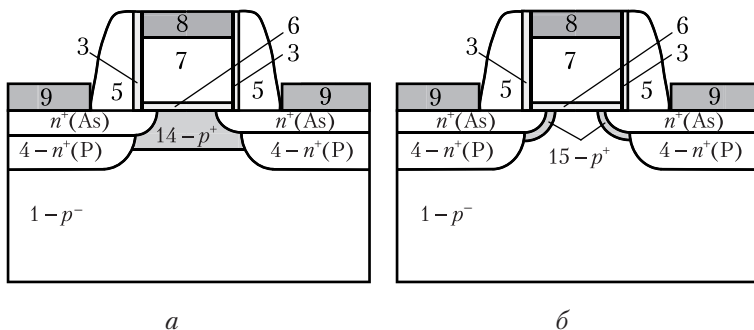


Рис. III.8.3. Структуры МДПТ с p^+ -слоем 14 (а) и с p^+ -гало-областями 15 (б).

Остальные обозначения соответствуют рис. III.8.1

Дальнейшее развитие этой идеи привело к замене p^+ -слоя тонкими цилиндрическим p^+ -областями, получившими название «гало» (области 15 на рис. III.8.3, б). Гало-области не только снижают площадь контактных p - n переходов между подложкой и контактными n^+ -областями стока и истока, но исключают большую часть емкости между подложкой и каналом, которая в эквивалентной схеме на рис. III.6.2 включена в емкости C_{bd} , C_{bs} . Формирование гало-областей основано на различии коэффициентов диффузии примесей и произво-

дится одновременно с формированием мелких контактных областей. Этот прием позволяет уменьшить длину канала до размеров, находящихся за пределами возможностей фотолитографии.

Аналогично выглядят структуры p -канальных МДПТ, для которых подложкой являются дополнительно сформированные n -карманы.

Помимо высоких характеристик транзистора важное значение имеет плотность компоновки элементов БИС. В комплементарной структуре на рис. III.8.2 каждый транзистор занимает на кристалле сравнительно большую площадь, что обусловлено несколькими причинами. Дополнительную площадь занимают карманы с омическими контактами к ним. При формировании кармана ионы примеси диффундируют в боковом направлении примерно на такое же расстояние, как в глубину полупроводника, что также увеличивает размеры структуры. Расстояние между p - n переходом карман—подложка и p - n переходом исток(сток)—подложка ближайшего n -канального транзистора должно превышать суммарную ширину ОПЗ этих переходов, чтобы исключить замыкание n - и n^+ -областей. Так как подложка легирована слабо, толщина ОПЗ достаточно велика (≈ 1 — 2 мкм). Кроме того, n -канальные транзисторы должны быть достаточно удалены от края кармана во избежание тиристорного эффекта защелки (см. параграф IV.6.5).

Чрезвычайно высокая плотность компоновки и минимальные паразитные емкости достигаются в структурах «кремний на изоляторе» (КНИ или SOI — silicon on insulator). Структура и энергетическая диаграмма такого транзистора показаны на рис. III.8.4. Технология изготовления КНИ описана в работах [20, 21] и др. Важным преимуществом структур КНИ является наличие потенциального барьера (диэлектрика) между активным полупроводниковым слоем (Si на рис. III.8.4, а) и подложкой. При малой толщине активного слоя он полностью перекрыт областью пространственного заряда. Если длина канала достаточно мала, температура электронов при рабочих напряжениях сток—исток может превышать 3000 К ($\approx 0,3$ эВ). В этих условиях толщина канала h (рис. III.8.4, б) оказывается сравнимой с толщиной диэлектрика. Таким образом, *эффективная толщина диэлектрика увеличивается*, что приводит к снижению крутизны ВАХ и коэффициента усиления по напряжению (см. параграф III.5.7). Диэлектрический слой нейтрализует отмеченный эффект.

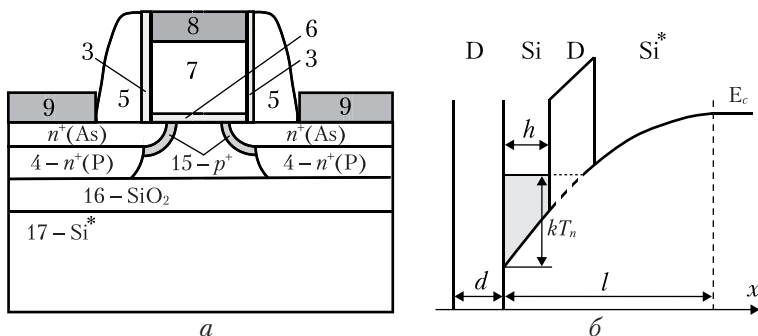


Рис. III.8.4. МДПТ «кремний на изоляторе» (КНИ):

a — структура; *б* — энергетическая диаграмма;

16 — изолятор (SiO_2); 17 — подложка (Si^*).

Остальные обозначения соответствуют рис. III.8.1, III.8.3, 6

На рис. III.8.5 представлена структура комплементарной транзисторной пары на сапфировой подложке (кремний на сапфире — КНС). Полупроводниковые слои толщиной $\approx 0,5$ мкм формируются методами гетероэпитаксии, а изолирующие слои SiO_2 — сквозным локальным окислением. Благодаря диэлектрической подложке паразитные емкости p - n переходов образуются только с торцев стока и истока. Это позволяет выбирать достаточно большие концентрации примеси в островках полупроводника между стоком и истоком, что обеспечивает высокое напряжение смыкания. Пороговое напряжение необходимой величины обеспечивается подлгированием поверхности.

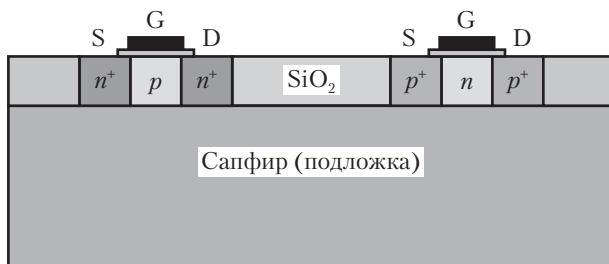


Рис. III.8.5. Комплементарная транзисторная пара на сапфировой подложке

Значительного повышения плотности компоновки можно добиться за счет расположения областей стока и истока в разных плоскостях — одного под другим. В таких структурах канал проходит в вертикальном направлении, поэтому они называются вертикальными. Вертикальные структуры являются одной из разновидностей так называемых трехмерных интегральных схем (ТИС) — одним из перспективных направлений микроэлектроники [22].

Современные субмикронные МДПТ имеют предельную частоту до 50—25 ГГц (n - и p -канал, соответственно). В лабораторных образцах достигнуты значения f_T более 150 ГГц (при рабочих напряжениях $< 1,5$ В).

Литература

1. Lilienfeld, J. E. Пат. США № 1745175 / J. E. Lilienfeld, 1930.
2. Heil, O. Пат. Великобрит. 439457 / O. Heil, 1935.
3. Shockley, W. Modulation on conductance of thin films of semiconductors by surface charge / W. Shockley, J. L. Pearson // Phys. Rev. — 1948. — Vol. 74. — P. 232.
4. Kahnd, D. Silicon — Silicon Dioxide field induced surface devices / D. Kahnd, M. M. Atalla // IRE Solid-St. Device Res. Conf., Carnegie Inst. Of Technology. Pittsburgh, Pa. — 1960.
5. Kahnd, D. A historical perspective on the development of MOS transistors and related devices / D. Kahnd // IEEE Trans. — 1976. Vol. ED-23. — P. 655.
6. Ihantola, H. K. Design theory of a surface field-effect transistors / H. K. Ihantola // Solid State Electron. — 1964. — Vol. 7. — P. 423.
7. Sah, C. T. Characteristics of the metal-oxide-semiconductor transistors / C. T. Sah // IEEE Trans. — 1964. — Vol. ED-11. — P. 324.
8. Shockley, W. An unipolar field-effect transistor / W. Shockley Proc. IRE. — 1952. — Vol. 40. — P. 1365.
9. Зи, С. М. Физика полупроводниковых приборов. Кн. 1 / С. М. Зи ; пер. с англ. — М. : Мир, 1984. — 456 с.
10. Зи, С. М. Физика полупроводниковых приборов. Кн. 2 / С. М. Зи ; пер. с англ. — М. : Мир, 1984. — 456 с.
11. Brews, J. R. Generalized guide for MOSFET miniaturisation / J. R. Brews [and other] // IEEE EDL. — 1980. — Vol. EDL-1. — P. 2.
12. Ван дер Зил, А. Шум / А. Ван дер Зил ; пер. с англ. — М. : Советское Радио, 1973. — 228 с.
13. Gupta, M. S. Selected papers on noise in circuits and systems / M. S. Gupta // IEEE Press, 1998.



14. *Красников, Г. Я.* Конструктивно-технологические особенности субмикронных МОП-транзисторов. Ч. 1 / Г. Я. Красников. — М. : Техносфера, 2002. — 416 с.
15. *Красников, Г. Я.* Конструктивно-технологические особенности субмикронных МОП-транзисторов. Ч. 2 / Г. Я. Красников. — М. : Техносфера, 2004. — 536 с.
16. *Sze, S. M.* Physics of Semiconductor Devices (3-rd edition) / S. M. Sze, K. Ng. Kwok. — Wiley-Interscience, 2007. — 815 p.
17. *Schimtz, J.* Ultra-shallow junction formation by outdiffusion from implanted oxide / J. Schimtz // IEEE IEDM 1998. — P. 1009–1012.
18. *Tasch, A. F.* Limitation of LDD types of structure in deep-submicrometer MOS technology / A. F. Tasch [and other] // IEEE EDL. — 1990. — Vol. 11. — N₀ 11. — P. 517–519.
19. *Gusev, E. P.* Ultrathin high-K gate stacks for advanced CMOS devices / E. P. Gusev [and other] // IEEE IEDM. — 2001. — Session 20.
20. *Hisamoto, D.* FD/GD SOI MOSFET/ D. Hisamoto // IEEE IEDM. — 2001. — Session 20.
21. *Matsumoto, T.* 70 nm SOI-CMOS of 135 GHz $f_{\max}$ with dual offset-implanted source-drain extension structure for RF/analog and logic application / T. Matsumoto [and other] // IEEE IEDM. — 2001. — Session 10.
22. Закон Мура и процветание человечества // Услуги и цены. — 2005. — № 22(193). — С. 8–12.



Глава IV

БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

Биполярный транзистор (transistor — transfer resistor — изменяемое сопротивление) является одним из важнейших активных полупроводниковых приборов. Транзистор изобретен в 1947 г. группой исследователей (Бардин, Браттейн, Шокли) фирмы «Bell Laboratories» (США) [1], удостоенных за это Нобелевской премии по физике. В 1949 г. опубликована классическая работа У. Шокли по теории транзисторов [2]. Непрерывное совершенствование технологии и параметров биполярных транзисторов привело к созданию на их основе первых интегральных схем. В настоящее время биполярные транзисторы наряду с МДПТ являются основными активными приборами микроэлектроники.

IV.1. Устройство и принцип действия биполярного транзистора

IV.1.1. Устройство биполярного транзистора

Устройство биполярного транзистора (БТ) схематически показано на рис. IV.1.1, *а*. Транзистор представляет собой трехслойную полупроводниковую структуру с чередующимся типом проводимости слоев. На рис. IV.1.1, *а* показан *n-p-n* транзистор. В *p-n-p* транзисторе слои имеют обратный тип проводимости. В дальнейшем для определенности мы будем рассматривать *n-p-n* транзисторы.

В структуре транзистора могут быть выделены следующие области:

- *база* — промежуточный полупроводниковый слой (*p*-типа на рис. IV.1.1, *а*);
- *эмиттер и коллектор* — крайние полупроводниковые слои (*n*-типа на рис. IV.1.1, *а*);



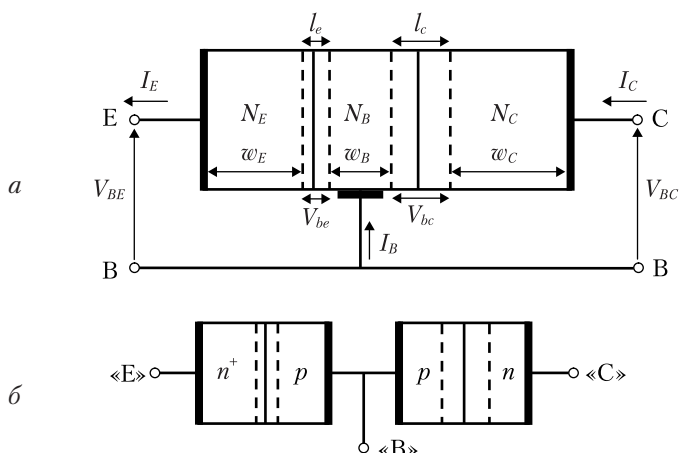


Рис. IV.1.1. Устройство биполярного транзистора (а) и попытка его имитации двумя встречно включенными диодами (б)

- эмиттерный и коллекторный p - n переходы шириной l_e и l_c ;
- омические контакты эмиттера, базы и коллектора с соответствующими электродами E (emitter), B (base), C (collector).

База, эмиттер и коллектор представляют собой электро-нейтральные области шириной w_B , w_E и w_C со средними результирующими концентрациями примеси N_B , N_E и N_C , соответственно. Эмиттер всегда легирован значительно сильнее, чем база (обычно в 50—1000 раз); коллектор может быть легирован сильнее или слабее, чем база.

Эмиттерный и коллекторный p - n переходы есть области пространственного заряда (ОПЗ), обедненные носителями заряда. Обычно площадь эмиттерного перехода меньше, чем коллекторного.

Из рис. IV.1.1, а очевидно, что биполярный транзистор, как и МДПТ, является *обратимым* прибором — эмиттер и коллектор можно поменять местами, сохранив принципиальную работоспособность прибора. Однако в связи с несимметричностью реальных транзисторов (степень легирования, площадь p - n переходов) при такой инверсии усилительные свойства транзистора резко ухудшаются.

Электрическое состояние транзистора определяется тремя напряжениями и тремя токами:

$$V_{BE} = V_B - V_E; \quad V_{BC} = V_B - V_C; \quad V_{CE} = V_C - V_E; \quad I_E; \quad I_C; \quad I_B,$$

где V_E , V_C и V_B — потенциалы эмиттера, базы и коллектора.

Напряжения и токи связаны уравнениями Кирхгофа:

$$V_{CE} = V_{BE} - V_{BC}; \quad I_E = I_C + I_B,$$

где положительные направления токов и напряжений указаны на рис. IV.1.1, а.

Для p - n - p транзистора положительные направления токов противоположны, а порядок индексов у напряжений инвертирован (например, $V_{EB} = V_E - V_B$). ВАХ транзистора описываются уравнениями, связывающими токи с напряжениями,

$$I_E(V_{BE}, V_{BC}); \quad I_C(V_{BE}, V_{BC}).$$

Таким образом, электрическое состояние транзистора определяется двумя независимыми переменными.

Заметим сразу, что *внешние* напряжения V_{BE} и V_{BC} между электрическими выводами транзистора отличаются от *внутренних* напряжений V_{be} и V_{bc} на эмиттерном и коллекторном переходах вследствие падений напряжений на сопротивлениях электронейтральных областей эмиттера, коллектора и базы от токов I_E , I_C и I_B .

IV.1.2. Принцип действия и режимы работы транзистора

Усилительные свойства транзистора наилучшим образом проявляются в *нормальном* режиме работы, когда эмиттерный переход открыт: $V_{be} > 0$, а коллекторный переход закрыт или находится под нулевым напряжением: $V_{bc} \leq 0$ (для p - n - p транзистора $V_{eb} > 0$, $V_{cb} \leq 0$). При этом в базу *инжектируются* из эмиттера неосновные носители заряда (электроны для p - n - p транзистора). Толщина базы w_B (см. рис. IV.1.1, а) должна быть достаточно мала по сравнению с диффузионной длиной неосновных носителей в базе L_B . При этом большая часть инжектированных в базу электронов достигает закрытого коллекторного перехода и *экстрагируется* (поглощается) коллектором. Названия областей эмиттера и коллектора отражают процессы эмиссии и собирания неосновных для базы носителей заряда. Название базовой области связано с тем, что в первых транзисторных структурах базовая область составляла первоначальную основу (базу) при изготовлении прибора.

В «хорошем» транзисторе обязательным является выполнение условий

$$w_B / L_B \ll 1; \quad (IV.1.1)$$

$$N_E / N_B \gg 1. \quad (IV.1.2)$$

В параграфе II.4.8 были введены понятия коэффициентов инжекции электронов γ_n и дырок γ_p , а также эффективности эмиттера γ , соответствующей коэффициенту инжекции тех носителей заряда, которые являются в эмиттере основными (для n - p - n транзистора $\gamma = \gamma_n$). Кроме того, было показано, что при условии (IV.1.2) эффективность эмиттера близка к 1 ($\gamma \approx 1$). Таким образом, при условии (IV.1.2) почти весь ток через эмиттерный переход переносится электронами ($I_{En} \approx I_E$).

В нормальном режиме работы ($V_{bc} \leq 0$) инжекция носителей заряда через коллекторный переход практически отсутствует, поэтому ток коллектора I_C соответствует потоку электронов, инжектированных в базу из эмиттера и экстрагированных коллектором. При условии (IV.1.1) число электронов, экстрагированных коллектором, много больше числа электронов, рекомбинирующих в базе. Поэтому $I_C \approx I_{En} \approx I_E$.

Таким образом, *в нормальном режиме работы ток коллектора с высокой точностью повторяет ток эмиттера.*

Принцип действия транзистора основан на том, что ток эмиттера зависит от напряжения на открытом эмиттерном переходе, поэтому его величиной можно управлять с помощью внешнего источника. Ток коллектора практически не зависит от напряжения на закрытом коллекторном переходе и может быть пропущен через достаточно большое внешнее сопротивление в цепи коллектора, что обеспечивает усиление по напряжению и по мощности.

В «плохом» транзисторе хотя бы одно из условий (IV.1.1), (IV.1.2) не выполнено. При этом ток коллектора не будет повторять с высокой точностью ток эмиттера, а при длинной базе ($w_B / L_B \gg 1$) вообще не будет от него зависеть (экстракция коллектором электронов, инжектированных из эмиттера, прекратится) и останется равным малому обратному току диода коллектор—база. Если, однако, длина базы хотя бы сравнима с диффузионной длиной неосновных носителей L_B , то экстракция электронов коллектором будут оставаться заметной и пропорциональность между токами коллектора и эмиттера сохранится. Этот эффект взаимодействия эмиттерного и



коллекторного p - n переходов и называется «транзисторным эффектом». Степень его проявления зависит от отношения w_B / L_B . В «хорошем» транзисторе он проявляется максимально сильно, а при условии $w_B / L_B \gg 1$ отсутствует. Поэтому попытка имитации транзистора двумя встречно включенными диодами (рис. IV.1.1, б) не может быть успешной — омические контакты к p -областям исключают взаимодействие p - n переходов и транзисторный эффект.

Основным параметром транзистора, характеризующим его усилительные свойства, является *коэффициент передачи эмиттерного тока*, который определяется как отношение тока коллектора к току эмиттера при нулевом напряжении база—коллектор

$$\alpha_N = I_C / I_E \big|_{V_{bc}=0}. \quad (\text{IV.1.3})$$

Отметим, что определение (IV.1.3) — лишь одно из возможных и впоследствии будет уточняться. В «хорошем» транзисторе обычно значение α_N составляет не менее 0,95.

Если входным электродом является база, а выходным — коллектор, то транзистор обеспечивает усиление тока. Действительно, при $V_{bc} = 0$ ток базы

$$I_B = I_E - I_C = I_C / \alpha_N - I_C = I_C(1 - \alpha_N) / \alpha_N \text{ и } I_C / I_B = \beta_N,$$

где

$$\beta_N = \alpha_N / (1 - \alpha_N) \quad (\text{IV.1.4})$$

— *коэффициент усиления базового тока*. Индекс N соответствует нормальному режиму работы транзистора. При $1 - \alpha_N \ll 1$ значение $\beta_N \gg 1$ и составляет обычно не менее 20.

В зависимости от полярности напряжений на p - n переходах возможны четыре режима работы транзистора:

- *нормальный режим* $V_{be} \geq 0; V_{bc} \leq 0;$
- *инверсный режим* $V_{be} < 0; V_{bc} > 0;$
- *режим насыщения* $V_{be} > 0; V_{bc} > 0;$
- *режим отсечки* $V_{be} < 0; V_{bc} < 0.$

В нормальном режиме, как отмечалось выше, усилительные свойства транзистора проявляются в наибольшей степени.

В инверсном режиме условие (IV.1.1) в большинстве случаев не выполнено. Кроме того, площадь коллекторного перехода S_C обычно заметно превышает площадь эмиттерного перехода S_E и часть неосновных носителей, инжектированных из эмиттера в базу, не может достигнуть коллектора. Поэтому *инверсный коэффициент передачи тока* существенно меньше 1:



в зависимости от конструкции транзистора $\alpha_I \approx 0,1 \div 0,8$, и значения *инверсного коэффициента усиления базового тока* лежат в диапазоне $\beta_I \approx 0,1 \div 4$. Заметим, что в соответствии с (IV.1.4) $\beta_I < 1$ при $\alpha_I < 0,5$, и $\beta_I \approx \alpha_I$ при $\alpha_I \ll 1$.

В режиме насыщения оба *p-n* перехода открыты и потенциалы всех электродов имеют близкие значения — транзистор «стягивается в потенциальную точку». В режиме отсечки оба *p-n* перехода закрыты и между электродами транзистора протекают только малые обратные токи.

IV.1.3. Разновидности биполярных транзисторов

Классификацию транзисторов можно провести по нескольким признакам.

С точки зрения *типа проводимости* существуют *n-p-n* и *p-n-p* транзисторы.

По характеру *распределения примеси в базе* N_B различают *дрейфовые* и *бездрейфовые* транзисторы. Если базовая область транзистора легирована однородно, в равновесном состоянии электрическое поле в базе отсутствует, и при достаточно малых токах основным механизмом переноса неосновных носителей заряда через базу является диффузия. При неоднородном легировании базовой области даже в равновесном состоянии в базе существует электрическое поле, которое компенсирует диффузионные токи носителей заряда, вызванные градиентами их концентраций. Это поле может оказывать существенное влияние на перенос неосновных носителей через базу, и такие транзисторы называют дрейфовыми.

В дискретных транзисторах, предназначенных для монтажа в корпусе, контакты к областям эмиттера, базы и коллектора расположены в разных плоскостях полупроводникового кристалла (как это показано на схематичном рис. IV.1.1, а). Такие конструкции транзисторов называются *непланарными*. В интегральных схемах все контакты транзистора должны быть расположены на верхней плоскости кристалла. Такие конструкции транзисторов называются *планарными*.

Можно также классифицировать транзисторы *по способу изготовления*. Промышленностью было сперва освоено производство *точечных* транзисторов, в которых эмиттер и коллектор выполнены в виде прижимного контакта тонкого металлического острия к полупроводниковому кристаллу базы. Точечные транзисторы имели чрезвычайно низкую воспроизводимость характеристик, и их выпуск давно прекращен.

С начала 50-х гг. XX в. широкое распространение получили *сплавные* германиевые, а затем и кремниевые транзисторы.

Структура сплавного германиевого $p-n-p$ транзистора представлена на рис. IV.1.2, *а*. Области эмиттера и коллектора создаются с помощью сплавления металла III группы (например, In) в шайбу из полупроводника n -типа (базу). После рекристаллизации полупроводника на границе с металлом возникают слои p -типа, легированные акцепторами (In), — эмиттер и коллектор. Площадь эмиттерного перехода обычно делается меньше, чем коллекторного, чтобы исключить инжекцию неосновных носителей из эмиттера в периферийные области базы. Длина базы w_B зависит от толщины исходной полупроводниковой пластины и не может быть сделана менее 20 мкм. Для получения достаточно высокого коэффициента передачи эмиттерного тока необходимо иметь в базе большую диффузионную длину неосновных носителей (в данном случае дырок) $L_B = \sqrt{D_B \tau_B}$. Поэтому область базы легируется слабо ($N_B \approx 10^{15} \text{ см}^{-3}$), чтобы увеличить время жизни дырок τ_B . Сплавные транзисторы являются бездрейфовыми.

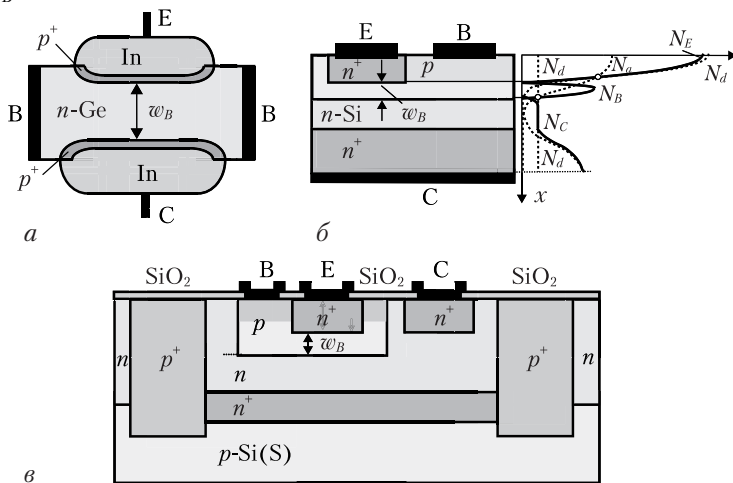


Рис. IV.1.2. Разновидности транзисторов:

- а* — сплавной бездрейфовый $p-n-p$: $w_B \approx 20 \text{ мкм}$, $N_B \approx 10^{15} \text{ см}^{-3}$, $N_E = N_C \approx 10^{18} \text{ см}^{-3}$; *б* — диффузионный дрейфовый $n-p-n$: $w_B \approx 1 \text{ мкм}$, $N_B \approx 10^{17} \text{ см}^{-3}$, $N_E > 10^{19} \text{ см}^{-3}$, $N_C \approx 10^{16} \text{ см}^{-3}$, $N_C^+ \approx 10^{18} \text{ см}^{-3}$; *в* — планарный дрейфовый $n-p-n$ с разделительной диффузией: $w_B \approx 0,2 \text{ мкм}$, $N_B \approx 3 \cdot 10^{17} \text{ см}^{-3}$, $N_E \approx 10^{20} \text{ см}^{-3}$, $N_C \approx 10^{16} \text{ см}^{-3}$, $N_C^+ \approx 10^{19} \text{ см}^{-3}$; $N_S \approx 10^{15} \text{ см}^{-3}$

Освоение диффузионной технологии привело к созданию *диффузионных* дрейфовых транзисторов (рис. IV.1.2, б). В таких транзисторах изменение типа проводимости полупроводниковых слоев осуществляется путем перекомпенсации исходного слаболегированного материала — n -коллектора (в n - p - n транзисторе). При этом концентрация акцепторов в базе уменьшается по направлению от эмиттера к коллектору. Для надежной перекомпенсации донорной примеси в области базы концентрация акцепторов должна быть достаточно велика, в результате чего время жизни и диффузионная длина неосновных носителей в базе значительно меньше, чем в сплавном транзисторе. Длина базы может быть уменьшена до ≈ 1 мкм. Область эмиттера формируется путем сплавления металла либо путем диффузии донорной примеси с предельно высокой концентрацией. Для получения хорошего омического контакта коллектора и снижения последовательного сопротивления тела коллектора тыльная сторона кристалла легируется дополнительно донорами (n^+ -коллектор). Диффузионные транзисторы первоначально изготавливались на основе германия; позднее была освоена кремниевая технология.

Современная кремниевая интегральная технология располагает такими методами формирования транзисторных структур, как эпитаксия, ионная имплантация, селективное анизотропное травление, молекулярно-лучевая эпитаксия, плазмохимические методы, самосовмещенная литография. Способы изготовления планарных интегральных кремниевых транзисторов и их конструкции непрерывно совершенствуются. В биполярных интегральных схемах используются в основном n - p - n транзисторы. Создание p - n - p транзисторов тоже возможно, однако они имеют значительно худшие параметры. В качестве примера на рис. IV.1.2, в показана структура *планарного* дрейфового n - p - n транзистора с *разделительной диффузией*. Эта структура не является оптимальной, однако она наиболее проста; далее мы будем использовать ее для иллюстраций. Более совершенные структуры биполярных транзисторов в ИС рассмотрены в разделе IV.8.

В транзисторной структуре на рис. IV.1.2, в подложкой (substrate — S) является слаболегированная пластина p -Si. Коллекторный слой n -типа создается методом эпитаксиального наращивания. База p -типа и эмиттерный слой n^+ -типа формируются методом селективной диффузии или селективной ионной имплантации. Концентрация акцепторов в базе уменьшается по направлению от эмиттера к коллектору,

поэтому планарный транзистор является дрейфовым. Длина базы обычно составляет $0,1-0,3$ мкм. Одновременно с эмиттером формируется контактная область n^+ -типа в коллекторе. Взаимная изоляция транзисторов в ИС осуществляется с помощью глубокой разделительной диффузии акцепторов с высокой концентрацией (p^+ -области). Для снижения сопротивления тела коллектора перед эпитаксиальным наращиванием n -слоя в подложке может быть сформирован «скрытый» n^+ -слой. Для его создания применяются донорные примеси с малым коэффициентом диффузии (например, As), чтобы предотвратить распыление n^+ -слоя при последующих высокотемпературных операциях. Акцепторной примесью в p -базе является обычно бор (В), а донорной примесью в эмиттере — фосфор (Р) или мышьяк (As).

Значения основных параметров типовых транзисторных структур приведены в табл. IV.1.1.

Таблица IV.1.1

**Значения основных параметров типовых
биполярных транзисторных структур**

Тип тран- зистора	Средняя концентрация примеси, см ⁻³			Толщина слоя, мкм		
	Эмиттер	База	Коллектор	Эмиттер	База	Коллектор
Сплавной бездрейфо- вый	$(1-5) \cdot 10^{18}$	$(0,5-2) \cdot 10^{15}$	$(1-5) \cdot 10^{18}$	10–50	20–30	10–50
Диффу- зионный дрейфо- вый	$(1-3) \cdot 10^{19}$	$(0,5-2) \cdot 10^{17}$	$n^- \approx 10^{16};$ $n^+ \approx 10^{18}$	0,5–1,5	0,5–1,5	$n^- - 10-20;$ $n^+ - 50-200$
Планар- ный дрей- фовый	$(1-8) \cdot 10^{20}$	$(1-5) \cdot 10^{18}$	$n^- \approx 10^{16};$ $n^+ \approx 10^{19}$	0,2–0,5	0,1–0,3	$n^- - 0,5-2;$ $n^+ - 0,5-1$

Интересно, что во всех рассмотренных типах транзисторов коэффициенты усиления базового тока примерно одинаковы ($\beta_N \approx 20-100$), несмотря на то, что длина базы в сплавных и планарных транзисторах различается примерно на два порядка. Причина состоит в различной степени выполнения условий (IV.1.1) и (IV.1.2). В сплавных транзисторах, имеющих сравнительно длинную базу, коэффициент усиления определяется степенью выполнения условия (IV.1.1). В диффузионных и планарных транзисторах условие (IV.1.1) выполнено значительно сильнее, чем (IV.1.2). Кроме того, для предельно



высокой концентрации примеси в эмиттере (вырожденный эмиттер) степень легирования эмиттера значительно слабее влияет на коэффициент усиления β_N . Этот вопрос подробно обсуждается в параграфе IV.3.2.

IV.1.4. Основные физические процессы в транзисторе (нормальный режим)

Физические процессы, происходящие в транзисторе в нормальном режиме работы ($V_{be} > 0$, $V_{bc} = 0$), иллюстрируются рис. IV.1.3, на котором представлены распределения неосновных носителей заряда в электронейтральных областях эмиттера, базы и коллектора, распределение токов и энергетическая диаграмма.

Концентрации неосновных носителей на границах p - n переходов определяются приложенными к переходам напряжениями (см. параграф I.3.4): на границах эмиттерного перехода ($V_{be} > 0$) $p_n > p_{n0}$, $n_p > n_{p0}$; на границах коллекторного перехода ($V_{bc} = 0$) $p_n = p_{n0}$, $n_p = n_{p0}$. Будем считать, что области базы, эмиттера и коллектора легированы однородно. В этом случае распределение неосновных носителей в базе при условии (IV.1.1) близко к линейному, как и в диоде с короткой базой (соотношение (I.5.166), параграф I.5.3). Токи неосновных носителей через p - n переходы имеют диффузионный механизм (см. параграф I.5.1) и пропорциональны градиентам концентрации неосновных носителей на границах с ОПЗ. Таким образом, величина этих токов соответствует тангенсам углов наклона функций $p_n(x)$ и $n_p(x)$ (т.е. производным dp_n/dx и dn_p/dx) на границах с ОПЗ (рис. IV.1.3, а). На границах базы с p - n переходами электронная составляющая коллекторного тока I_{Cn} несколько меньше, чем эмиттерного тока I_{En} , так как небольшая часть электронов, инжектированных из эмиттера, рекомбинирует в базе, не успев достигнуть коллекторного перехода. Поэтому функция dn_p/dx в базе имеет слабовогнутый характер.

Распределение токов в транзисторе схематично иллюстрируется рис. IV.1.3, б. Величина каждой составляющей тока представлена шириной соответствующей «ленты». Течения токов показаны стрелками, направление которых соответствует потокам дырок и противоположно потокам электронов. Носители заряда, создающие каждую составляющую тока, также схематично показаны на энергетической диаграмме (рис. IV.1.3, в).

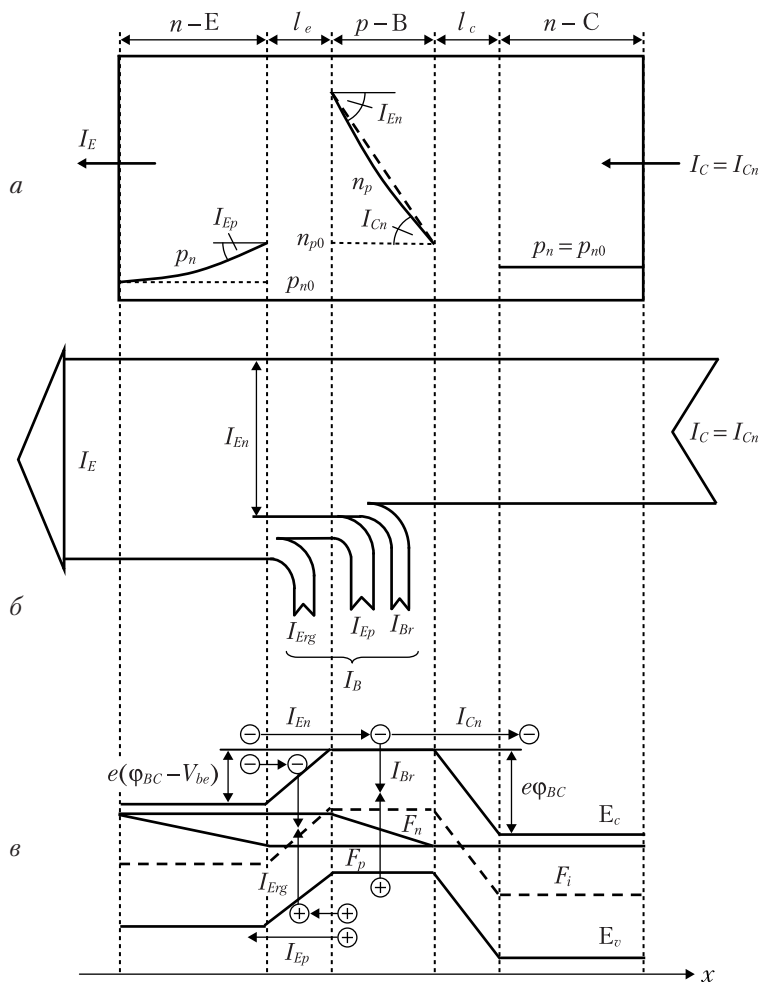


Рис. IV.1.3. Нормальный режим работы транзистора:

a — распределения неосновных носителей заряда в электронейтральных областях; *б* — распределение токов; *в* — энергетическая диаграмма

Ток эмиттера является суммой трех составляющих:

$$I_E = I_{En} + I_{Ep} + I_{Erg},$$

где I_{En} и I_{Ep} — токи инжекции электронов и дырок через эмиттерный переход, а I_{Erg} — ток, связанный с рекомбинацией и термогенерацией носителей заряда в эмиттерном переходе

(см. параграф I.2.2). Отметим, что из трех составляющих эмиттерного тока лишь одна, а именно I_{En} , связана с транзисторным эффектом. Две другие — I_{Ep} и I_{Erg} — замыкаются по цепи эмиттер—база (рис. IV.1.3, б, в) и не отражаются на величине коллекторного тока. Поскольку напряжение база—коллектор в нашем примере равно нулю, инжекция электронов из коллектора в базу и дырок из базы в коллектор отсутствует. Поэтому при $V_{bc} = 0$ ток коллектора имеет только электронную составляющую

$$I_C = I_{Cn}.$$

Этот ток равен электронной составляющей тока эмиттера I_{En} за вычетом тока I_{Br} , обусловленного рекомбинацией электронно-дырочных пар в электронейтральной базе. Некоторая часть электронов, инжектированных в базу из эмиттера, успевает рекомбинировать с основными носителями — дырками. Рекомбинация дырок в базе восполняется соответствующим потоком дырок из базового контакта. Поэтому ток I_{Br} , как и токи I_{Ep} , I_{Erg} , замыкается по цепи эмиттер—база (см. рис. IV.1.3, б, в) и не попадает в цепь коллектора.

Таким образом, при $V_{be} > 0$, $V_{bc} = 0$ ток коллектора составляет

$$I_C = I_{Cn} = I_E - I_{Ep} - I_{Erg} - I_{Br} \quad (\text{IV.1.5})$$

(см. рис. IV.1.3, б, в). При запираании коллекторного перехода ($V_{bc} < 0$) к коллекторному току добавляется только весьма малый ток, связанный с экстракцией неосновных носителей. Отсюда следует, что в нормальном режиме работы ($V_{be} \geq 0$, $V_{bc} \leq 0$) ток коллектора зависит только от напряжения на эмиттерном переходе.

Физическая причина этого факта состоит в том, что для электронов, инжектируемых из эмиттера в базу, эмиттерный переход с контактной разностью потенциалов ϕ_{BE} является потенциальным барьером (см. рис. IV.1.3, б, в), высота которого $e(\phi_{BE} - V_{be})$ снижается с ростом напряжения V_{be} , а коллекторный переход — потенциальной ямой, глубина которой $e(\phi_{BC} - V_{bc})$ хотя и зависит от напряжения V_{bc} , но не влияет на экстракцию переходом носителей заряда.

На энергетической диаграмме (см. рис. IV.1.3, в) показано расположение квазиуровней Ферми F_n и F_p в транзисторе при $V_{be} > 0$, $V_{bc} = 0$. В коллекторе и в коллекторном переходе концентрации электронов и дырок имеют равновесные значения ($V_{bc} = 0$), поэтому $F_n = F_p$. В эмиттере и в базе кон-

центрации неосновных носителей повышены вследствие инжекции электронов в базу из эмиттера, а дырок — в эмиттер из базы ($V_{be} > 0$). Ввиду электронейтральности эмиттера и базы настолько же повышены концентрации основных носителей, поэтому $F_n > F_p$. В эмиттерном переходе квазиуровни Ферми почти параллельны (см. параграф I.3.2) и сливаются только на омическом контакте к эмиттеру (см. параграф I.5.3).

IV.1.5. Эффективность эмиттера и коэффициент переноса

Из определения (IV.1.3) с учетом соотношения (IV.1.5) следует, что коэффициент передачи эмиттерного тока можно представить в виде

$$\alpha_N = \frac{I_C}{I_E} \bigg|_{V_{bc}=0} = \frac{I_C}{I_{En}} \cdot \frac{I_{En}}{I_E} \bigg|_{V_{bc}=0},$$

или $\alpha_N = \kappa_N \gamma_N,$ (IV.1.6)

где $\kappa_N = I_C / I_{En} \big|_{V_{bc}=0}$ (IV.1.7)

— коэффициент переноса,

$$\gamma_N = \frac{I_{En}}{I_E} \bigg|_{V_{bc}=0} = \frac{1}{1 + (I_{Ep} + I_{Erg}) / I_{En}} \bigg|_{V_{bc}=0} \quad (IV.1.8)$$

— *эффективность эмиттера* — доля тока неосновных для базы носителей заряда в полном токе через эмиттерный переход. В отличие от формул (II.4.41а, б) в (IV.1.8) понятие эффективности эмиттера включает ток рекомбинации — генерации I_{Erg} в эмиттерном переходе, который в модели идеализированного диода не учитывался. В параграфе I.6.4 было показано, что вклад тока I_{Erg} может быть существенным только в кремниевых приборах при малом токе.

Коэффициент переноса можно определить как *вероятность переноса через базу электронов, инжектированных через эмиттерный переход, до коллекторного перехода без рекомбинации*.

Из (IV.1.6) следует, что близкий к 1 коэффициент передачи эмиттерного тока α_N может быть получен только в случае, когда и коэффициент переноса, и эффективность эмиттера близки к 1. Первое условие обеспечивается неравенством (IV.1.1), второе — неравенством (IV.1.2).



Основные выводы

1. Биполярный транзистор представляет собой систему двух близко расположенных взаимодействующих p - n переходов. Транзисторный эффект состоит в собирании коллекторным переходом носителей, инжектированных в базу из эмиттера.

2. В зависимости от полярности напряжений на p - n переходах возможны четыре режима работы транзистора: нормальный режим, инверсный режим, режим насыщения и режим отсечки. Усилительные свойства транзистора в наибольшей степени проявляются в нормальном режиме: $V_{be} \geq 0$; $V_{bc} \leq 0$.

3. Транзисторы различаются по типу проводимости (n - p - n и p - n - p), по характеру легирования базы (бездрейфовые и дрейфовые), а также по конструкции (сплавные, диффузионные и планарные).

4. Основными параметрами транзистора, характеризующими его усилительные свойства, являются коэффициент передачи эмиттерного тока α_N и коэффициент усиления базового тока $\beta_N = \alpha_N / (1 - \alpha_N)$. Значение коэффициента α_N должно быть возможно ближе к 1.

5. Коэффициент передачи эмиттерного тока может быть представлен в виде произведения коэффициента переноса на эффективность эмиттера. Коэффициент переноса характеризует рекомбинационные потери при переносе неосновных носителей через базу. Эффективность эмиттера характеризует потери эмиттерного тока, связанные с инжекцией неосновных носителей из базы в эмиттер, а также с рекомбинацией носителей в эмиттерном переходе.

6. Близкие к 1 значения коэффициента переноса и эффективности эмиттера обеспечиваются выполнением условий $w_B / L_B \ll 1$ и $N_E / N_B \gg 1$ соответственно.

IV.2. Модель идеализированного транзистора Эберса—Молла

IV.2.1. Модель Эберса—Молла

Дж. Эберсом и Дж. Моллом была впервые предложена модель электронного прибора, позволяющая получить его ВАХ в простой аналитической форме [3]. Интересно отметить, что в это время полупроводниковые приборы еще не занимали доминирующего положения в электронике и активная (с усилением мощности) обработка электрических сигналов осуществлялась в основном электровакуумными лампами. Их ВАХ представлялись исключительно в графической форме.



Таким образом, модель биполярного транзистора Эберса—Молла явилась первым шагом к аналитическому, а впоследствии компьютерному проектированию электронных схем без использования объемных справочников.

В отличие от теории идеализированного МДПТ Шокли, основанной на физическом подходе (см. раздел III.3), модель Эберса—Молла имеет феноменологический характер и использует как параметры нормальный и инверсный коэффициенты передачи тока, не прибегая к их определению физическими методами. Модель описывает произвольный режим работы транзистора, т.е. является нелинейной моделью для большого сигнала.

В модели Эберса—Молла используется большинство допущений классической теории Шокли идеализированного полупроводникового диода (см. параграф II.4.1): малость поперечных размеров по сравнению с продольными (перпендикулярными плоскостям p - n переходов), однородность легирования областей базы и эмиттера, низкий уровень инжекции, пренебрежение токами генерации—рекомбинации носителей заряда в p - n переходах и диффузионный механизм токов неосновных носителей. Дополнительными допущениями являются:

- эквипотенциальность базы в плоскостях p - n переходов;
- постоянство нормального и инверсного коэффициентов передачи тока α_N , α_I .

Первое из этих допущений сводит к одномерной задаче анализа в принципе неодномерного прибора (направление тока в базе перпендикулярно направлению переноса неосновных носителей через базу от эмиттера к коллектору). Это допущение предполагает, что напряжение на эмиттерном и коллекторном переходах постоянно по всей площади переходов и зависит только от внешних напряжений V_{BE} и V_{BC} между электродами транзистора (см. рис. IV.1.1).

Второе допущение игнорирует то обстоятельство, что при изменении напряжений V_{be} и V_{bc} на p - n переходах изменяется ширина p - n переходов, а следовательно, и длина электро нейтральных областей базы w_B , эмиттера w_E и коллектора w_C (см. рис. IV.1.1). Изменения длины базы w_B должны, безусловно, изменять значения коэффициентов передачи тока. Особенно существенно изменение длины базы w_B при изменении напряжения V_{bc} , так как в нормальном режиме коллекторный переход закрыт и напряжение $V_{bc} < 0$ может изменяться весьма значительно.



В основе модели Эберса—Молла лежит идея разложения токов через эмиттерный и коллекторный переходы на *инжектируемую* и *собираемую* составляющие:

$$I_E = I_1(V_{be}) - I_{Ecol}(V_{bc}); \quad (IV.2.1a)$$

$$I_C = -I_2(V_{bc}) + I_{Ccol}(V_{be}). \quad (IV.2.1б)$$

Здесь $I_1(V_{be})$ и $I_2(V_{bc})$ — инжектируемые составляющие токов эмиттера и коллектора, соответственно. Уравнения (IV.2.1a, б) несимметричны относительно знаков «+» и «-», что связано с несимметрией направлений эмиттерного и коллекторного токов, принятых за положительные (см. рис. IV.1.1). *Инжектируемые токи являются функциями напряжений на тех переходах, через которые происходит инжекция.* Инжектируемые токи имеют как электронные составляющие (инжекция электронов в базу из эмиттера и коллектора), так и дырочные (инжекция дырок в эмиттер и коллектор из базы). Собираемые токи являются чисто электронными и пропорциональны токам, инжектированным через противоположные переходы, с коэффициентами передачи токов α_I и α_N :

$$I_{Ccol}(V_{be}) = I_{Ccoln}(V_{be}) = \alpha_N I_1(V_{be}); \quad (IV.2.2a)$$

$$I_{Ecol}(V_{bc}) = I_{Ecoln}(V_{bc}) = \alpha_I I_2(V_{bc}). \quad (IV.2.2б)$$

Таким образом, *собираемые токи являются функциями напряжений на переходах, противоположных тем, через которые они протекают.*

Эта идея поясняется рис. IV.2.1, на котором изображены распределения равновесных и неравновесных неосновных носителей заряда в квазинейтральных областях базы, эмиттера и коллектора. Для простоты на рисунке не показаны области *p-n* переходов, а все квазинейтральные области считаются короткими, так что распределения носителей заряда практически линейны. Рисунок IV.2.1 соответствует режиму насыщения ($V_{be} > 0$, $V_{bc} > 0$), поэтому концентрации носителей на границах с эмиттерным и коллекторным переходами больше равновесных значений. На границах с омическими контактами к эмиттеру и коллектору концентрации неосновных носителей имеют равновесные значения p_{n0} . Электронные составляющие I_{En} и I_{Cn} токов через эмиттерный и коллекторный переходы имеют диффузионный механизм и пропорциональны градиентам концентрации электронов в

базе на границах с переходами (т.е. тангенсам углов наклона функции $n_p(x)$, отмеченных на рис. IV.2.1 двойными дугами). Дырочные составляющие I_{Ep} и I_{Cp} представлены на рис. IV.2.1 соответствующими углами наклона функций $p_{n1}(x)$ в эмиттере и $p_{n2}(x)$ в коллекторе.

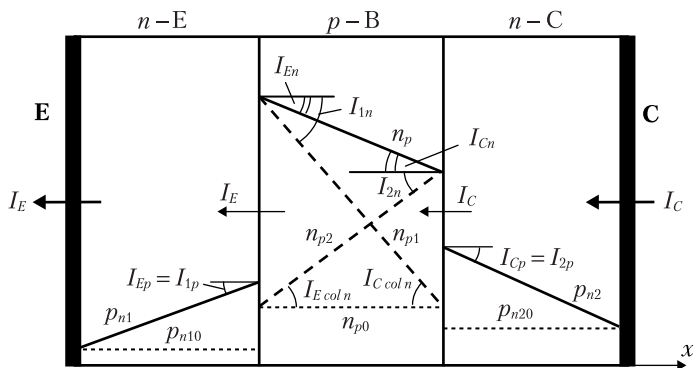


Рис. IV.2.1. Распределения равновесных и неравновесных неосновных носителей заряда в квазинейтральных областях базы, эмиттера и коллектора при $V_{be}, V_{bc} > 0$

Разложение токов I_E и I_C на инжектируемые и собираемые составляющие соответствует представлению о том, что для фиксированных напряжений V_{be} и V_{bc} распределение электронов в базе является суммой

$$n_p(x) = n_{p1}(x) + n_{p2}(x), \quad (IV.2.3)$$

где $n_{p1}(x)$ и $n_{p2}(x)$ — распределения электронов при $V_{bc} = 0$ и $V_{be} = 0$, соответственно. Функции $n_{p1}(x)$ и $n_{p2}(x)$ показаны на рис. IV.2.1 штриховыми линиями.

Как отмечено выше, собираемые составляющие токов коллектора и эмиттера являются чисто электронными: $I_{Ccol} = I_{Ccoln}$; $I_{Ecol} = I_{Ecoln}$. На рис. IV.2.1 эти токи представлены соответствующими углами наклона функций $n_{p1}(x)$ и $n_{p2}(x)$ в базе на границах с коллекторным и эмиттерным переходами.

Инжектируемые токи эмиттера и коллектора I_1 и I_2 имеют как электронные, так и дырочные составляющие. На рис. IV.2.1 электронные составляющие I_{1n} и I_{2n} представлены углами наклона функций $n_{p1}(x)$ и $n_{p2}(x)$ в базе на границах с эмиттерным и коллекторным переходами. Наконец, дырочные составляющие инжектируемых токов I_{1p} и I_{2p} совпадают с

дырочными составляющими полных токов эмиттера и коллектора I_{Ep} и I_{Cp} , о которых уже говорилось.

Заметим, что функции $n_p(x)$, $n_{p1}(x)$ и $n_{p2}(x)$ являются решениями биполярного уравнения непрерывности в p -базе, которое для принятого допущения о низком уровне инжекции имеет вид (I.7.86) (см. раздел I.7)

$$\frac{\partial n_p}{\partial t} = -\frac{\Delta n_p}{\tau_n} + \mu_n E \frac{\partial n_p}{\partial x} + D_n \frac{\partial^2 n_p}{\partial x^2}. \quad (\text{IV.2.4})$$

Соотношение (IV.2.3) справедливо только в том случае, когда в линейном дифференциальном уравнении (IV.2.4) коэффициенты τ_n , $\mu_n E$ и D_n постоянны (не зависят от n_p и координаты x). Это условие выполнено для принятых допущений о низком уровне инжекции (τ_n , μ_n и D_n постоянны) и диффузионном механизме токов неосновных носителей ($E = 0$). Таким образом, сама идея модели Эберса—Молла, несмотря на ее плодотворность, не является универсальной.

Из соотношений (IV.2.1а, б) и (IV.2.2а, б) следует

$$I_E = I_1(V_{be}) - \alpha_I I_2(V_{bc}); \quad (\text{IV.2.5a})$$

$$I_C = \alpha_N I_1(V_{be}) - I_2(V_{bc}). \quad (\text{IV.2.5б})$$

Напомним, что распределения электронов в базе $n_{p1}(x)$ и $n_{p2}(x)$ соответствуют условиям $V_{bc} = 0$ и $V_{be} = 0$, соответственно. При этом на границе с коллекторным переходом $n_{p1} = n_{p0}$, а на границе с эмиттерным переходом $n_{p2} = n_{p0}$ (см. рис. IV.2.1). Эти граничные условия в точности совпадают с граничными условиями на омических контактах в полупроводниковом диоде (см. параграф II.4.3). Допущения модели Эберса—Молла идентичны допущениям модели идеализированного диода, поэтому ВАХ инжектируемых токов $I_1(V_{be})$ и $I_2(V_{bc})$ должны совпадать с ВАХ идеализированного диода

$$I_1(V_{be}) = I_{1S}(e^{V_{be}/\Phi_T} - 1); \quad (\text{IV.2.6a})$$

$$I_2(V_{bc}) = I_{2S}(e^{V_{bc}/\Phi_T} - 1). \quad (\text{IV.2.6б})$$

Уравнениям (IV.2.5а, б) и (IV.2.6а, б) соответствует эквивалентная схема биполярного транзистора Эберса—Молла, представленная на рис. IV.2.2.

Уравнения (IV.2.5а, б) и (IV.2.6а, б) вместе с очевидным равенством

$$I_B = I_E - I_C \quad (\text{IV.2.7})$$



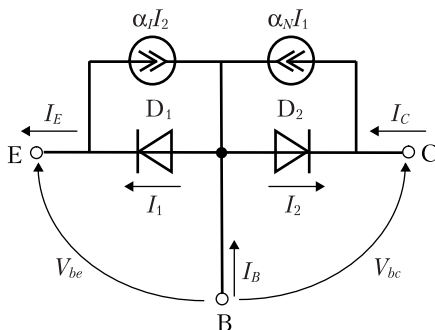


Рис. IV.2.2. Эквивалентная схема биполярного транзистора Эберса—Молла

полностью определяют ВАХ транзистора в рамках модели Эберса—Молла. Подстановка (IV.2.5а, б) в (IV.2.6а, б) и (IV.2.7) дает

$$I_E = I_{1S}(e^{V_{be}/\Phi_T} - 1) - \alpha_I I_{2S}(e^{V_{bc}/\Phi_T} - 1); \quad (\text{IV.2.8a})$$

$$I_C = \alpha_N I_{1S}(e^{V_{be}/\Phi_T} - 1) - I_{2S}(e^{V_{bc}/\Phi_T} - 1); \quad (\text{IV.2.8б})$$

$$I_B = (1 - \alpha_N) I_{1S}(e^{V_{be}/\Phi_T} - 1) + (1 - \alpha_I) I_{2S}(e^{V_{bc}/\Phi_T} - 1), \quad (\text{IV.2.8в})$$

где I_{1S} и I_{2S} — тепловые токи эмиттерного и коллекторного переходов. Иногда их называют тепловыми токами эмиттерного и коллекторного диодов. Уравнения (IV.2.8) получили название *уравнений Эберса—Молла*.

Согласно модели Эберса—Молла ВАХ транзистора определяются четырьмя параметрами: α_N , α_I , I_{1S} и I_{2S} . Уравнения Эберса—Молла линейны относительно аргументов $(e^{V_{be}/\Phi_T} - 1)$ и $(e^{V_{bc}/\Phi_T} - 1)$, что является следствием линейности биполярных уравнений непрерывности в базе, эмиттере и коллекторе. Согласно общей теории линейных цепей из четырех параметров идеализированного транзистора независимыми являются только три. Можно показать, что для бездрейфовых транзисторов справедливо соотношение

$$\alpha_N I_{1S} = \alpha_I I_{2S}. \quad (\text{IV.2.9})$$

Это соотношение выполнено также и для дрейфовых транзисторов при условии, что коэффициенты переноса в прямом и инверсном включениях одинаковы: $\kappa_N = \kappa_I$. Ниже мы покажем, что это условие не вносит существенных ограничений,

так как в дрейфовых транзисторах коэффициенты переноса практически можно считать равными 1.

В МДП-транзисторе статический ток затвора равен нулю, поэтому в рамках идеализированной модели (не учитывающей ток подложки) его ВАХ описывается одним уравнением, связывающим ток стока (или равный ему ток истока) с двумя напряжениями между электродами. В биполярном транзисторе ток базы конечен, поэтому для описания его ВАХ необходимы два независимых уравнения, представляющие *входные* и *выходные* характеристики.

Способ представления ВАХ зависит от способа включения транзистора.

При включении по схеме с *общей базой* (ОБ, рис. IV.2.3, а) входным напряжением V_{in} является напряжение база—эмиттер V_{be} , входным током I_{in} — ток эмиттера I_E , выходным напряжением V_{out} — напряжение база—коллектор V_{bc} и выходным током I_{out} — ток коллектора I_C . Семейство входных характеристик представляет собой зависимости входного тока I_E от входного напряжения V_{be} с параметром — выходным напряжением V_{bc} :

$$I_E(V_{be}, V_{bc}).$$

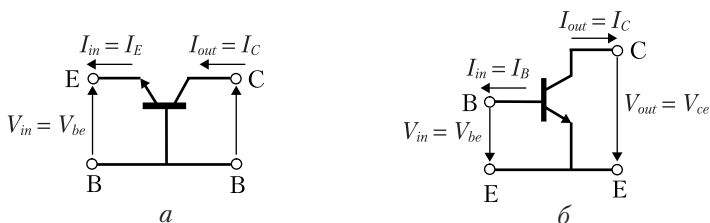


Рис. IV.2.3. Включение биполярного транзистора по схемам с общей базой ОБ (а) и с общим эмиттером ОЭ (б)

Семейство выходных характеристик представляет собой зависимости выходного тока I_C от выходного напряжения V_{bc} с параметром — входным током I_E :

$$I_C(V_{bc}, I_E).$$

При включении по схеме с *общим эмиттером* (ОЭ, рис. IV.2.3, б) входным напряжением V_{in} является напряжение база—эмиттер V_{be} , входным током I_{in} — ток базы I_B , выходным напряжением V_{out} — напряжение коллектор—эмиттер V_{ce} и выходным током I_{out} — ток коллектора I_C . Семейство вход-

ных характеристик представляет собой зависимости входного тока I_B от входного напряжения V_{be} с параметром — выходным напряжением V_{ce} :

$$I_B(V_{be}, V_{ce}).$$

Семейство выходных характеристик представляет собой зависимости выходного тока I_C от выходного напряжения V_{ce} с параметром — входным током I_B :

$$I_C(V_{ce}, I_B).$$

IV.2.2. Статические характеристики идеализированного транзистора в схеме ОБ

При включении ОБ *входные* ВАХ описываются уравнением (IV.2.8а). Вид входных ВАХ представлен на рис. IV.2.4, а, б для случаев $V_{bc} = 0$, $V_{bc} < 0$ и $V_{bc} > 0$ в мелком и крупном масштабах. При $V_{bc} = 0$ из (IV.2.8а) получаем диодную характеристику

$$I_E = I_{1S}(e^{V_{be}/\Phi_T} - 1).$$

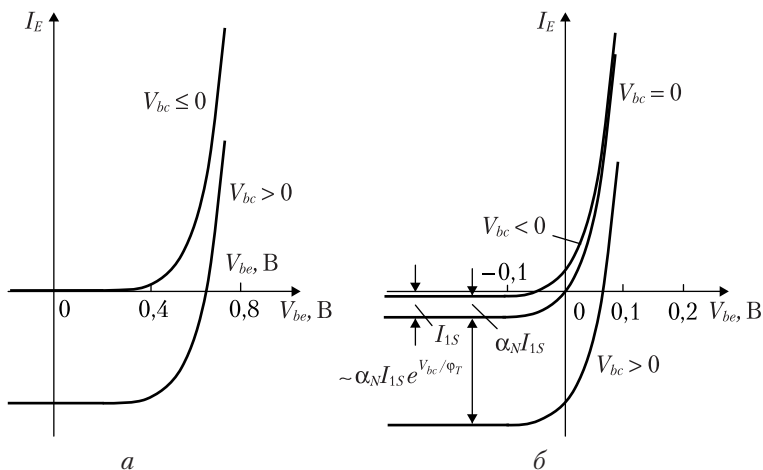


Рис. IV.2.4. Входные ВАХ идеализированного кремниевого транзистора при включении ОБ в мелком (а) и крупном (б) масштабах

При $V_{bc} < 0$ входные ВАХ сдвигаются в сторону больших токов на малую величину $\alpha_I I_{2S}(e^{V_{bc}/\Phi_T} - 1)$. При $V_{bc} < -3\Phi_T \approx 80$ мВ величина этого сдвига составляет $\alpha_I I_{2S} = \alpha_N I_{1S}$ и входные

характеристики не зависят от V_{bc} . При $V_{bc} > 0$ входные ВАХ сдвигаются в сторону меньших токов на величину $\alpha_I I_{2s}(e^{V_{bc}/\Phi_T} - 1) = \alpha_N I_{1s}(e^{V_{bc}/\Phi_T} - 1)$. Эта величина уже не является малой и экспоненциально зависит от V_{bc} .

Выходные ВАХ (рис. IV.2.5) могут быть получены из уравнений (IV.2.8а) и (IV.2.8б) путем исключения напряжения V_{be} . С учетом (IV.2.9) получаем

$$I_C = \alpha_N I_E - (1 - \alpha_N \alpha_I) I_{2s}(e^{V_{bc}/\Phi_T} - 1),$$

или
$$I_C = \alpha_N I_E - I_{C0}(e^{V_{bc}/\Phi_T} - 1), \quad (\text{IV.2.10})$$

где
$$I_{C0} = I_{2s}(1 - \alpha_N \alpha_I) \quad (\text{IV.2.11})$$

— *тепловой ток коллектора*. При графическом изображении выходных ВАХ по оси абсцисс откладывается напряжение $-V_{bc}$ (см. рис. IV.2.5), чтобы в нормальном режиме ВАХ находились в 1-ом квадранте. Соотношение (IV.2.10) показывает, что при $V_{bc} = 0$ ток коллектора $I_C = \alpha_N I_E$. Это соответствует определению коэффициента передачи эмиттерного тока (IV.1.3).

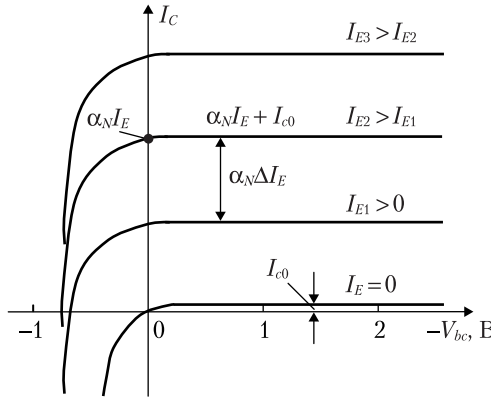


Рис. IV.2.5. Выходные ВАХ идеализированного кремниевого транзистора при включении ОБ

Выходные ВАХ пересекают ось абсцисс ($I_C = 0$) при напряжении

$$V_{bc} = \Phi_T \ln(1 + \alpha_N I_E / I_{C0}). \quad (\text{IV.2.12})$$

Обычно транзисторы работают в нормальном режиме при $V_{bc} < -3\Phi_T$. При этом из (IV.2.10) получаем

$$I_C|_{V_{bc} < -3\Phi_T} = \alpha_N I_E + I_{C0}. \quad (\text{IV.2.13a})$$

Структура выражения (IV.2.13а) показывает, что при закрытом коллекторном переходе ток коллектора складывается из *управляемого* тока $\alpha_N I_E$ и *неуправляемого* тока I_{C0} . В этом режиме коллекторный ток не зависит от напряжения на коллекторном переходе и *идеализированный транзистор является идеальным источником тока*. Практически трудно обеспечить нулевое напряжение V_{bc} , так как оно представляет собой напряжение на *коллекторном переходе* внутри транзистора (см. рис. IV.1.1) и не может быть измерено. Поэтому измерения коэффициента α_N производят при хорошо закрытом коллекторном переходе, измеряя приращение тока коллектора при заданном приращении тока эмиттера

$$\alpha_N = \Delta I_C / \Delta I_E.$$

Из соображений симметрии ясно, что в инверсном режиме при $V_{be} < -3\phi_T$ выполняется равенство

$$I_E \big|_{V_{be} < -3\phi_T} = \alpha_I I_C + I_{E0}, \quad (\text{IV.2.13б})$$

$$\text{где} \quad I_{E0} = I_{1S}(1 - \alpha_N \alpha_I) \quad (\text{IV.2.14})$$

— *тепловой ток эмиттера*. Соотношения (IV.2.11)–(IV.2.14) с учетом (IV.2.9) показывают, что для описания выходных ВАХ транзистора в схеме ОБ удобно заменить параметры I_{1S} и I_{2S} параметрами I_{E0} и I_{C0} , которые связаны между собой уравнением

$$\alpha_N I_{E0} = \alpha_I I_{C0}. \quad (\text{IV.2.15})$$

IV.2.3. Статические характеристики идеализированного транзистора в схеме ОЭ

При включении ОЭ *входные* ВАХ могут быть получены из уравнения (IV.2.8в) с учетом очевидного равенства $V_{bc} = V_{be} - V_{ce}$:

$$I_B = I_{1S} \frac{\beta_N + \beta_I}{\beta_I(\beta_N + 1)} \left[e^{(V_{be} - \Delta V)/\phi_T} - 1 \right], \quad (\text{IV.2.16})$$

где $\Delta V = \phi_T \ln[(\beta_N + \beta_I) / (\beta_N e^{-V_{ce}/\phi_T} + \beta_I)]$. Таким образом, входные ВАХ аналогичны ВАХ идеализированного диода с тепловым током $I_S^* = I_{1S}(\beta_N + \beta_I) / \beta_I(\beta_N + 1)$, сдвинутым вдоль оси напряжений V_{be} на величину ΔV (рис. IV.2.6, а, б). Обычно $\beta_N \gg \beta_I$, и $I_S^* \approx I_{1S} / \beta_I$. Если $V_{ce} = 0$, то $\Delta V = 0$ и входная ВАХ проходит через начало координат. При $V_{ce} > \sim [3 + \ln(\beta_N / \beta_I)] \phi_T$ ВАХ



сдвигается вправо на малую величину $\Delta V = \varphi_T \ln(1 + \beta_N / \beta_I) \approx \varphi_T \ln(\beta_N / \beta_I)$. При $V_{ce} < 0$ входные ВАХ сдвигаются в сторону меньших напряжений. Эта величина уже не является малой и при $V_{ce} \approx -3\varphi_T$ составляет $\Delta V \approx -V_{ce}$.

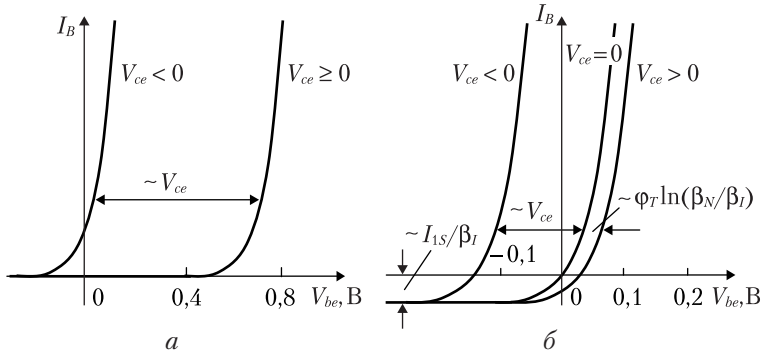


Рис. IV.2.6. Входные ВАХ идеализированного кремниевого транзистора при включении ОЭ в мелком (а) и крупном (б) масштабах

Выходные ВАХ (рис. IV.2.7) с учетом очевидного равенства $V_{bc} = V_{be} - V_{ce}$ могут быть получены из уравнений (IV.2.8б) и (IV.2.8в) путем исключения напряжения V_{be}

$$I_C = (\beta_N I_B + I_{C0}^*) \left[1 - \frac{1 + \beta_I / \beta_N + I_B / (\beta_N I_B + I_{C0}^*)}{1 + (\beta_I / \beta_N) e^{V_{ce} / \varphi_T}} \right], \quad (\text{IV.2.17})$$

где $I_{C0}^* = I_{C0}(1 + \beta_N)$ — тепловой ток коллектора в схеме ОЭ. Отметим следующие свойства выходных ВАХ в схеме ОЭ, представленных на рис. IV.2.7.

1. При $I_B = 0$ и $V_{ce} > \approx 3\varphi_T$ ток коллектора

$$I_C|_{I_B=0; V_{ce} > 3\varphi_T} = I_{C0}^*. \quad (\text{IV.2.19})$$

2. При $I_B \gg \beta_N I_{C0}^*$ и напряжении

$$V_{ce} > \approx 3\varphi_T \ln(1 + 1 / \beta_I) \quad (\text{IV.2.20})$$

ток коллектора не зависит от напряжения V_{ce} и составляет

$$I_C = \beta_N I_B + I_{C0}^*. \quad (\text{IV.2.21a})$$

Этот режим соответствует закрытому коллекторному переходу $V_{bc} < -3\varphi_T$, т.е. нормальному режиму работы транзистора.

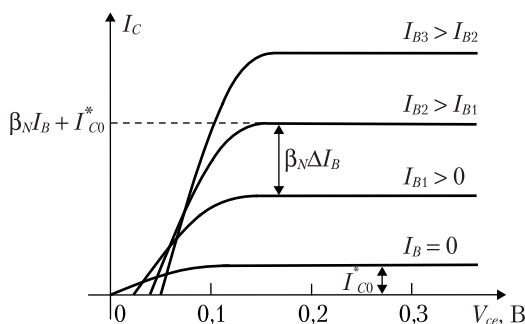


Рис. IV.2.7. Выходные ВАХ идеализированного кремниевого транзистора при включении ОЭ

Структура выражения (IV.2.21a) показывает, что при закрытом коллекторном переходе ток коллектора складывается из управляемого тока $\beta_N I_B$ и неуправляемого тока I_{C0}^* . Идеализированный транзистор является идеальным источником тока. Коэффициент усиления тока β_N можно измерить, используя соотношение

$$\beta_N = \Delta I_C / \Delta I_B.$$

Из соображений симметрии ясно, что в инверсном режиме при $-V_{ce} > \approx 3\phi_T \ln(1 + 1/\beta_N)$ выполняется равенство

$$I_E = \beta_I I_B + I_{E0}^*, \quad (\text{IV.2.21б})$$

где
$$I_{E0}^* = I_{E0}(1 + \beta_I) \quad (\text{IV.2.22})$$

— тепловой ток эмиттера в схеме ОЭ. Соотношения (IV.2.17)—(IV.2.22) с учетом (IV.2.9) показывают, что для описания выходных ВАХ транзистора в схеме ОЭ удобно заменить параметры I_{E0} и I_{C0} параметрами I_{E0}^* и I_{C0}^* , которые связаны между собой уравнением

$$\beta_N I_{E0}^* = \beta_I I_{C0}^*. \quad (\text{IV.2.23})$$

3. Область выходных ВАХ, где $I_C > 0$ (в отличие от включения ОБ) целиком располагается в 1-ом квадранте (см. рис. IV.2.7). Нулевому значению коллекторного тока соответствует напряжение

$$V_{ce} = \phi_T \ln \left(1 + \frac{\beta_N I_B / \beta_I}{\beta_N I_B + I_{C0}^*} \right) \xrightarrow{\beta_N I_B \gg I_{C0}^*} \phi_T \ln(1 + 1/\beta_I). \quad (\text{IV.2.24})$$

IV.2.4. Тепловые токи коллектора и эмиттера

Уравнение Эберса—Молла (IV.2.86) для коллекторного тока показывает, что тепловой ток коллекторного перехода (коллекторного диода D_2 на рис. IV.2.2) I_{2S} равен току коллектора при нулевом напряжении база—эмиттер ($V_{be} = 0$) и закрытом коллекторном переходе ($V_{bc} < -3\phi_T$) (рис. IV.2.8, а):

$$I_C|_{V_{be}=0, V_{bc}<-3\phi_T} = I_{2S}.$$

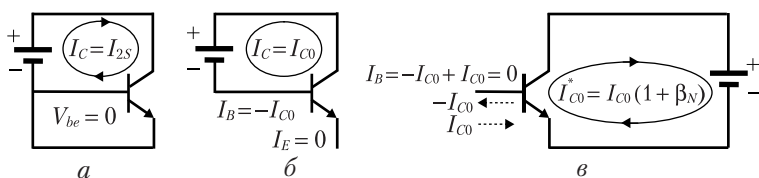


Рис. IV.2.8. Схемы измерения тепловых токов I_{2S} (а), I_{C0} (б) и I_{C0}^* (в)

Из соотношения (IV.2.13а) и рис. IV.2.5 следует, что при включении транзистора по схеме ОБ тепловой ток коллектора I_{C0} может быть измерен как ток коллектора при «оборванном» эмиттере ($I_E = 0$) и закрытом коллекторном переходе ($V_{bc} < -3\phi_T$) (рис. IV.2.8, б). Аналогично ток I_{E0} измеряется как ток эмиттера при $I_C = 0$, $V_{be} < -3\phi_T$.

Из (IV.2.19) и рис. IV.2.7 следует, что тепловой ток коллектора I_{C0}^* в схеме ОЭ может быть измерен как ток коллектора при «оборванной» базе ($I_B = 0$) и напряжении $V_{ce} > 3\phi_T$ (рис. IV.2.8, в). Аналогично ток I_{E0}^* измеряется как ток эмиттера при $I_B = 0$, $V_{ce} < -3\phi_T$.

Из соотношения (IV.2.18) следует, что в схеме включения ОЭ тепловой ток в $1 + \beta_N$ раз больше, чем в схеме ОБ: $I_{C0}^* = I_{C0}(1 + \beta_N)$. Этот результат можно объяснить, сравнивая схемы измерения токов I_{C0} и I_{C0}^* (см. рис. IV.2.8, б, в). Эти токи измеряются при одинаковых напряжениях, но различных значениях тока базы. В схеме измерения на рис. IV.2.8, б ток базы равен току коллектора с обратным знаком: $I_B = -I_{C0}$. Нулевой ток «оборванной» базы при измерении тока I_{C0}^* можно представить как сумму встречных токов I_{C0} и $-I_{C0}$ (см. рис. IV.2.8, в). При наличии только второго из этих слагаемых ($I_B = -I_{C0}$) воспроизводятся условия измерения в схеме на рис. IV.2.8, б, и ток коллектора должен иметь значение I_{C0} .

1. Напряжения V_{be} и V_{bc} соответствуют граничным концентрациям электронов в базе $n_p(0)$ и $n_p(w_B)$ (при низком уровне инжекции связаны с ними граничными условиями Шокли (II.3.7), см. параграф II.3.3).

2. Токи I_E и I_C пропорциональны градиентам концентрации электронов на границах базы, т.е. тангенсам углов наклона функции $n_p(x)$ при $x=0$ и $x=w_B$ (см. рис. IV.2.9).

3. Ток базы I_E при нашем предположении (и в рамках модели Эберса—Молла) связан только с рекомбинацией избыточных электронов в базе с концентрацией $\Delta n_p(x) = n_p(x) - n_{p0}$. Ток I_B пропорционален полному количеству избыточных электронов в базе, т.е. затемненной на рис. IV.2.9 площади.

Описанный метод во многих случаях помогает получить ясное представление о связи электрических характеристик транзистора с физическими процессами. В качестве примера на рис. IV.2.10 показана геометрическая интерпретация тепловых токов I_{2S} , I_{C0} и I_{C0}^* .

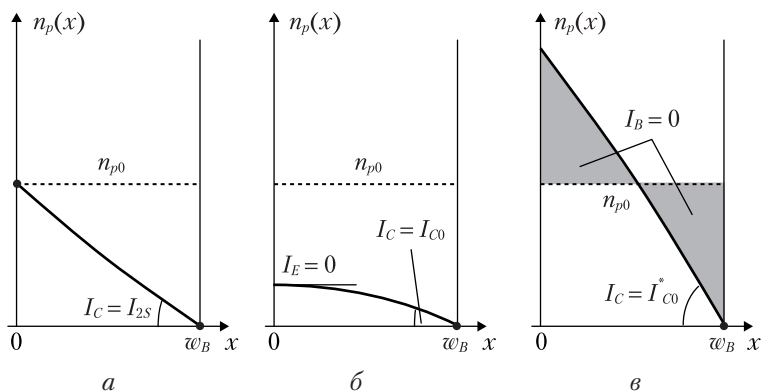


Рис. IV.2.10. Геометрическая интерпретация тепловых токов:

а — I_{2S} ($V_{be}=0$; $V_{bc} < -3\Phi_T$); б — I_{C0} ($V_{be}=0$; $V_{bc} < -3\Phi_T$);

в — I_{C0}^* ($I_B=0$; $V_{bc} < -3\Phi_T$)

Основные выводы

1. Простое математическое описание статических характеристик идеализированного биполярного транзистора дает модель Эберса—Молла. В основе модели Эберса—Молла лежит идея разложения токов через эмиттерный и коллекторный переходы на *инжектируемую* и *собираемую* составляющие. ВАХ транзистора при произвольном режиме могут быть получены из уравнений Эберса—Молла (IV.2.8а–в).

2. Для описания ВАХ биполярного транзистора необходимы два независимых уравнения, представляющих собой *входные* и *выходные* характеристики. При включении по схеме ОБ входные характеристики представляют собой функции $I_E(V_{be}, V_{bc})$, а выходные — функции $I_C(V_{bc}, I_E)$. При включении по схеме ОЭ входные характеристики представляют собой функции $I_B(V_{be}, V_{ce})$, а выходные — функции $I_C(V_{ce}, I_B)$.

3. ВАХ идеализированного транзистора определяются четырьмя параметрами: α_N , α_I , I_{1S} и I_{2S} , связанных соотношением $\alpha_N I_{1S} = \alpha_I I_{2S}$. При включении ОБ вместо тепловых токов эмиттерного и коллекторного переходов I_{1S} и I_{2S} удобно использовать тепловые токи эмиттера и коллектора $I_{E0} = I_{1S}(1 - \alpha_N \alpha_I)$ и $I_{C0} = I_{2S}(1 - \alpha_N \alpha_I)$, связанные соотношением $\alpha_N I_{E0} = \alpha_I I_{C0}$. При включении ОЭ удобно использовать тепловые токи $I_{E0}^* = I_{E0}(1 + \beta_I)$ и $I_{C0}^* = I_{C0}(1 + \beta_N)$, связанные соотношением $\beta_I I_{E0}^* = \beta_N I_{C0}^*$.

4. Входные ВАХ при $V_{bc} = 0$ (ОБ) и $V_{ce} = 0$ (ОЭ) подобны ВАХ идеализированного диода. В схеме ОБ при $V_{bc} < 0$ входные ВАХ сдвигаются в сторону больших токов на малую величину $\alpha_I I_{2S} = \alpha_N I_{1S}$, а при $V_{bc} > 0$ — в сторону меньших токов на величину $\approx \alpha_N I_{1S} e^{V_{bc}/\Phi_T}$. В схеме ОЭ при $V_{ce} > 0$ входные ВАХ сдвигаются в сторону больших напряжений на малую величину $\approx \Phi_T$, а при $V_{ce} < 0$ — в сторону меньших напряжений на величину $\approx -V_{ce}$.

5. В нормальном режиме работы транзистора выходной ток (ток коллектора) не зависит от выходного напряжения и складывается из управляемого тока ($\alpha_N I_E$ для включения ОБ и $\beta_N I_B$ для включения ОЭ) и неуправляемого тока (I_{C0} для включения ОБ и I_{C0}^* для включения ОЭ). При включении ОБ область положительных значений тока коллектора соответствует выходным напряжениям $-V_{bc} \approx V_{be}$, при включении ОЭ — выходным напряжениям $V_{ce} \approx 0$.

IV.3. Параметры идеализированного транзистора

В этом разделе будут получены соотношения, связывающие параметры идеализированной модели Эберса—Молла с электрофизическими и геометрическими параметрами транзисторной структуры. Основу физики биполярного транзистора составляет классическая теория Шокли [2], хотя она была разработана для сплавных бездрейфовых германиевых транзисторов, не содержащих вырожденных полупроводниковых слоев в одномерном приближении. Современные



интегральные транзисторы во многом отличаются от этих сравнительно простых приборов. Необходимые уточнения и обобщения будут сделаны в последующих разделах, однако некоторые коррекции целесообразно сделать уже в этой главе. На основе материала главы I могут быть без труда учтены эффекты вырождения, которые проявляются в основном в наиболее сильно легированной области эмиттера. Другой особенностью современных транзисторов является неоднородное распределение легирующей примеси в электронейтральных областях, в первую очередь в базе. Неоднородность распределения результирующей примеси приводит к существованию «встроенного» (не связанного с внешними воздействиями) электрического поля. При этом механизм переноса неосновных носителей уже не может считаться чисто диффузионным, а теоретические результаты сильно усложняются. Особенности и модели дрейфовых транзисторов будут рассмотрены в разделе IV.6. Это не означает, что результаты данной главы применимы только к сплавным бездрейфовым транзисторам. Просто вместо реальных примесных профилей мы будем использовать усредненные значения концентраций результирующей примеси, полагая их не зависящими от координат в пределах данного квазинейтрального слоя. Разумеется, дрейфовый механизм движения неосновных носителей при этом не учитывается. Поэтому такой подход соответствует *бездрейфовому приближению*, погрешности которого будут выявлены позже.

Основным параметром, определяющим усилительные свойства биполярного транзистора, является коэффициент передачи эмиттерного тока, равный произведению коэффициента переноса и эффективности эмиттера. Эти параметры будут рассмотрены в первую очередь.

IV.3.1. Коэффициент переноса

В соответствии с определением (IV.1.7) коэффициент переноса есть отношение электронных токов через коллекторный и эмиттерный переходы при нулевом напряжении база—коллектор

$$\kappa_N = I_{Cn} / I_{En} \big|_{V_{bc}=0}. \quad (\text{IV.3.1})$$

Токи I_{En} и I_{Cn} определяются как электронные токи в p -базе на границах с коллекторным и эмиттерным переходами



(см. рис. IV.1.3, а). Таким образом, коэффициент переноса определяется только свойствами базы. На рис. IV.3.1 представлено распределение электронов в p -базе при $V_{bc} = 0$, $V_{be} > 0$. Плоскости $x = 0$ и $x = w_B$ соответствуют границам базы с эмиттерным и коллекторным переходами. В бездрейфовом приближении при низком уровне инжекции токи являются диффузионными

$$I_{En} = S_E j_n(0) = e S_E D_B dn_p / dx(0); \quad (IV.3.2a)$$

$$I_{Cn} = S_E j_n(w_B) = e S_E D_B dn_p / dx(w_B), \quad (IV.3.2б)$$

где S_E — площадь эмиттерного перехода; D_B — коэффициент диффузии неосновных носителей (электронов) в базе.

Распределение электронов $n(x)$ находится путем решения биполярного уравнения непрерывности в базе. Для p -базы идеализированного транзистора и p -базы идеализированного диода биполярные уравнения непрерывности одинаковы. Граничные условия на границе с эмиттерным переходом для транзистора и диода также одинаковы (граничное условие Шокли). При $V_{bc} = 0$ в базе транзистора $n_p(w_B) = n_{p0}$ (см. рис. IV.3.1), что совпадает с граничным условием в базе диода на границе с омическим контактом. Следовательно, распределение электронов $n(x)$ определяется соотношением (II.4.15а)

$$\Delta n_p(x) = n_p(x) - n_{p0} = \Delta n_p(0) \cdot \frac{\text{sh}[(w_B - x) / L_B]}{\text{sh}(w_B / L_B)}, \quad (IV.3.3)$$

$$\text{где} \quad \Delta n_p(0) = n_{p0}(e^{V_{be}/\Phi_T} - 1); \quad (IV.3.4)$$

$$L_B = \sqrt{D_B \tau_N} \quad (IV.3.5)$$

и τ_N — диффузионная длина и время жизни неосновных носителей — электронов в базе. Обозначение τ_N соответствует нормальному режиму работы транзистора. В силу асимметрии реального транзистора в инверсном режиме время жизни $\tau_I \neq \tau_N$.

Подстановка (IV.3.2а, б) в (IV.3.1) с учетом (IV.3.3) дает

$$\kappa_N = 1 / \text{ch}(w_B / L_B). \quad (IV.3.6a)$$

В «хорошем» транзисторе всегда выполнено условие (IV.1.1): $w_B \ll L_B$. При этом

$$\kappa_N \approx 1 / (1 + w_B^2 / 2L_B^2) \approx 1 - w_B^2 / 2L_B^2. \quad (IV.3.6б)$$



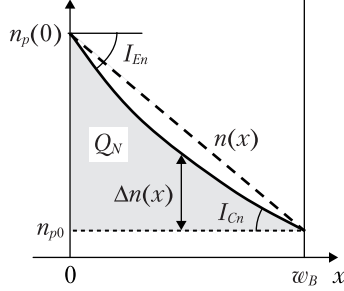


Рис. IV.3.1. Распределение электронов в p -базе при $V_{bc} = 0$, $V_{be} > 0$

С учетом (IV.3.5) соотношение (IV.3.6б) можно записать в виде

$$\kappa_N \approx 1 / (1 + T_N / \tau_N) \approx 1 - T_N / \tau_N, \quad (\text{IV.3.6в})$$

где
$$T_N = w_B^2 / 2D_B. \quad (\text{IV.3.7})$$

Заметим, что приближенное соотношение (IV.3.6в) можно получить из простых физических соображений, не прибегая к решению уравнения непрерывности. При $w_B \ll L_B$ большинство электронов успевает пролетать базу без рекомбинации. Поэтому ток $I_n(x)$ и, следовательно, градиент $dn_p/dx(x)$ слабо зависят от координаты x , и функция $n(x)$ весьма близка к линейной (штриховая линия на рис. IV.3.1). Абсолютная величина полного заряда избыточных электронов в базе Q_N и диффузионный ток I_{Cn} при этом составляют

$$Q_N = eS_E \int_0^{w_B} \Delta n_p(x) dx \approx \Delta n_p(0) w_B / 2; \quad (\text{IV.3.8})$$

$$I_{Cn} = eS_E dn/dx(0) \approx eS_E \Delta n(0) / w_B. \quad (\text{IV.3.9})$$

В отсутствие рекомбинации отношение $Q_N / I_{Cn} = w_B^2 / 2D_B$ есть среднее время пролета неосновных носителей заряда через базу длиной w_B . Этот результат показывает, что определенный в (IV.3.7) параметр T_N имеет смысл среднего времени диффузии неосновных носителей заряда через базу в нормальном направлении (от эмиттера к коллектору). Таким образом,

$$I_{Cn} = Q_N / T_N. \quad (\text{IV.3.10a})$$

Отношение Q_N / τ_N есть составляющая тока базы I_{Br} , связанная с рекомбинацией избыточных электронов в базе (см. рис. IV.1.3)

$$I_{Br} = Q_N / \tau_N. \quad (\text{IV.3.10б})$$

Электронный ток эмиттера есть сумма токов I_{Cn} и I_{Br} :

$$I_{En} = I_{Cn} + I_{Br} = Q_N (1 / T_N + 1 / \tau_N). \quad (\text{IV.3.10в})$$

Подставляя (IV.3.10а) и (IV.3.10в) в (IV.3.1), получаем соотношение (IV.3.6в), которое было выведено из уравнения непрерывности, учитывающего рекомбинацию в базе. Отметим, что соотношения (IV.3.8), (IV.3.9) и (IV.3.10а) для коллекторного тока и заряда в базе получены *без учета рекомбинации* и потому являются приближенными. Рекомбинация учтена только в соотношении (IV.3.10б). Разумеется, полное пренебрежение рекомбинацией дало бы значение $\kappa_N = 1$.

Соотношение (IV.3.6в) по точности эквивалентно (IV.3.6б). Оно, однако, может быть (и будет) использовано и для более сложных транзисторных структур, когда решение уравнения непрерывности в базе связано с математическими трудностями. Отметим также, что для бездрейфовых транзисторов условие (IV.1.1) $w_B \ll L_B$ эквивалентно условию

$$T_N \ll \tau_N. \quad (\text{IV.3.11})$$

Условие (IV.3.11) имеет, однако, более общий характер при надлежащем вычислении времени пролета T_N .

IV.3.2. Тепловые токи эмиттерного диода и эффективность эмиттера

Для электронной и дырочной составляющих тепловых токов эмиттерного диода I_{1S} и эффективности эмиттера γ_N справедливы соотношения, аналогичные соотношениям (II.4.21), (II.4.26)–(II.4.28), которые получены для идеализированного диода (см. раздел II.4):

$$I_{1Sn} = e S_E n_i^2 / G_B; \quad (\text{IV.3.12а})$$

$$I_{1Sp} = e S_E n_i^2 / G_E, \quad (\text{IV.3.12б})$$

$$I_{1S} = I_{1Sn} + I_{1Sp} = e S_E n_i^2 (G_B^{-1} + G_E^{-1}). \quad (\text{IV.3.13})$$

где S_E — площадь эмиттерного перехода.

Числа Гуммеля в эмиттере и базе определяются соотношением (II.4.27). Ввиду условий (IV.1.1) и (IV.1.2) база в транзисторах всегда короткая и невырожденная, поэтому

$$G_B = \frac{N_B w_B}{D_B}. \quad (\text{IV.3.14а})$$



В эмиттере условие $w_E \ll L_E$ может не выполняться. Кроме того, эмиттер практически всегда сильно вырожден. При этом эффективная концентрация примеси N_{Eef} , определяющая равновесную концентрацию неосновных носителей (дырок), близка к максимальному значению $N_{dm} = N_c e^{-\Delta E_d / kT}$ (соотношение (I.3.18)), и в соответствии с (II.4.27)

$$G_E = \frac{N_{dm} L_E}{D_E} \cdot \text{th} \frac{w_E}{L_E}. \quad (\text{IV.3.14б})$$

Эффективность эмиттера определяется соотношением (II.4.43)

$$\gamma_N = \frac{1}{1 + G_B / G_E} \approx 1 - \frac{G_B}{G_E}. \quad (\text{IV.3.15})$$

IV.3.3. Влияние коэффициента переноса и эффективности эмиттера на усилительные свойства транзистора

Поскольку в «хорошем» транзисторе коэффициенты $\alpha_N = \kappa_N \gamma_N$, κ_N и γ_N близки к 1, коэффициент усиления базового тока $\beta_N = \alpha_N / (1 - \alpha_N) \gg 1$ можно представить в виде

$$\beta_N^{-1} \approx \beta_\kappa^{-1} + \beta_\gamma^{-1}, \quad (\text{IV.3.16})$$

где с учетом (IV.3.6в), (IV.3.15) и (IV.3.14)

$$\beta_\kappa = \frac{\kappa_N}{1 - \kappa_N} \approx \frac{1}{1 - \kappa_N} \approx \frac{\tau_N}{T_N} = \frac{2L_B^2}{w_B^2} \quad (\text{IV.3.17a})$$

и

$$\beta_\gamma = \frac{\gamma_N}{1 - \gamma_N} \approx \frac{1}{1 - \gamma_N} \approx \frac{G_E}{G_B} = \frac{L_E}{w_B} \cdot \frac{N_{dm} D_B}{N_B D_E} \cdot \text{th} \frac{w_E}{L_E}. \quad (\text{IV.3.17б})$$

— коэффициенты усиления базового тока, связанные с коэффициентом переноса и эффективностью эмиттера, соответственно.

В реальных транзисторах, как правило, один из коэффициентов β_κ и β_γ оказывается значительно меньшим другого. При этом значение коэффициента усиления базового тока β_N равно меньшему из значений β_κ и β_γ .

Значение коэффициента $\beta_\kappa = 2(L_B / w_B)^2$ (IV.3.17a) зависит только от свойств базы. Возможность его повышения связана в основном с уменьшением длины базы w_B . Диф-

фузионная длина $L_B = \sqrt{D_B \tau_N}$ возрастает при снижении степени легирования базы за счет повышения и коэффициента диффузии $D_B = \Phi_T \mu_n$, и времени жизни неосновных носителей в базе. Оба этих параметра довольно слабо зависят от концентрации примеси (см. разделы II.4, II.5). Значение коэффициента β_γ зависит от свойств как базы, так и эмиттера. Как следует из (IV.3.176), коэффициент $\beta_\gamma \sim (w_B N_B)^{-1}$ также возрастает с уменьшением длины базы (слабее, чем $\beta_\kappa \sim w_B^{-2}$) и степени ее легирования. Увеличение длины эмиттера повышает коэффициент β_γ только до тех пор, пока она остается меньшей диффузионной длины дырок в эмиттере $L_E = \sqrt{D_E \tau_E}$. В вырожденном n -эмиттере значение коэффициента диффузии дырок составляет $\approx 1 \text{ см}^2/\text{с}$ (см. табл. I.3.1), время жизни при определяющем механизме Оже-рекомбинации составляет от 1 до 10 нс (см. раздел I.5) и диффузионная длина $L_E \approx 0,5 \text{ мкм}$.

Согласно данным табл. I.4.1, IV.1.1 и аппроксимации (I.4.8) в *сплавных транзисторах* (типичные значения параметров для кремниевого n - p - n транзистора: $D_B \approx 40 \text{ см}^2/\text{с}$, $D_E \approx 3 \text{ см}^2/\text{с}$, $L_B \approx 100 \text{ мкм}$, $L_E \approx 50 \text{ мкм}$, $w_B \approx 20 \text{ мкм}$, $w_E \approx 30 \text{ мкм}$, $N_B \approx 10^{15} \text{ см}^{-3}$, $N_E \approx 3 \cdot 10^{18} \text{ см}^{-3}$) $\beta_\kappa \approx 50 \ll \beta_\gamma \approx 300$, и коэффициент усиления базового тока определяется значением коэффициента переноса: $\beta_N \approx \beta_\kappa \approx 50$.

В *планарных транзисторах* (типичные значения параметров: $D_B \approx 15 \text{ см}^2/\text{с}$, $D_E \approx 1 \text{ см}^2/\text{с}$, $L_B \approx 10 \text{ мкм}$, $L_E \approx 0,5 \text{ мкм}$, $w_B \approx 0,2 \text{ мкм}$, $w_E \approx 0,3 \text{ мкм}$, $N_B \approx 4 \cdot 10^{17} \text{ см}^{-3}$, $N_E \approx 10^{20} \text{ см}^{-3}$) $\beta_\kappa \approx 10^4 \gg \beta_\gamma \approx 200$, и коэффициент усиления базового тока определяется значением эффективности эмиттера: $\beta_N \approx \beta_\gamma \approx 200$.

Отметим в заключение, что n - p - n транзисторы имеют более высокий коэффициент усиления, чем p - n - p , так как коэффициент диффузии электронов примерно в 2,5 раза выше, чем дырок.

IV.3.4. Инверсные параметры

При расчете инверсных параметров транзистора следует учитывать асимметрию его структуры, схематично показанной на рис. IV.3.2.

Тепловой ток коллекторного перехода складывается из электронной и дырочной составляющих:



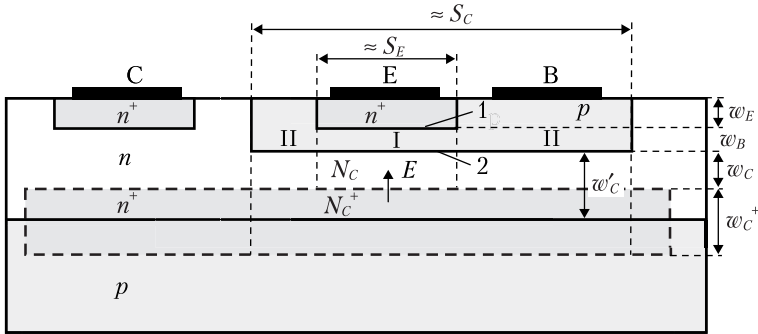


Рис. IV.3.2. Схематичное изображение структуры планарного ***n-p-n*** транзистора:

I — активная база; II — пассивная база; 1 — эмиттерный переход;
2 — коллекторный переход. На границе *n*- и *n*⁺-коллектора
стрелкой показано «встроенное» электрическое поле

$$I_{2S} = I_{2Sn} + I_{2Sp}. \quad (\text{IV.3.18})$$

Дырочная составляющая определяется свойствами коллектора.

При наличии скрытого *n*⁺-слоя в коллекторе в области перехода от *n*-коллектора к *n*⁺-слою формируется «встроенное» электрическое поле, которое направлено в сторону *n*-коллектора и уравнивает диффузионный поток электронов из *n*⁺-слоя в *n*-коллектор (рис. IV.3.3)¹. Это поле препятствует проникновению дырок, инжектированных в коллектор из базы, в *n*⁺-слой, и граница *n*-коллектор — *n*⁺-коллектор является «отражающей» для дырок. При этом дырки распределены в *n*-коллекторе практически равномерно (на рис. IV.3.3 $I_{2p}(w_{Cn}) \sim dp_n / dx(w_{Cn}) = 0$) и дырочный ток I_{2p} определяется скоростью рекомбинации заряда избыточных дырок Q_C в коллекторе

$$I_{2p} = Q_C / \tau_c \approx eS_C w_{Cn} \Delta p_n(0) / \tau_c. \quad (\text{IV.3.19})$$

Здесь τ_c — время жизни дырок в *n*-коллекторе; S_C — площадь коллекторного перехода, а граничная концентрация избыточных дырок на границе с коллекторным переходом ($x = 0$) определяется граничным условием Шокли

$$\Delta p_n(0) = (n_i^2 / N_C)(e^{V_{bc}/\Phi_T} - 1).$$

¹ Действие электрического поля на дырки в неоднородно легированном полупроводнике *n*-типа детально обсуждается в параграфе III.6.1.

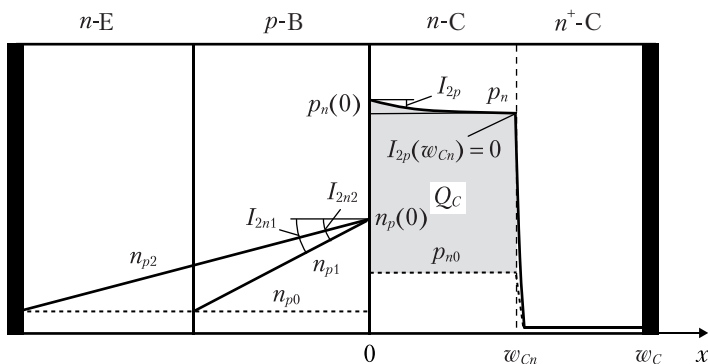


Рис. IV.3.3. Инверсный режим работы транзистора со скрытым n^+ -слоем в коллекторе ($V_{be} = 0$; $V_{bc} > 0$).

Распределения неосновных носителей в активной базе (n_{p1}), пассивной базе (n_{p2}) и коллекторе (p_n)

Таким образом, ВАХ тока I_{2p} имеет вид

$$I_{2p} = I_{2Sp} (e^{V_{bc}/\varphi_T} - 1),$$

где
$$I_{2Sp} \approx \frac{eS_C n_i^2 w_{Cn}}{N_C \tau_C} = \frac{eS_C n_i^2}{G_C}, \quad (\text{IV.3.20})$$

$$G_C = N_C \tau_C / w_{Cn} \quad (\text{IV.3.21a})$$

— модифицированное число Гуммеля для коллектора с n^+ -слоем.

В отсутствие скрытого n^+ -слоя для сравнительно слабо легированного n -коллектора выполнено условие $w'_C \ll L_C$ (см. рис. IV.3.2), и, по аналогии с (IV.3.14a), число Гуммеля в коллекторе составляет

$$G_C = N_C w_C / D_C. \quad (\text{IV.3.21б})$$

Электронная составляющая I_{2Sn} теплового тока I_{2S} определяется свойствами базы, в которой можно выделить активную область I (под эмиттером) и пассивную область II (см. рис. IV.3.2). Соответственно, ток I_{2Sn} делится на активную и пассивную составляющие

$$I_{2Sn} = I_{2SnI} + I_{2SnII}. \quad (\text{IV.3.22})$$

Активная составляющая I_{2SnI} совпадает с электронным током I_{1Sn} (IV.3.12a), так как определяется свойствами той же активной области базы (см. рис. IV.3.2)

$$I_{2SnI} = I_{1Sn} = eS_E n_i^2 / G_B, \quad (\text{IV.3.23a})$$

где число Гуммеля G_B определено в (IV.3.14a). Пассивная составляющая I_{2SnII} определяется свойствами пассивной области базы II площадью $S_C - S_E$ (см. рис. IV.3.2). Для оценочных расчетов можно полагать, что омический контакт к базе занимает большую часть этой площади. При этом

$$I_{2SnII} = e(S_C - S_E)n_i^2 / G_{BII}, \quad (IV.3.23б)$$

где
$$G_{BII} = \frac{N_{BII}(w_B + w_E)}{D_{BII}}. \quad (IV.3.24)$$

Средняя концентрация примеси N_{BII} в пассивной базе больше, чем в активной (см. рис. IV.1.2, б); соответственно, коэффициент диффузии электронов D_{BII} — несколько меньше.

Суммируя соотношения (IV.3.20) и (IV.3.23а, б), получаем

$$I_{2S} = en_i^2 \left[\frac{S_E}{G_B} + \frac{S_C}{G_C} + \frac{S_C - S_E}{G_{BII}} \right], \quad (IV.3.25)$$

где числа Гуммеля G_B , G_{BII} и G_C определены в (IV.3.14а), (IV.3.24) и (IV.3.21а, б).

Инверсный коэффициент передачи тока α_I может быть найден из соотношения (IV.2.9)

$$\alpha_I = \alpha_N I_{IS} / I_{2S}. \quad (IV.3.26)$$

Отметим, что наличие скрытого n^+ -слоя в коллекторе значительно уменьшает инжекцию дырок из базы в коллектор. Это приводит к уменьшению дырочной составляющей I_{2Sp} теплового тока I_{2S} и, как следствие, к увеличению инверсного коэффициента передачи тока α_I . Обычно даже при наличии скрытого n^+ -слоя в коллекторе $G_C < G_E$, а площадь коллекторного перехода S_C значительно больше, чем эмиттерного S_E (см. рис. IV.3.2). Сравнение соотношений (IV.3.25) и (IV.3.13) показывает, что за счет этого значение I_{2S} значительно больше, чем I_{IS} . Таким образом, инверсный коэффициент передачи тока α_I значительно меньше нормального α_N и не может считаться близким к 1. В планарных транзисторах, как правило, $\alpha_I < 0,5$ и инверсный коэффициент усиления тока базы $\beta_I = \alpha_I / (1 - \alpha_I) < 1$.

Основные выводы

1. В бездрейфовом приближении коэффициент переноса определяется соотношением (IV.3.6а) и составляет $\kappa_N = 1 / \text{ch}(w_B / L_B)$. Приблизительно он является функцией отношения времени пролета



неосновных носителей через базу T_N к их времени жизни в базе τ_N : $\kappa_N \approx 1 - T_N / \tau_N$. Это соотношение имеет более общий характер при надлежащем вычислении времени пролета.

2. Эффективность эмиттера определяется отношением чисел Гуммеля в базе и эмиттере (IV.3.15): $\gamma_N = 1 / (1 + G_B / G_E)$. Число Гуммеля в базе описывается простой формулой $G_B = N_B w_B / D_B$ (IV.3.14a). При вычислении числа Гуммеля в эмиттере (IV.3.14б) необходимо учитывать вырождение эмиттера.

3. Коэффициент усиления базового тока можно представить в виде: $\beta_N^{-1} \approx \beta_K^{-1} + \beta_\gamma^{-1}$, где $\beta_K = \kappa_N / (1 - \kappa_N)$ и $\beta_\gamma = \gamma_N / (1 - \gamma_N)$ — коэффициенты усиления, связанные с коэффициентом переноса и эффективностью эмиттера. В сплавных транзисторах $\beta_N \approx \beta_K$; в планарных транзисторах $\beta_N \approx \beta_\gamma$.

4. Инверсные параметры транзистора в значительной степени определяются асимметрией его структуры. Наличие скрытого n^+ -слоя в коллекторе значительно снижает тепловой ток I_{2sp} и повышает инверсный коэффициент передачи тока α_I . В любом случае значение α_I значительно меньше α_N и не может считаться близким к 1. В планарных транзисторах инверсный коэффициент усиления тока базы не превышает 1.

IV.4. Частотные и импульсные свойства коэффициентов передачи тока

IV.4.1. Механизм внутренней инерционности биполярного транзистора и форма представления частотных и импульсных характеристик

Как всякий усилительный прибор, биполярный транзистор обладает некоторой инерционностью, ограничивающей частотный диапазон усиления и скорость передачи импульсных сигналов. Принципиальная внутренняя инерционность заложена в самом транзисторном эффекте. Она связана с конечным временем переноса неосновных носителей через базу, т.е. с конечным временем установления стационарных распределений носителей в базе и эмиттере $n_p(x)$ и $p_n(x)$ при мгновенном изменении напряжения база—эмиттер. Таким образом, коэффициент передачи эмиттерного тока и коэффициент усиления тока базы должны зависеть от скорости изменения входных сигналов.

В разделе III.1 коэффициент передачи эмиттерного тока и его составляющие — коэффициент переноса и эффективность



эмиттера — считались *статическими* параметрами и определены для нормального режима соотношениями (IV.1.6)–(IV.1.8) как отношения соответствующих токов при нулевом напряжении V_{bc} . При этом выходной коллекторный ток при $V_{bc} < 0$ является суммой *управляемого* и *неуправляемого* токов (для включения ОБ $I_C = \alpha_N I_E + I_{C0}$). Анализ частотных свойств прибора основан на исследовании малых токов и напряжений — сигналов, — не содержащих постоянных (неуправляемых) составляющих. Для этого упомянутые выше коэффициенты удобно представить в *дифференциальной форме*, определив их для нормального режима работы транзистора производными

$$\alpha = \kappa \gamma = \partial I_C / \partial I_E \big|_{dV_{bc}=0}; \quad (IV.4.1)$$

$$\kappa = \partial I_C / \partial I_{En} \big|_{dV_{bc}=0}; \quad (IV.4.2)$$

$$\gamma = \partial I_{En} / \partial I_E \big|_{dV_{bc}=0}. \quad (IV.4.3)$$

В определениях (IV.4.1)–(IV.4.3) опущен индекс N , так как частотные характеристики, как правило, исследуются только в нормальном режиме. Дифференциальные параметры могут отличаться от статических, если они не являются константами (как принято в модели идеализированного транзистора), не зависящими от режима по постоянному току (на самом деле такая зависимость существует).

При дифференциальном определении параметров для коэффициента усиления тока базы остается в силе соотношение

$$\beta = \alpha / (1 - \alpha). \quad (IV.4.4a)$$

Как принято в теории линейных цепей, частотные характеристики коэффициентов передачи тока α и β исследуются на гармоническом входном сигнале — токах эмиттера или базы, представленных в комплексной экспоненциальной форме $Ie^{j\omega t}$. Импульсные характеристики исследуются на стандартном импульсном входном сигнале в виде ступенчатого перепада входного тока (эмиттера или базы). Частотные характеристики коэффициентов передачи тока

$$\dot{\alpha}(\omega) = \dot{\kappa}(\omega) \dot{\gamma}(\omega), \quad (IV.4.4б)$$

$$\dot{\beta}(\omega) = \dot{\alpha}(\omega) / [1 - \dot{\alpha}(\omega)] \quad (IV.4.4в)$$



представляются в виде амплитудно-частотных и фазо-частотных характеристик (АЧХ и ФЧХ), определенных как частотные зависимости их модуля $|\dot{\alpha}, \dot{\beta}|(\omega)$ и фазы $\varphi_{\alpha, \beta}(\omega)$. Импульсные характеристики представляются в виде операторных изображений тех же коэффициентов, которые формально получаются из частотных характеристик заменой $i\omega \rightarrow s$, где s — оператор Хэвисайда–Карсона.

IV.4.2. Частотная характеристика коэффициента переноса

Частотная характеристика коэффициента переноса может быть получена путем решения нестационарного биполярного уравнения непрерывности в базе, определяющего распределение малой переменной составляющей концентрации неосновных носителей $\dot{n}_p^1(x, \omega)$. После этого, по аналогии со стационарными уравнениями (IV.3.2а, б), коэффициент переноса определяется как отношение

$$\dot{\kappa}(\omega) = \frac{\partial \dot{n}_p(x, \omega) / \partial x|_{(w_B, \omega)}}{\partial \dot{n}_p(x, \omega) / \partial x|_{(0, \omega)}}.$$

Процедура нахождения функции $\dot{n}_p^1(x, \omega)$ была выполнена при анализе частотных свойств полупроводникового диода в параграфе II.7.4. Было показано, что функция $\dot{n}_p^1(x, \omega)$ совпадает с решением стационарного уравнения непрерывности $n_p(x)$ с заменой диффузионной длины неосновных носителей в базе L_B на ее комплексную величину, определенную в (II.7.27).

Таким образом, частотная зависимость коэффициента переноса может быть получена из (IV.3.6а) с заменой диффузионной длины на ее комплексное значение:

$$\dot{\kappa}(\omega) = 1 / \operatorname{ch}[w_B / \dot{L}_B(\omega)], \quad (\text{IV.4.5})$$

$$\text{где} \quad \dot{L}_B = L_B / \sqrt{1 + i\omega\tau_N}. \quad (\text{IV.4.6})$$

Разлагая гиперболический косинус в (IV.4.5) по степеням $w_B / \dot{L}_B(\omega)$ и ограничиваясь линейным или квадратичным по частоте членом, получаем с учетом (IV.3.6б, в)

$$\dot{\kappa}(\omega) \approx \frac{1}{1 + (i\omega\tau_N)(w_B^2 / 2L_B^2) + (i\omega\tau_N)^2(w_B^2 / 2L_B^2)^2 / 6};$$



$$\dot{\kappa}(\omega) \approx \frac{\kappa}{1 + i\omega\kappa T_N}; \quad (\text{IV.4.7a})$$

$$\begin{aligned} \dot{\kappa}(\omega) &\approx \frac{\kappa}{1 + i\omega\kappa T_N + (i\omega\kappa T_N)^2 / 6} \approx \\ &\approx \frac{\kappa}{[1 + i\omega\kappa T_N(1 - \xi)] \cdot [1 + i\omega\kappa T_N \xi]}, \end{aligned} \quad (\text{IV.4.7б})$$

или
$$\dot{\kappa}(\omega) \approx \frac{\kappa e^{-i\omega\kappa T_N \xi}}{1 + i\omega\kappa T_N(1 - \xi)}, \quad (\text{IV.4.7в})$$

где время пролета через базу $T_N = \omega_B^2 / 2D_B$ определено в (IV.3.7),

$$\xi = (1 - 1/\sqrt{3}) / 2 \approx 0,21. \quad (\text{IV.4.8})$$

Соотношение (IV.4.7в) представляет собой более удобную для моделирования, чем (IV.4.7б), и достаточно точную аппроксимацию (IV.4.5).

IV.4.3. Частотная характеристика эффективности эмиттера

Изначально эффективность эмиттера определена соотношением (IV.1.8) как доля электронной составляющей в полном токе через эмиттерный переход при $V_{bc} = 0$. В принципе ее частотная характеристика может быть получена таким же путем, как для коэффициента переноса, — решением стационарных биполярных уравнений непрерывности в базе и эмиттере и отысканием функций $\dot{n}_p^1(x, \omega)$ и $\dot{n}_n^1(x, \omega)$. Результаты такого анализа показывают, что в первом приближении частотная характеристика эффективности эмиттера имеет вид, аналогичный (IV.4.7а):

$$\dot{\gamma}(\omega) \approx \frac{\gamma}{1 + i\omega\tau_\gamma}, \quad (\text{IV.4.9})$$

где
$$\tau_\gamma \approx \frac{(1 - \gamma)T_E}{1 + 2T_E / \tau_E}, \quad (\text{IV.4.10})$$

$T_E = \omega_E^2 / 2D_E$ — время пролета дырок через эмиттер; τ_E — время жизни дырок в эмиттере. Как правило, благодаря множителю $1 - \gamma \ll 1$ в (IV.4.10)

$$\tau_\gamma \ll T_N. \quad (\text{IV.4.11})$$



Это позволяет считать, что *эффективность эмиттера не зависит от частоты* вплоть до частот $\omega \approx 1/T_N$, когда величина $|\dot{\kappa}(\omega)|$ становится заметно меньшей 1.

Физическое объяснение этого факта состоит в том, что коэффициент переноса определяется значением коллекторного тока при фиксированном значении электронного тока эмиттера, а эффективность эмиттера — соотношением электронного и дырочного токов эмиттера при фиксированном значении *напряжения* на эмиттерном переходе. Различие в степени инерционности этих коэффициентов поясняется рис. IV.4.1, где изображены переходные процессы установления стационарного распределения электронов в базе $n_p(x)$ при скачкообразном включении электронного тока эмиттера I_{En} (а) и напряжения база—эмиттер V_{be} (б) при $V_{bc} = 0$. В первом случае постоянное значение I_{En} соответствует постоянному углу наклона функции $n_p(x)$ при $x=0$; во втором случае постоянное значение V_{be} соответствует постоянному значению $n_p(0)$. В обоих случаях возрастание коллекторного тока $I_C(t)$ в ряд последовательных моментов времени $0 < t_1 < t_2 < t_3 < t_4 < \infty$ соответствует углу наклона функции $n_p(x)$ при $x = w_B$. Из рисунка ясно видно, что установление стационарного значения коллекторного тока в случае (б) происходит значительно быстрее.

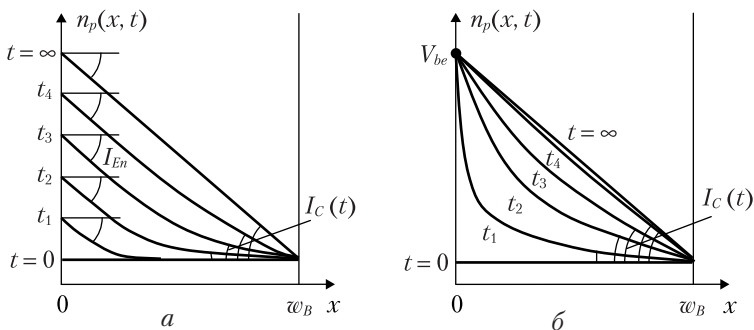


Рис. IV.4.1. Процессы установления стационарного распределения электронов в базе $n_p(x)$ при скачкообразном включении электронного тока эмиттера I_{En} (а) и напряжения база—эмиттер V_{be} при $V_{bc} = 0$ (б); $0 < t_1 < t_2 < t_3 < t_4 < \infty$

IV.4.4. Частотные и импульсные свойства коэффициента передачи эмиттерного тока

При условии (IV.4.11) частотная характеристика коэффициента передачи эмиттерного тока определяется соотношением (IV.4.5) либо приближениями (IV.4.7а), (IV.4.7в)¹:

$$\dot{\alpha}(\omega) \approx \frac{\alpha}{1 + i\omega\tau_\alpha} = \frac{\alpha}{1 + i\omega / \omega_\alpha}; \quad (\text{IV.4.12a})$$

$$\dot{\alpha}(\omega) \approx \frac{\alpha e^{-i\omega\xi_\alpha\tau_\alpha}}{1 + i\omega\tau_\alpha(1 - \xi_\alpha)} = \frac{\alpha e^{-i\xi_\alpha\omega / \omega_\alpha}}{1 + i(1 - \xi_\alpha)\omega / \omega_\alpha}, \quad (\text{IV.4.12б})$$

где $\tau_\alpha = \alpha T_N$ (IV.4.13)

— постоянная времени коэффициента передачи эмиттерного тока,

$$\omega_\alpha = 1 / \tau_\alpha \quad (\text{IV.4.14})$$

— верхняя граничная частота коэффициента передачи эмиттерного тока.

Из (IV.4.12а) получаем простейший вид АЧХ и ФЧХ коэффициента передачи эмиттерного тока

$$\frac{|\dot{\alpha}(\omega)|}{\alpha} \approx \frac{1}{\sqrt{1 + (\omega / \omega_\alpha)^2}}; \quad (\text{IV.4.15a})$$

$$\varphi_\alpha(\omega) \approx -\text{Arctg}(\omega / \omega_\alpha). \quad (\text{IV.4.15б})$$

Более точные выражения для АЧХ и ФЧХ дает соотношение (IV.4.12б):

$$\frac{|\dot{\alpha}(\omega)|}{\alpha} \approx \frac{1}{\sqrt{1 + (\xi_\alpha\omega / \omega_\alpha)^2}}; \quad (\text{IV.4.16a})$$

$$\varphi_\alpha(\omega) \approx -\text{Arctg}(\xi_\alpha\omega / \omega_\alpha) - (1 - \xi_\alpha)\omega / \omega_\alpha. \quad (\text{IV.4.16б})$$

АЧХ и ФЧХ коэффициента передачи эмиттерного тока представлены на рис. IV.4.2. Как очевидно из рисунка, простое приближение (IV.4.15а) хорошо воспроизводит АЧХ, но точность приближения (IV.4.15б) для ФЧХ значительно ниже.

Операторные изображения переходных характеристик коэффициента передачи эмиттерного тока соответствуют частотным характеристикам (IV.4.12а, б)

¹ Значения коэффициента ξ_α в диффузионном и дрейфовом (см. параграф IV.6.4) транзисторах различны. В рассматриваемом случае диффузного транзистора $\xi_\alpha = \xi$.



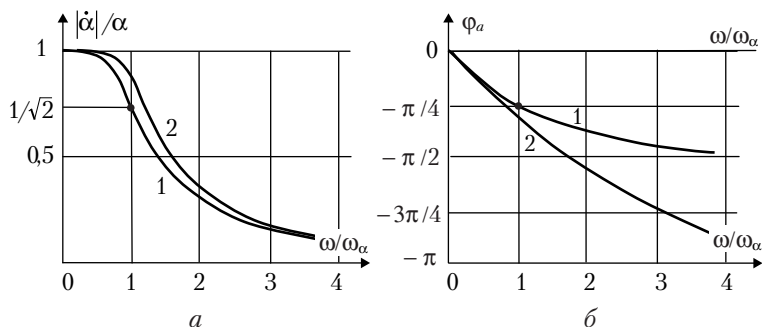


Рис. IV.4.2. Амплитудно-частотные (а) и фазо-частотные (б) характеристики коэффициента передачи эмиттерного тока:
1 — формулы (IV.4.15а, б); 2 — формулы (IV.4.16а, б)

$$\dot{\alpha}(s) \approx \frac{\alpha}{1 + s\tau_{\alpha}}; \quad (\text{IV.4.17a})$$

$$\dot{\alpha}(s) \approx \frac{\alpha e^{-s\xi_{\alpha}\tau_{\alpha}}}{1 + (1 - \xi_{\alpha})s\tau_{\alpha}}. \quad (\text{IV.4.17б})$$

Оригиналы этих изображений имеют вид

$$\alpha(t) \approx \alpha(1 - e^{-t/\tau_{\alpha}}) \quad (\text{IV.4.18a})$$

$$\alpha(t) \approx \begin{cases} 0, & t < \xi_{\alpha}\tau_{\alpha}, \\ \alpha(1 - e^{-t/(1-\xi_{\alpha})\tau_{\alpha}}), & t \geq \xi_{\alpha}\tau_{\alpha}. \end{cases} \quad (\text{IV.4.18б})$$

Переходные характеристики (IV.4.18а, б) представлены на рис. IV.4.3 кривыми 1 и 2; пунктиром показана точная функция $\alpha(t)$, полученная из (IV.4.5). Конечная длительность фронта коллекторного тока $I_C(t) = \alpha(t)I_E$ при ступенчатом изменении управляющего тока I_E связана с диффузионным механизмом переноса неосновных носителей через базу.

Соотношение (IV.4.18а) показывает, что *среднее* время диффузии через базу составляет $T_N \approx \tau_{\alpha}$, однако носители движутся с различными скоростями и достигают коллекторного перехода в разное время. Пространственный фронт инжектированных в базу носителей (рис. IV.4.1, а) за среднее время диффузии $T_N \approx \tau_{\alpha}$ размывается на величину $\Delta x_{\text{dif}} \approx \sqrt{2D_B T_N}$, которая с учетом (IV.3.7) составляет $\Delta x \approx w_B$. Поэтому коллекторный ток $I_C(t)$ плавно нарастает от нуля до установившегося

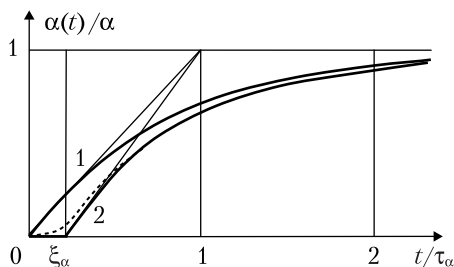


Рис. IV.4.3. Переходные характеристики коэффициента передачи эмиттерного тока:

1 — приближение (IV.4.18a); 2 — приближение (IV.4.18б);
пунктир — точная функция

значения αI_E за время порядка среднего времени диффузии неосновных носителей через базу T_N .

Соотношение (IV.4.18б) можно интерпретировать следующим образом. «Наиболее быстрые электроны» диффундируют через базу за время *диффузионной задержки*

$$t_{d \text{ dif}} \approx \xi_\alpha T_N. \quad (\text{IV.4.19})$$

После этого остальные носители, движущиеся с различными скоростями, собираются коллекторным переходом за время $(1 - \xi_\alpha)T_N$.

Проведенный анализ показывает, что инерционность коэффициента передачи эмиттерного тока определяется только свойствами базы, а именно — средним временем пролета неосновных носителей от эмиттерного перехода до коллекторного.

IV.4.5. Частотные и импульсные свойства коэффициента усиления тока базы

Частотные характеристики коэффициента усиления базового тока определяются соотношением (IV.4.6): $\beta(\omega) = \dot{\alpha}(\omega) / [1 - \dot{\alpha}(\omega)]$. Использование приближений (IV.4.12a, б) дает

$$\beta(\omega) \approx \frac{\beta}{1 + i\omega\tau_\beta} = \frac{\beta}{1 + i\omega / \omega_\beta}; \quad (\text{IV.4.20a})$$

$$\beta(\omega) \approx \frac{\beta e^{-i\omega\tau_\alpha}}{1 + i\omega\tau_\beta} = \frac{\beta e^{-i\omega / \omega_\alpha}}{1 + i\omega / \omega_\beta}, \quad (\text{IV.4.20б})$$

где
$$\tau_\beta = (1 + \beta)\tau_\alpha \quad (\text{IV.4.21})$$

— постоянная времени коэффициента усиления базового тока,

$$\omega_\beta = \omega_\alpha / (1 + \beta) \quad (\text{IV.4.22})$$

— верхняя граничная частота коэффициента усиления базового тока. Из (IV.4.21) и (IV.4.22) следует, что $\tau_\beta \gg \tau_\alpha$ и $\omega_\beta \ll \omega_\alpha$, поэтому точность частотных характеристик (IV.4.20a) и (IV.4.20б) практически одинакова.

Операторное изображение переходной характеристики коэффициента передачи базового тока соответствуют (IV.4.20a)

$$\dot{\beta}(s) \approx \frac{\beta}{1 + s\tau_\beta}. \quad (\text{IV.4.23})$$

Оригинал этого изображения имеет вид

$$\beta(t) \approx \beta(1 - e^{-t/\tau_\beta}). \quad (\text{IV.4.24})$$

Амплитудно-частотная $|\dot{\beta}(\omega)|$ и переходная $\beta(t)$ характеристики представлены на рис. IV.4.4, а, б, где для сравнения показаны АЧХ и переходная характеристика коэффициента передачи эмиттерного тока (IV.4.16a) и (IV.4.18a).

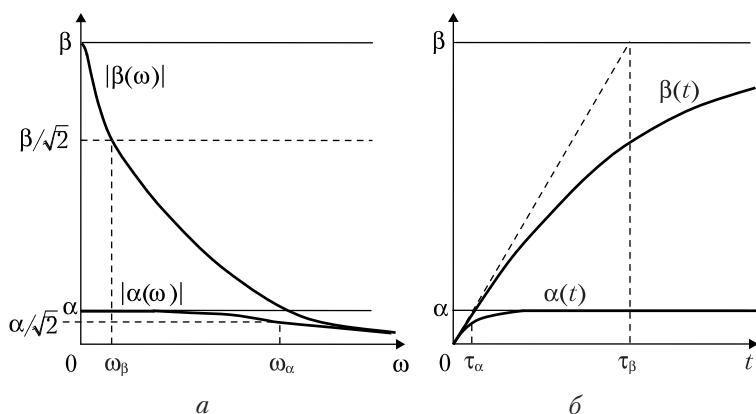


Рис. IV.4.4. Амплитудно-частотные (а) и переходные (б) характеристики коэффициентов α и β .

Для $\dot{\alpha}(\omega)$ и $\alpha(t)$ использованы приближения (IV.4.12a) и (IV.4.18a); для $\dot{\beta}(\omega)$ и $\beta(t)$ — приближения (IV.4.20a) и (IV.4.24)

Соотношения (IV.4.21) и (IV.4.22) показывают, что в общем случае инерционность коэффициента усиления базового тока определяется средним временем пролета неосновных носителей через базу T_N , а также статическим значением коэффициента усиления β . Таким образом, она зависит как от свойств базы, так и от свойств эмиттера (влияющих на значение β). Исключение составляет случай, когда значение эффективности эмиттера γ_N значительно ближе к 1, чем значение коэффициента переноса κ_N . При этом $\alpha_N \approx \kappa_N$, $\beta_N \approx \beta_\kappa = \tau_N / T_N$ (IV.3.17a), и постоянная времени τ_β имеет смысл времени жизни неосновных носителей в базе: $\tau_\beta \approx \tau_N$. Этому факту можно дать простое физическое объяснение. При $\gamma_N = 1$ инжекция из базы в эмиттер отсутствует, и при выключении тока базы заряд избыточных носителей в базе Q_N (см. рис. IV.3.1) исчезает только за счет рекомбинации, т.е. за время τ_N . Коллекторный ток пропорционален этому заряду (IV.3.10a), следовательно, время его выключения (постоянная времени τ_β) также соответствует времени жизни τ_N .

IV.4.6. Сравнительный анализ частотных и импульсных характеристик коэффициентов α и β

Отметим, что коэффициент α соответствует коэффициенту усиления транзистора по току в схеме включения ОБ, а коэффициент β — коэффициенту усиления транзистора по току в схеме ОЭ. Эти коэффициенты связаны соотношением $\beta = \alpha / (1 - \alpha)$ или $\alpha = \beta / (1 + \beta)$, поэтому их отношение составляет

$$\beta / \alpha = 1 + \beta. \quad (\text{IV.4.25})$$

Сравнение частотных и переходных характеристик (IV.4.12a), (IV.4.18a) (для коэффициента α) и (IV.4.20a), (IV.4.24) (для коэффициента β), а также данных рис. IV.4.4 показывает, что *при включении транзистора по схеме ОЭ выигрыш в усилении тока достигается ценой равной потери быстродействия*:

$$\omega_\beta / \omega_\alpha = \tau_\alpha / \tau_\beta = \alpha / \beta = 1 / (1 + \beta). \quad (\text{IV.4.26})$$

Физическая причина резкого снижения модуля коэффициента усиления базового тока $\beta = \dot{\alpha} / (1 - \dot{\alpha})$ с частотой $\omega > \omega_\beta$ (см. рис. IV.4.4) состоит главным образом не в уменьшении



модуля коэффициента передачи $|\dot{\alpha}(\omega)|$, а в увеличении его фазового сдвига $\varphi_\alpha(\omega)$. Это соображение поясняется рис. IV.4.5, где изображены векторные диаграммы токов $\dot{I}_E$, $\dot{I}_C$ и $\dot{I}_B = \dot{I}_E - \dot{I}_C$ при $\omega \rightarrow 0$ (рис. IV.4.5, а) и при $\omega = \omega_\alpha$ (рис. IV.4.5, б). При $\omega \rightarrow 0$ токи $\dot{I}_C$ и $\dot{I}_E$ совпадают по фазе и ток $\dot{I}_B$ является их арифметической разностью. При $\omega = \omega_\alpha$ фазовый сдвиг токов $\dot{I}_C$ и $\dot{I}_E$ составляет $\varphi_\alpha = -\pi/4$ (соотношение (IV.4.166), рис. IV.4.2, б). Рисунок IV.4.5, б показывает, что увеличение модуля тока $|\dot{I}_B|$, а следовательно, и снижение $|\dot{\beta}| = |\dot{I}_C / \dot{I}_B|$ определяется в основном фазовым сдвигом $\varphi_\alpha = -\pi/4$.

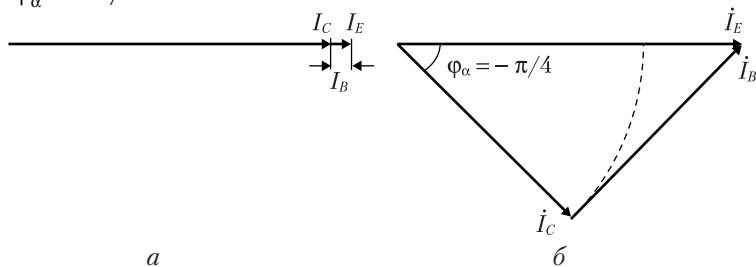


Рис. IV.4.5. Векторные диаграммы токов $\dot{I}_E$, $\dot{I}_C$ и $\dot{I}_B$:
а — $\omega = 0$; б — $\omega = \omega_\alpha$

Отметим также, что на очень высокой частоте ($\omega \gg \omega_\alpha$) или в области очень малого времени ($t \ll \tau_\alpha$), когда усилительные свойства транзистора практически потеряны, значения $\dot{\alpha}(\omega)$ и $\dot{\beta}(\omega)$, а также $\alpha(t)$ и $\beta(t)$ одинаковы. Из (IV.4.12а) и (IV.4.20а), а также из (IV.4.20а) и (IV.4.24) с учетом (IV.4.21), (IV.4.22) получаем

$$\begin{aligned} \dot{\beta}(\omega) &\xrightarrow{\omega \gg \omega_\alpha} \dot{\alpha}(\omega) \xrightarrow{\omega \gg \omega_\alpha} \alpha \omega_\alpha / i\omega; \\ \beta(t) &\xrightarrow{t \ll \tau_\alpha} \alpha(t) \xrightarrow{t \ll \tau_\alpha} \alpha t / \tau_\alpha. \end{aligned}$$

Физическое объяснение этого факта заключается в том, что при $\omega \gg \omega_\alpha$ или при $t \ll \tau_\alpha$ абсолютная величина коллекторного тока много меньше, чем базового и эмиттерного, которые, следовательно, практически одинаковы: $I_C \ll I_B \approx I_E$. Это означает, что транзистор «не чувствует», в какой схеме включения он находится — ОБ (входной ток I_E) или ОЭ (входной ток I_B).

IV.4.7. Диффузионные емкости в транзисторе

Как показано выше, инерционные свойства коэффициентов передачи α и β связаны с процессами накопления и рассасывания заряда избыточных носителей в базе. Эти свойства были учтены путем представления коэффициентов α и β в виде функций частоты управляющего сигнала или времени. Такой метод крайне неудобен для компьютерного анализа переходных процессов в линейных транзисторных схемах в связи с необходимостью непрерывного выполнения обратного преобразования Фурье, а для нелинейных устройств вообще не применим. Возможен иной подход, связанный с использованием понятия *диффузионной емкости*.

Такой подход был использован при анализе инерционных свойств диода (см. параграф II.7.3). В биполярном транзисторе идея Эберса и Молла разделения коллекторного и эмиттерного токов на инжектируемые и собираемые составляющие позволяет распространить его на произвольный режим работы.

На рис. IV.4.6 схематично показаны распределения неосновных носителей заряда в электронейтральных областях базы, эмиттера и коллектора для режима насыщения ($V_{be, bc} > 0$). Согласно модели Эберса–Молла распределения неосновных носителей, инжектированных в базу из эмиттера и коллектора, аддитивны. Следовательно, аддитивны и заряды избыточных носителей в базе Q_{BN} и Q_{BI} , инжектированные из эмиттера и коллектора.

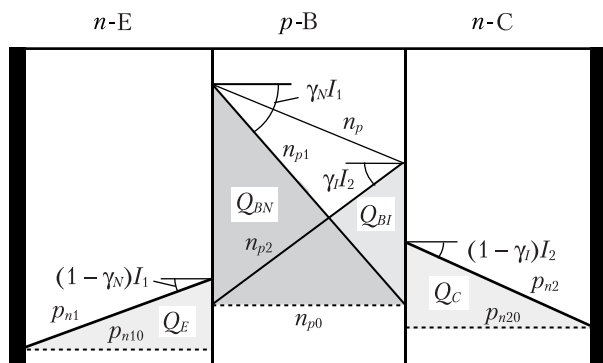


Рис. IV.4.6. Заряды Q_{BN} , Q_{BI} , Q_E и Q_C в диффузионных емкостях биполярного транзистора ($V_{be, bc} > 0$)

В нормальном режиме ($V_{bc} = 0$) отличны от нуля заряды Q_{BN} в базе и Q_E в эмиттере, зависящие от напряжения V_{be} . Эта ситуация анализировалась в параграфе II.7.3 в применении к полупроводниковому диоду. Полученные результаты — соотношения (II.7.18), (II.7.17а, б) — могут быть использованы и в транзисторе для определения диффузионной емкости C_{DN} , которая заряжена напряжением V_{be} . Отметим сразу, что в транзисторе эффективность эмиттера γ_N и коэффициент переноса κ_N всегда близки к 1. Поэтому зарядом Q_E , накопленным в эмиттере, можно пренебречь по сравнению с зарядом в базе ($Q_E \ll Q_{BN}$). При этом из (II.7.17а) получаем

$$C_{DN} = \gamma_N (I_1 + I_{1S}) / \varphi_T (T_N^{-1} + \tau_N^{-1}) = \alpha_N T_N (I_1 + I_{1S}) / \varphi_T,$$

где учтено, что $\alpha_N = \kappa_N \gamma_N$, $\kappa_N = 1 / (1 + T_N / \tau_N)$. Токи I_1 и I_{1S} относятся к эмиттерному диоду D_1 модели Эберса—Молла (рис. IV.2.2). Дифференциальное сопротивление диода D_1 составляет $r_1 = \varphi_T / (I_1 + I_{1S})$ (соотношение (II.4.37), параграф II. 4.7). Таким образом,

$$C_{DN} = \alpha_N T_N / r_1. \quad (\text{IV.4.27a})$$

Диффузионная емкость C_{BDI} соответствует заряду Q_{BI} и заряжена напряжением V_{bc} .

$$C_{BDI} = \alpha_I T_I / r_2, \quad (\text{IV.4.27б})$$

где T_I — время пролета электронов через базу в инверсном направлении (от коллекторного перехода до эмиттерного).

Накопление заряда Q_C в коллекторе (рис. IV.4.6) также приближенно может быть описано диффузионной емкостью коллектора C_{DC} , однако в планарных интегральных транзисторах, имеющих двухслойную структуру коллектора со скрытым n^+ -слоем (рис. IV.1.2) следует учитывать действие «встроенного» электрического поля на границе между n - и n^+ -слоями (см. раздел IV.3). Из соотношения (IV.3.19) следует, что диффузионная емкость коллектора составляет $C_{DC} = dQ_C / dV_{bc} = \tau_C dI_{2p} / dV_{be}$, или

$$C_{DC} = (1 - \gamma_I) \tau_C / r_2. \quad (\text{IV.4.27в})$$

В слаболегированном n -коллекторе время жизни дырок τ_C значительно превышает инверсное время пролета T_I , поэтому $C_{DC} \gg C_{BDI}$, и суммарная диффузионная емкость, связанная с инжекцией электронов и дырок через коллекторный переход, составляет

$$C_{DI} = C_{BDI} + C_{DC} \approx C_{DC} = (1 - \gamma_I) \tau_c / r_2. \quad (\text{IV.4.28})$$

В эквивалентной схеме транзистора диффузионные емкости C_{DN} и C_{DI} включены параллельно диодам D_1 и D_2 (см. рис. IV.2.2), поэтому произведения $r_1 C_{DN}$ и $r_2 C_{DI}$ есть постоянные времени заряда этих емкостей. Соотношение (IV.4.27) показывает, что постоянная времени $r_1 C_{DN} = \alpha_N T_N$ заряда емкости C_{DN} в точности равна постоянной времени коэффициента передачи эмиттерного тока $\tau_\alpha = \alpha_N T_N$ (IV.4.13). Отсюда следует, что *введение диффузионной емкости C_{DN} эквивалентно приближенному учету частотной зависимости коэффициента передачи эмиттерного тока α_N соотношением (IV.4.12a)*. Аналогично *диффузионная емкость C_{DI} учитывает частотную зависимость инверсного коэффициента передачи тока α_I* .

IV.4.8. Накопление заряда в коллекторе

Следует отметить, что в планарных интегральных транзисторах, имеющих двухслойную структуру коллектора (см. рис. IV.1.2), накопление заряда в коллекторе не может быть отражено корректно одной диффузионной емкостью C_{DC} , так как сравнительно слабо легированная n -область не может считаться эквипотенциальной. Кроме того, процесс осложняется тем, что в области перехода от n -коллектора к n^+ -слою формируется «встроенное» электрическое поле, способствующее накоплению заряда Q_C избыточных дырок в n -коллекторе.

Модель коллекторной области транзистора предложена в работе Кулла [4]. Мы не будем описывать ее подробно, ограничившись изложением метода анализа и его результатов и сохранив большинство обозначений [4]. Модель основана на следующих допущениях:

- 1) весь заряд Q_C сосредоточен в n -области;
- 2) n -коллектор электронейтрален;
- 3) рекомбинация в n -коллекторе отсутствует;
- 4) ток дырок I_p равен нулю;
- 5) граничные условия для n -области имеют вид:

$p_{cb} = p_n(x_{cb}) = p_{n0} e^{V_{bc}/\Phi_T}$; $p_{cw} = p_n(x_{cw}) = p_{n0} e^{V_{bcw}/\Phi_T}$, где смысл обозначений ясен из рис. IV.4.7, а, б.

Анализ производится в следующем порядке:

- 1) напряженность электрического поля находится из условия $I_p = 0$;



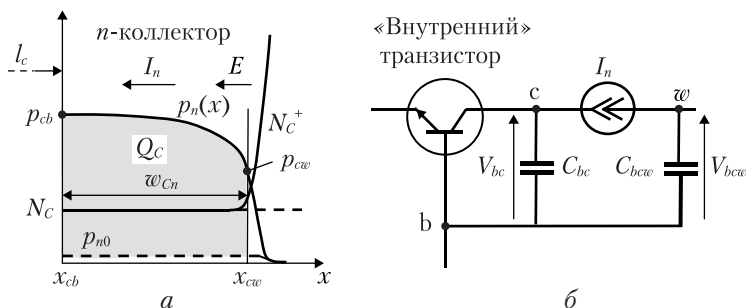


Рис. IV.4.7. Накопление заряда дырок в n -коллекторе Q_C при $V_{bc} > 0$ (а) и эквивалентная схема n -коллектора (б)

- 2) записывается уравнение для электронного тока I_n , включающее диффузионную и дрейфовую составляющие;
- 3) из уравнения для тока I_n находится распределение избыточных носителей заряда $\Delta n_n(x) = \Delta p_n(x)$ при заданных выше граничных условиях и ток I_n ;
- 4) вычисляется заряд избыточных дырок в коллекторе

$$Q_C(V_{bc}, V_{bcw}) = eS_E \int_{x_{cb}}^{x_{cw}} \Delta p_n(x) dx.$$

В результате анализа получены ВАХ n -коллектора $I_n(V_{bc}, V_{bcw})$ и вольт-кулонная характеристика $Q_C(V_{bc}, V_{bcw})$ в виде

$$I_n R_{c0} = \varphi_T \left[K(V_{bc}) - K(V_{bcw}) - \ln \frac{1 + K(V_{bc})}{1 + K(V_{bcw})} \right] + V_{bc} - V_{bcw}; \quad (\text{IV.4.29})$$

$$Q_C = Q_{bc}(V_{bc}) + Q_{bcw}(V_{bcw}); \quad (\text{IV.4.30})$$

$$Q_{bc} = Q_{c0} [K(V_{bc}) - 1 - \gamma/2]; \quad (\text{IV.4.31a})$$

$$Q_{bcw} = Q_{c0} [K(V_{bcw}) - 1 - \gamma/2], \quad (\text{IV.4.31б})$$

где $R_{c0} = \omega_{cn} / S_E e N_C \mu_n$ — равновесное сопротивление n -коллектора; $Q_{c0} = e S_E \omega_c N_C$ — равновесный заряд электронов в n -коллекторе; $K(V) = (1 + \gamma e^{V/\varphi_T})^{1/2}$, $\gamma = (2n_i / N_C)^2$.

Главным принципиально важным результатом анализа является соотношение (IV.4.30). Оно показывает, что накопленный в n -коллекторе заряд Q_C может быть представлен в виде суммы двух составляющих, каждая из которых зависит только от одного напряжения. Это позволяет синтезировать эквивалентную схему n -коллектора (см. рис. IV.4.7, б).

Схема содержит управляемый двумя напряжениями источник тока и две диффузионные емкости коллектора. Емкость $C_{bc}(V_{bc}) = dQ_{bc} / dV_{bc}$ включена между точками «внутренней» базы b и «внутреннего» коллектора c , напряжение между которыми является единственным аргументом заряда в емкости Q_{bc} . То же относится к емкости $C_{bcw}(V_{bcw}) = dQ_{bcw} / dV_{bcw}$, включенной между точками «внутренней» базы b и «наружного» коллектора w .

Следует отметить, что этот практически важный результат не вполне обоснован теоретически. В частности, не обоснованы допущения 2) об электронейтральности n -коллектора и граничные условия, принятые в допущении 5). Тем не менее ввиду принципиальной простоты модель Кулла используется в широко распространенных SPICE-подобных системах автоматизированного проектирования ИС. Проведенные автором модели эксперименты [4] показывают, что она хорошо описывает режим «квазинасыщения» транзистора, когда напряжение между внешними выводами базы и коллектора $V_{bc} < 0$, а напряжение на переходе база–коллектор $V_{bc} > 0$. Другие методы учета накопления избыточного заряда неосновных носителей в коллекторной области описаны в работах [5, 6].

Основные выводы

1. Инерционные свойства коэффициента передачи эмиттерного тока связаны в основном с накоплением заряда в базе, т.е. с инерционностью коэффициента переноса. Эффективность эмиттера практически не зависит от частоты.

2. Частотные и передаточные характеристики коэффициента передачи эмиттерного тока и коэффициента усиления базового тока приближенно описываются соотношениями $\dot{\alpha}(\omega) \approx \alpha / (1 + i\omega / \omega_\alpha)$, $\dot{\alpha}(s) \approx \alpha / (1 + s\tau_\alpha)$, $\dot{\beta}(\omega) \approx \beta / (1 + i\omega / \omega_\beta)$, $\dot{\beta}(s) \approx \beta / (1 + s\tau_\beta)$, где $\omega_\alpha = 1 / \tau_\alpha$, $\omega_\beta = 1 / \tau_\beta$.

3. Постоянные времени τ_α и τ_β связаны соотношением $\tau_\beta = (1 + \beta)\tau_\alpha$. Это означает, что при включении транзистора по схеме ОЭ усиление тока достигается ценой равной потери быстрогодействия.

4. Постоянная времени τ_α соответствует среднему времени пролета неосновных носителей заряда через базу T_N . При единичной эффективности эмиттера постоянная времени τ_β соответствует времени жизни неосновных носителей заряда в базе.



5. Инерционность коэффициентов передачи тока может быть учтена введением диффузионных емкостей, включенных параллельно эмиттерному и коллекторному переходам. В планарных транзисторах со скрытым n^+ -слоем в коллекторе большая часть инверсного заряда Q_i накапливается в n -коллекторе.

6. В режиме насыщения накопление заряда в коллекторе описывается моделью коллекторной области, содержащей управляемый двумя напряжениями источник тока и две диффузионные емкости коллектора.

IV.5. Действие факторов, не учтенных в идеализированной модели транзистора

IV.5.1. Факторы, не учтенные в идеализированной модели

Модель Эберса—Молла описывает свойства идеализированного транзистора и основана на выполнении ряда допущений, перечисленных в параграфе IV.2.1. В реальном транзисторе эти допущения нарушаются в той или иной степени. Наиболее существенны следующие факторы, не учтенные в идеализированной модели.

1. Электрически нейтральные области базы, эмиттера и коллектора имеют некоторые сопротивления R_B , R_E и R_C . Очевидным следствием является различие напряжений на эмиттерном и коллекторном p - n переходах V_{be} , V_{bc} и напряжений, приложенных к внешним выводам транзистора V_{BE} , V_{BC} . Этот фактор достаточно просто учитывается в моделях транзистора путем включения дополнительных сопротивлений в цепи внешних выводов, хотя и затрудняет интерпретацию экспериментальных результатов. Более существенным является то обстоятельство, что направление тока базы перпендикулярно плоскости p - n переходов. Это приводит к нарушению эквипотенциальности базы в плоскостях p - n переходов. В результате напряжения на переходах оказываются различными внутри транзистора. Для коллекторного перехода этот фактор сказывается на ВАХ только в режиме насыщения, так как при закрытом коллекторном переходе напряжение V_{bc} не влияет на ВАХ. Однако неоднородное распределение напряжения на эмиттерном переходе приводит к существенным нежелательным последствиям.



2. При малых значениях токов в кремниевых транзисторах становится существенной роль токов, обусловленных рекомбинацией—генерацией носителей заряда в области переходов. Особенно велика роль тока рекомбинации в эмиттерном переходе, так как этот ток, не учтенный в идеализированной модели, влияет на эффективность эмиттера и, следовательно, на коэффициент усиления тока.

3. При больших значениях токов нарушаются условия низкого уровня инжекции в слаболегированных областях базы и коллектора (в планарных транзисторах). В полупроводниковых диодах этот фактор приводит к искажению ВАХ, но в транзисторах он может также влиять на эффективность эмиттера и на коэффициент усиления тока.

4. При изменении напряжений на p - n переходах изменяется их ширина, что приводит к изменению толщины электро-нейтральной базы. Это в свою очередь вызывает изменение коэффициентов передачи тока. Особенно существенно изменение ширины коллекторного перехода, поскольку в нормальном режиме она сравнима с длиной базы, а напряжение на закрытом коллекторном переходе обычно изменяется в широких пределах.

5. В закрытом коллекторном переходе и на его границе с базой концентрация носителей заряда, создающих коллекторный ток, весьма мала. При большой плотности тока должна быть велика их скорость. Как показано в разделе I.4, дрейфовая скорость носителей заряда ограничена. Диффузионная скорость также в принципе ограничена тепловой скоростью движения частиц. Эффект ограничения скорости переноса носителей должен сказаться на усилительных и инерционных свойствах биполярного транзистора подобно тому, как он сказывается на характеристиках МДПТ.

Действие перечисленных факторов рассматривается в этом разделе.

IV.5.2. Эффект Эрли

Эффект Эрли состоит в модуляции длины базы напряжением на коллекторном переходе [7]. Механизм эффекта Эрли схематично представлен на рис. IV.5.1. При изменении напряжения V_{bc} на величину $dV_{bc} < 0$ коллекторный переход расширяется в сторону базы на величину $dl_c > 0$, что приводит к уменьшению длины базы на величину $dw_B < 0$. Действие эффекта Эрли проявляется на ВАХ транзистора в виде двух следствий:



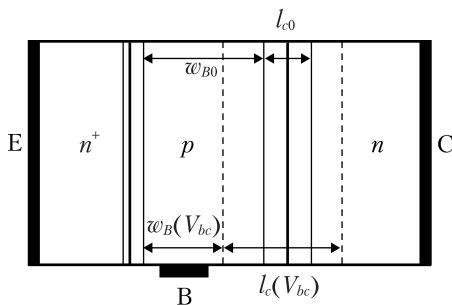


Рис. IV.5.1. Модуляция длины базы напряжением на коллекторном переходе (эффект Эрли): $V_{bc} < 0$

1) конечное выходное сопротивление в нормальном режиме;

2) внутренняя обратная связь.

Изменение длины базы приводит к изменению коэффициента передачи эмиттерного тока α_N и, следовательно, к изменению коэффициента усиления базового тока β_N . При увеличении запирающего напряжения $-V_{bc} > 0$ коллекторный переход расширяется, база становится короче и коэффициенты α_N , β_N возрастают. В результате при заданной величине тока эмиттера или базы возрастает ток коллектора.

В нормальном режиме при закрытом коллекторном переходе ($V_{bc} < -3\phi_T$) для фиксированного тока эмиттера (включение ОБ) ток коллектора составляет (IV.2.13а)

$$I_C = \alpha_N I_E + I_{C0}. \quad (IV.5.1)$$

Полагая коэффициент α_N постоянным, влияние эффекта Эрли можно учесть, положив конечным сопротивление закрытого коллекторного перехода r_c

$$I_C = \alpha_N I_E + I_{C0} - V_{bc} / r_c. \quad (IV.5.2)$$

Из (IV.5.1) следует, что сопротивление r_c определяется соотношением

$$\frac{1}{r_c} = \left. \frac{\partial I_C}{\partial (-V_{bc})} \right|_{dI_E=0} = -I_E \frac{d\alpha_N}{dV_{bc}}, \quad (IV.5.3)$$

где

$$\frac{d\alpha_N}{dV_{bc}} = \frac{d\alpha_N}{dw_B} \cdot \frac{dw_B}{dl_c} \cdot \frac{dl_c}{dV_{bc}}.$$

Для фиксированного тока базы (включение ОЭ) подстановка $I_E = I_C - I_B$ в (IV.5.2) дает

$$I_C = \frac{\alpha_N}{1 - \alpha_N} I_B + \frac{I_{C0}}{1 - \alpha_N} - \frac{V_{bc}}{r_c(1 - \alpha_N)},$$

или
$$I_C = \beta_N I_B + I_{C0}^* - V_{bc} / r_c^*, \quad (\text{IV.5.4})$$

где
$$r_c^* = r_c / (1 + \beta_N) \quad (\text{IV.5.5})$$

— сопротивление коллекторного перехода в схеме ОЭ.

Сопротивление r_c можно рассчитать, используя соотношения из параграфов II.3.4, IV.3.1 и IV.3.2 для определения ширины коллекторного перехода l_c и коэффициента передачи эмиттерного тока α_N . Мы оценим сопротивление r_c исходя из следующих простых соображений.

Пусть при малом напряжении $-V_{bc} = V_{bc0} \approx 3\phi_T$ значение коэффициента передачи эмиттерного тока составляет $\alpha_N(-V_{bc0}) = \alpha_{N0}$. При некотором напряжении $-V_{bc} = V_{pinch} > 0$ коллекторный переход смыкается с эмиттерным, длина базы становится равной нулю и $\alpha_N(-V_{pinch}) = 1$. Напряжение V_{pinch} называется *напряжением прокола базы*.

Из (IV.5.1) следует, что при повышении напряжения $-V_{bc}$ на величину $\Delta V_{bc} = V_{pinch} - V_{bc0} \approx V_{pinch}$ коллекторный ток повышается на величину $\Delta I_C = (1 - \alpha_{N0}) I_E$. Это означает, что сопротивления коллекторного перехода в схеме ОБ $r_c = \Delta V_{bc} / \Delta I_C$ и в схеме ОЭ $r_c^* = r_c / (1 + \beta_N) \approx r_c / (1 + \beta_{N0})$ (IV.5.5) составляют

$$r_c \approx V_{pinch} / (1 - \alpha_{N0}) I_E = (1 + \beta_{N0}) V_{pinch} / I_E; \quad (\text{IV.5.6a})$$

$$r_c^* \approx V_{pinch} / I_E, \quad (\text{IV.5.6б})$$

где $\beta_{N0} = \alpha_{N0} / (1 - \alpha_{N0})$.

В области базы расположена часть $l_{cB} = \theta_B l_c$ полной ширины коллекторного перехода l_c , которая определяется соотношением (II.3.9)

$$l_c(V_{bc}) = l_{c0} \sqrt[k]{(\phi_{BC} - V_{bc}) / \phi_{BC}}, \quad (\text{IV.5.7})$$

где ϕ_{BC} — контактная разность потенциалов p - n перехода база — коллектор. При напряжении $-V_{bc} = V_{bc0}$ длина базы составляет $w_B(-V_{bc0}) \approx w_B(0) = w_{B0}$ (см. рис. IV.5.1), а ширина коллекторного перехода $l_c(-V_{bc0}) \approx l_{c0}$. При напряжении прокола базы $-V_{bc} = V_{pinch}$ длина базы $w_B(-V_{pinch}) = 0$, а ширина коллекторного перехода $l_c(-V_{pinch}) \approx l_{c0} + \theta w_{B0}$. Подставляя это значение в (IV.5.7), получаем: $l_{c0} + \theta w_{B0} = l_{c0} \sqrt[k]{(\phi_{BC} + V_{pinch}) / \phi_{BC}}$. Отсюда



$$V_{pinch} \approx \varphi_{BC} \left[(1 + \theta w_{B0} / l_{c0})^{1/k} - 1 \right]. \quad (IV.5.8)$$

Если коллекторный переход является ступенчатым, $k=2$, $\theta = 1 / (1 + N_B / N_C)$. Для линейного (симметричного) перехода $k=3$, $\theta=0,5$. Для планарного транзистора с линейным коллекторным переходом при $w_{B0} = 0,15$ мкм, $l_{c0} = 0,1$ мкм и $\varphi_{BC} = 0,75$ В напряжение прокола базы составляет $V_{pinch}^* \approx 3,3$ В. При $I_E = 0,1$ мА и $\beta_{N0} = 70$ из (IV.5.6а, б) получаем

$$r_c^* \approx 33 \text{ кОм}; r_c \approx 230 \text{ кОм}.$$

Влияние эффекта Эрли на выходные ВАХ транзистора при включении ОБ и ОЭ иллюстрируется рис. IV.5.2. Модуляция длины базы напряжением на коллекторном переходе приводит к тому, что в нормальном режиме *транзистор не является идеальным источником тока*, как это следует из модели Эберса—Молла. Соотношения (IV.5.6а, б) показывают, что *сопротивления коллекторного перехода r_c и r_c^* обратно пропорциональны току эмиттера*. Отсюда следует, что прямые, экстраполирующие выходные ВАХ до нулевого тока коллектора, должны пересечь ось напряжения в одной точке при напряжении, абсолютное значение которого называется *напряжением Эрли V_{IRL}* . Для включения ОБ наклон выходных характеристик $1/r_c$ весьма мал и напряжение Эрли очень велико. Для включения ОЭ наклон выходных характеристик $1/r_c^*$ больше в $1 + \beta_N$ раз (IV.5.5), и напряжение Эрли можно определить, используя экспериментальные данные. Напряжение Эрли может служить единственным параметром, который полностью характеризует эффект Эрли.

Из рис. IV.5.2, б следует, что

$$V_{IRL} = I_C(0)r_c^*, \quad (IV.5.9)$$

где ток $I_C(0)$ соответствует пересечению прямой, экстраполирующей выходную ВАХ в схеме ОЭ с осью тока. Поскольку $\beta_N \gg 1$, в (IV.5.6б) можно приближенно положить $I_E \approx I_C(0)$. Сравнение (IV.5.6б) и (IV.5.9) приводит к выводу, что при включении ОЭ *напряжение Эрли близко к напряжению прокола базы*.

Следует отметить, что приведенные выкладки имеют приближенный характер. В частности, ниоткуда не следует, что выходные ВАХ линейны. В схеме ОБ это действительно очень близко к истине, так как коэффициент α_N в (IV.5.2) изменяется незначительно ($\alpha_{N0} \leq \alpha_N \leq 1$). В схеме ОЭ выходные ВАХ не могут быть линейными. Действительно, при $-V_{bc} \rightarrow V_{pinch}$ ($V_{ce} = -V_{bc} + V_{be} \rightarrow V_{pinch} + V_{be}$) коэффициент передачи

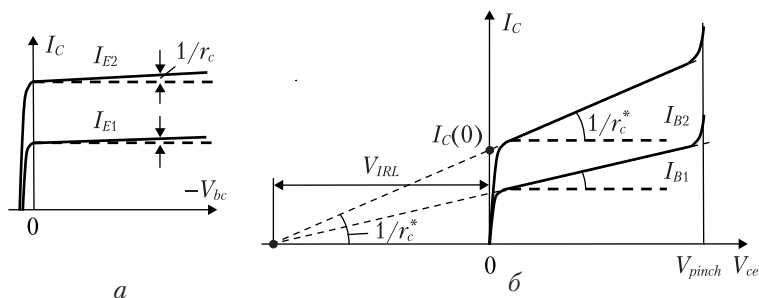


Рис. IV.5.2. Влияние эффекта Эрли на выходные ВАХ транзистора при включении ОБ (а) и ОЭ (б)

$\alpha_N \rightarrow 1$, $\beta_N \rightarrow \infty$ и при фиксированном значении тока базы $I_C \rightarrow \infty$ (рис. IV.5.2, б). Это означает, что в схеме ОЭ *прокол базы является видом пробоя*; в этом режиме транзистор неработоспособен. В этом смысле эффект прокола базы близок к эффекту смыкания канала в МДП-транзисторах (см. параграф III.5.3).

Прокол базы накладывает довольно жесткие ограничения на возможности уменьшения длины базы. В приведенном выше примере при $w_{B0} = 0,15$ мкм и $l_{c0} = 0,1$ мкм напряжение прокола составило 3,3 В. Если уменьшить длину базы до 0,1 мкм либо увеличить равновесную ширину коллекторного перехода до 0,15 мкм, то в соответствии с (IV.5.8) напряжение прокола снизится до 0,8 В, что в большинстве случаев неприемлемо. Таким образом, уменьшение длины базы возможно только при достаточно малой ширине коллекторного перехода, т.е. при повышении степени легирования базы и коллектора.

Отметим основные моменты, характеризующие деформацию выходных ВАХ транзистора под действием эффекта Эрли.

1. Выходные ВАХ транзистора имеют конечное выходное сопротивление.

2. В схеме ОЭ выходное сопротивление r_c^* не зависит от коэффициента усиления базового тока β_N ; в схеме ОБ выходное сопротивление r_c в $1 + \beta_N$ больше.

3. Прямые, экстраполирующие выходные ВАХ до нулевого тока коллектора, пересекают ось напряжения при напряжении Эрли V_{IRL} .

4. В схеме ОЭ напряжение Эрли близко к напряжению прокола базы V_{pinch} .

5. В схеме ОЭ прокол базы является видом пробоя.

Вторым следствием эффекта Эрли является существование в транзисторе *внутренней обратной связи*. Обратная связь проявляется в том, что при фиксированном токе эмиттера I_E изменение напряжения база—коллектор V_{bc} вызывает изменение напряжения база—эмиттер V_{be} . Механизм внутренней обратной связи поясняется рис. IV.5.3, где принято $V_{be} > 0$; $V_{bc} < -3\varphi_T$. В соответствии с граничными условиями Шокли (II.3.7) напряжения на переходах определяют граничные концентрации неосновных носителей

$$n_p(0) = n_{p0} e^{V_{be}/\varphi_T} > n_{p0}; \quad (\text{IV.5.10a})$$

$$n_p(w_B) = n_{p0} e^{V_{bc}/\varphi_T} = 0. \quad (\text{IV.5.10б})$$

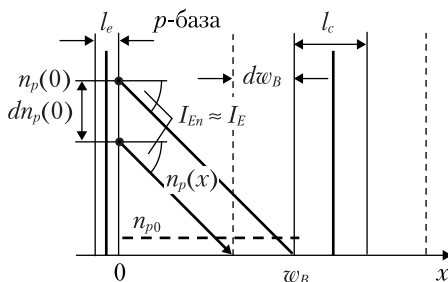


Рис. IV.5.3. Механизм внутренней обратной связи как следствия эффекта Эрли.

На рисунке принято $V_{be} > 0$, $V_{bc} < 0$, $I_{En} \approx I_E = \text{const}$

При увеличении обратного напряжения $-V_{bc}$ коллекторный переход расширяется, что приводит к уменьшению длины базы. В нормальном режиме ток через эмиттерный переход практически полностью является электронным

$$I_{En} = \gamma_N I_E \approx I_E.$$

Как не раз отмечалось выше, фиксированному значению тока I_{En} соответствует фиксированный наклон практически линейной функции $n_p(x)$ при $x = 0$ (см. рис. IV.5.3). Поэтому при сужении базы снижается граничная концентрация $n_p(0)$, что в соответствии с (IV.5.10a) означает снижение напряжения V_{be} .

Глубина обратной связи характеризуется *коэффициентом обратной связи*:

$$\mu = \left. \frac{\partial V_{be}}{\partial V_{bc}} \right|_{dI_E=0} = \frac{dV_{be}}{dn_p(0)} \cdot \frac{dn_p(0)}{d\varpi_B} \cdot \frac{d\varpi_B}{dV_{bc}}. \quad (\text{IV.5.11})$$

Из (IV.5.10а) следует, что $dV_{be}/dn_p(0) = \varphi_T/n_p(0)$, а поскольку функция $n_p(x)$ линейна, $dn_p(0)/d\varpi_B = n_p(0)/\varpi_B$ (см. рис. IV.5.3). При этом из (IV.5.11) получаем

$$\mu = \frac{\varphi_T}{\varpi_B} \cdot \frac{d\varpi_B}{dV_{bc}}. \quad (\text{IV.5.12})$$

Обратная связь является *отрицательной*, так как при повышении выходного напряжения $-V_{bc}$ входное напряжение V_{be} снижается. Типичные значения коэффициента обратной связи составляют 10^{-4} – 10^{-3} , и при практических расчетах в большинстве случаев действием обратной связи пренебрегают.

На рис. IV.5.4 показана эквивалентная схема Эберса—Молла (см. рис. IV.2.2) для нормального режима работы транзистора ($V_{be} > 0$; $V_{bc} < -3\varphi_T$), учитывающая эффект Эрли. В соответствии с уравнением (IV.5.2) генератор тока $\alpha_N I_1$ и диод D_2 заменены генераторами тока $\alpha_N I_E$ и I_{C0} . Генератор малого тока $\alpha_I I_2 = -\alpha_I I_{2S} = -\alpha_N I_{1S} \ll I_1 \approx I_E$ из цепи эмиттера исключен. Эффект модуляции длины базы учтен сопротивлением коллекторного перехода r_c , а внутренняя обратная связь — генератором ЭДС μV_{bc} , включенным в цепь эмиттера последовательно.

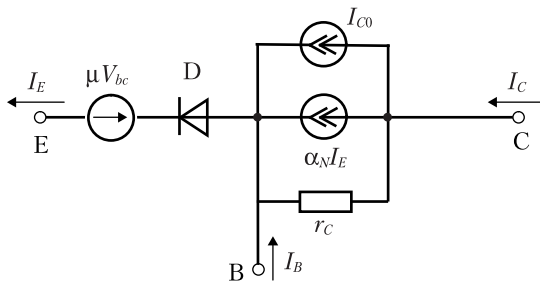


Рис. IV.5.4. Эквивалентная схема Эберса—Молла для нормального режима работы транзистора ($V_{be} > 0$; $V_{bc} < -3\varphi_T$), учитывающая эффект Эрли

IV.5.3. Сопротивления базы и коллектора

Линии тока базы и коллектора в планарном транзисторе для нормального режима схематично показаны на рис. IV.5.5.

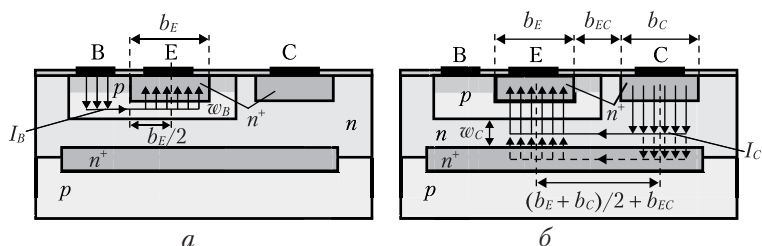


Рис. IV.5.5. Линии тока базы (а) и коллектора (б) в планарном транзисторе (нормальный режим)

Отметим, что для всех составляющих тока базы (см. рис. IV.1.3) линии тока одинаковы. Сопротивление базы R_B складывается из сопротивлений *активной* базы R_{B1} (под эмиттером, рис. IV.5.5, а) длиной b_E с удельной электропроводностью $\sigma_B = eN_B \mu_{Bp}$ и остальной *пассивной* части R_{B2} . Основной вклад вносит сопротивление тонкой и сравнительно слабо легированной активной базы (см. рис. IV.5.5, а), которое можно оценить формулой

$$R_{B1} = b_E / 2\sigma_B w_B Z_E,$$

где Z_E — размер базы, нормальный к плоскости рисунка. Обычно сопротивление базы лежит в диапазоне от 0,3 до 2 кОм. Сопротивление базы повышается с ростом запирающего напряжения на коллекторном переходе вследствие эффекта Эрли. Для снижения сопротивления базы могут использоваться два омических контакта, расположенных по разные стороны эмиттера. При этом сопротивление активной базы уменьшается в четыре раза, а пассивной — вдвое.

При отсутствии скрытого n^+ -слоя в коллекторе сопротивление коллектора R_C определяется сопротивлением тонкого горизонтального на рис. IV.5.5, б n -слоя под областью базы, в котором концентрация легирующей примеси N_C минимальна (см. табл. IV.1.1). В этом случае значение R_C может составлять 100–200 Ом. При введении n^+ -слоя линии тока в высокоомном n -слое располагаются вертикально (см. рис. IV.5.5, б), и сопротивление коллектора составляет от 5 до 20 Ом. Следует отметить, что при использовании двух омических контактов к

базе сопротивление коллектора возрастает, так как увеличивается расстояние b_{EC} до эмиттера (см. рис. IV.5.5, б). В мощных транзисторах, рассчитанных на большой ток коллектора, для снижения сопротивления R_C также могут быть созданы два контакта по разные стороны эмиттера либо прямоугольный контакт вокруг p -области базы.

IV.5.4. Эффект оттеснения эмиттерного тока

Эффект оттеснения эмиттерного тока обусловлен нарушением эквипотенциальности базы в плоскости, параллельной плоскости эмиттерного перехода. Он проявляется при значительных токах эмиттера и, следовательно, базы, когда падение напряжения на сопротивлении активной базы R_{B1} сравнимо с температурным потенциалом ϕ_T . На рис. IV.5.6 показаны линии тока базы, а также эпюры распределений потенциала базы (а) и плотности тока через эмиттерный переход (б) в нормальном режиме работы транзистора. Кривые 1, 2, 3 соответствуют токам эмиттера $I_{E1} < I_{E2} < I_{E3}$. Потенциал эмиттера принят равным нулю.

При малом токе эмиттера $I_E = I_{E1}$ ток базы I_{B1} также мал, и падение напряжения ΔV_{b1} от тока I_{B1} на активном сопротивлении базы R_{B1} мало: $\Delta V_{b1} \ll \phi_T$. База остается эквипотенциальной в направлении оси y (рис. IV.5.6, а, прямая 1). Напряжение база—эмиттер $V_{be} = V_b$ во всех точках эмиттерного перехода одинаково, следовательно, плотность эмиттерного тока j_e также одинакова (рис. IV.5.6, б, прямая 1) — эмиттер инжектирует носители заряда в базу однородно по всей поверхности эмиттерного перехода.

При увеличении тока I_E ток базы I_B возрастает. Вследствие падения напряжения на сопротивлении R_{B1} потенциал базы V_b снижается по мере удаления от базового контакта (рис. IV.5.6, а, кривые 2, 3). Плотность эмиттерного тока экспоненциально зависит от напряжения V_{be} : $j_e \sim e^{V_{be}/\phi_T}$. Поэтому неравномерность распределения плотности тока $j_e(y)$ значительно больше, чем неравномерность распределения потенциала базы $V_b(y)$ (см. рис. IV.5.6, б, кривые 2, 3). Таким образом, плотность эмиттерного тока максимальна на краю эмиттера, ближайшем к базовому контакту. Ток эмиттера с возрастанием его величины как бы оттесняется к базовому контакту. Если конструкция транзистора содержит два контакта к базе, расположенных по разные стороны эмиттера, ток эмиттера оттесняется из середины эмиттера к обоим контактам симметрично.

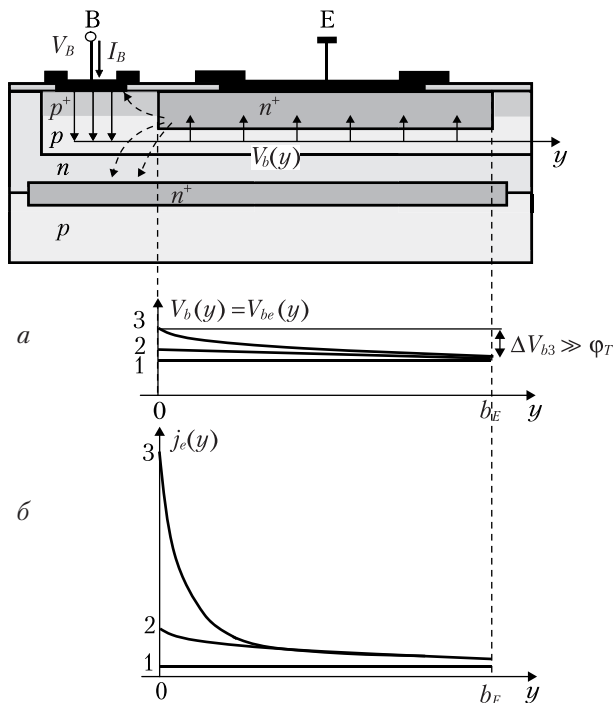


Рис. IV.5.6. Линии тока базы и эшоры распределений потенциала базы (а) и плотности тока через эмиттерный переход (б) в нормальном режиме работы транзистора.

Кривые 1, 2, 3 соответствуют токам эмиттера $I_{E1} < I_{E2} < I_{E3}$

Эффект оттеснения начинает проявляться, когда произведение $I_B R_{B1}$ становится сравнимым с величиной температурного потенциала ϕ_T . Оттеснение эмиттерного тока приводит к тому, что при высоком уровне тока I_E значительная часть этого тока инжектируется через боковые стенки эмиттера вблизи контакта к базе (пунктир на рис. IV.5.6).

Следует отметить, что эффект оттеснения эмиттерного тока является следствием принципиальной неоднородности линий тока в транзисторе — направление дырочного тока базы в активной базе перпендикулярно направлению электронного тока коллектора. Поэтому этот эффект практически не удастся адекватно учесть коррекцией эквивалентной схемы.

IV.5.5. Влияние режима работы транзистора на коэффициенты передачи тока

В нормальном режиме работы транзистора его электрическое состояние определяется двумя электрическими аргументами, в качестве которых принято использовать ток эмиттера I_E и напряжение коллектор—база V_{bc} . Действие факторов, не учтенных в идеализированной модели транзистора, приводит к тому, что эти аргументы влияют на величину коэффициентов передачи тока.

Влияние напряжения V_{bc} на коэффициенты α_N и β_N проявляется через эффект Эрли и учитывается коррекцией эквивалентной схемы транзистора (см. параграф IV.5.2). Значительно сложнее учесть зависимости $\alpha_N(I_E)$ и $\beta_N(I_E)$. Экспериментальные исследования показывают, что коэффициент усиления базового тока β_N снижается как при уменьшении, так и при достаточном увеличении тока эмиттера, достигая расчетных значений (см. раздел IV.3) при некоторой оптимальной плотности эмиттерного тока $\approx (1-5) \cdot 10^3 \text{ А/см}^2$ (рис. IV.5.7). Зависимость $\beta_N(I_E)$ определяется целым рядом факторов, действие которых при низкой и высокой плотности тока проявляется различным образом.

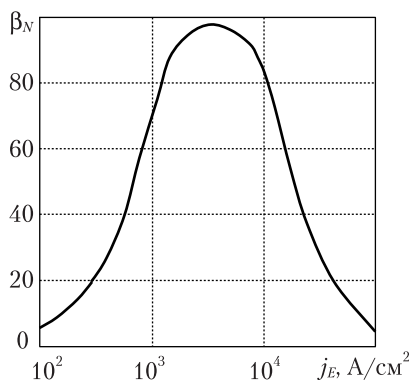


Рис. IV.5.7. Типичная зависимость коэффициента усиления базового тока от плотности эмиттерного тока для кремниевого транзистора

При низкой плотности эмиттерного тока определяющим является влияние на эффективность эмиттера тока рекомбинации—генерации I_{Erg} в области эмиттерного перехода (см. рис. IV.1.3), не учтенного в идеализированной модели тран-

зистора. Эффективность эмиттера определена соотношением (IV.1.8), которое при $V_{bc} = 0$ можно представить в виде

$$\gamma_N = \frac{\gamma_{N0}}{1 + I_{Erg} / (I_{En} + I_{Ep})} = \frac{\gamma_{N0}}{1 + I_{Erg} / I_1}, \quad (IV.5.13)$$

где коэффициент $\gamma_{N0} = I_{En} / (I_{En} + I_{Ep})$ соответствует определению эффективности эмиттера в идеализированной модели, а $I_1 = I_{En} + I_{Ep}$ — ток диода D_1 в модели Эберса—Молла (см. рис. IV.2.2), ВАХ которого имеет вид (IV.2.6а). При $V_{be} > 3\phi_T$

$$I_1 = I_{1S} e^{V_{be} / \phi_T}. \quad (IV.5.14a)$$

Это выражение получено для низкого уровня инжекции и в нашем случае низкой плотности эмиттерного тока безусловно справедливо. ВАХ тока рекомбинации—генерации I_{Erg} при $V_{be} > 3\phi_T$ определена соотношением (II.5.11)

$$I_{Erg} \approx I_{Er0} e^{V_{be} / m\phi_T}, \quad (IV.5.14б)$$

где $m = 2$. Соотношения (IV.5.14а, б) показывают, что при уменьшении напряжения V_{be} , т.е. тока эмиттера, ток инжекции

$$I_1 = I_{En} + I_{Ep} \quad (IV.5.15)$$

уменьшается быстрее, чем ток рекомбинации I_{Erg} , что и приводит к потере усиления.

При *высокой плотности эмиттерного тока* ток рекомбинации I_{Erg} пренебрежимо мал и для эффективности эмиттера остается справедливым соотношение

$$\gamma_N = 1 / (1 + I_{Ep} / I_{En}). \quad (IV.5.16)$$

Снижение усилительных свойств транзистора при этом обусловлено двумя главными факторами: повышением уровня инжекции в базе и эффектом оттеснения эмиттерного тока.

При высоком уровне инжекции нарушается граничное условие Шокли на границе базы с эмиттерным переходом ($x = 0$ на рис. IV.5.3) и граничное значение концентрации электронов определяется соотношением (II.3.8б): $n_p(0) = n_i e^{V_{be} / 2\phi_T}$. Такой же характер имеет ВАХ тока неосновных для базы носителей — формула (II.5.27)

$$I_{En} = (2en_i D_B / w_B) e^{V_{be} / 2\phi_T}.$$

В эмиттере уровень инжекции всегда остается низким и

$$I_{Ep} = I_{1Sp} e^{V_{be} / \phi_T}.$$



Таким образом, при повышении напряжения V_{be} , т.е. тока эмиттера, дырочный ток растет быстрее, чем электронный: $I_{Ep} / I_{En} \sim e^{V_{be}/2\Phi_T}$. Согласно (IV.5.16) это приводит к снижению эффективности эмиттера и коэффициентов α_N , β_N .

Эффект оттеснения эмиттерного тока приводит к целому ряду нежелательных последствий. Во-первых, при оттеснении эмиттерного тока к контакту базы часть электронов инжектируется в пассивную базу (пунктир на рис. IV.5.6). Для этих электронов увеличивается время пролета до коллекторного перехода T_N либо они могут вообще рекомбинировать на поверхности пассивной базы, не достигнув коллекторного перехода. Во-вторых, поверхность пассивной базы легирована сильнее, чем активный объем, поэтому время жизни инжектированных в эту область электронов меньше. Эти факторы приводят к увеличению среднего значения T_N и снижению среднего времени жизни τ_N , в результате чего снижается коэффициент переноса $\kappa_N = 1 / (1 + T_N / \tau_N)$.

Наконец, оттеснение тока к периферии эмиттерного перехода снижает его эффективную площадь, что приводит к повышению плотности тока на периферии эмиттерного перехода. В результате в краевой области активной базы, через которую протекает основная часть тока, уровень инжекции повышается при меньшем значении эмиттерного тока, что ведет к снижению эффективности эмиттера.

Все эти факторы объясняют снижение усилительных свойств транзистора при повышении $I_{Ep} / I_{En} \sim e^{V_{be}/2\Phi_T}$ и плотности эмиттерного тока (см. рис. IV.5.7).

IV.5.6. Ограничение скорости носителей заряда в базе и в коллекторном переходе

В общем случае плотность электронного тока, независимо от его механизма, описывается выражением

$$j_n = -en\bar{v}_n, \quad (\text{IV.5.17a})$$

где $\bar{v}_n$ — средняя скорость электронов в направлении тока. Плотность диффузионного тока электронов определяется соотношением

$$j_n = eD_B dn / dx, \quad (\text{IV.5.176})$$

Сравнение выражений (IV.5.17a) и (IV.5.176) показывает, что *диффузионная скорость* электронов в направлении оси x составляет

$$v_{n \text{ dif } x} = -\frac{D_n}{n} \frac{dn}{dx}. \quad (\text{IV.5.18})$$

В модели идеализированного транзистора ток неосновных носителей в p -базе считается диффузионным, причем на границах с ОПЗ используются граничные условия Шокли

$$n_p(0, w_B) = n_{p0} e^{V_{bc}/\Phi_T}. \quad (\text{IV.5.19a})$$

Нарушение этого условия на границе с эмиттерным переходом ($x=0$) при высоком уровне инжекции уже обсуждалось, однако и при низком уровне инжекции оно может быть некорректным. В частности, на границе с коллекторным переходом ($x=w_B$) в нормальном режиме работы транзистора при $V_{bc} < -3\Phi_T$ полагалось $n_p(w_B)=0$. Как следует из (IV.5.18), диффузионная скорость электронов $v_{n \text{ dif }}(w_B)$ при этом бесконечна. Абсурдность этого вывода объясняется тем, что соотношение (IV.5.17б) является феноменологическим. На микроскопическом уровне диффузионная скорость электронов есть средняя скорость их теплового движения в направлении снижения их концентрации. При $n_p(w_B)=0$ она не может превышать среднюю скорость теплового движения в выбранном направлении

$$v_{Tn} = \sqrt{3kT/m_n} \approx 10^7 \text{ см/с}.$$

Через коллекторный переход электроны переносятся электрическим полем, средняя напряженность которого $\bar{E} = (\Phi_{bc} - V_{bc})/l_c$ при $V_{bc} < 0$ обычно составляет не менее $(3-7) \cdot 10^4$ В/см. В таких полях скорость электронов можно считать равной предельной дрейфовой скорости (см. раздел I.4):

$$v_{Sn} \approx 10^7 \text{ см/с}.$$

Совпадение значений v_{Tn} и v_{Sn} , по-видимому, случайно, тем более, что их температурные зависимости различны. Однако для простых качественных оценок можно положить скорости электронов в плоскости $x=w_B$ и в коллекторном переходе равными $v_m \approx 10^7$ см/с. Тогда в нормальном режиме работы транзистора ($V_{be} > 0$, $V_{bc} < -3\Phi_T$) как на границе базы с коллекторным переходом, так и в самом коллекторном переходе концентрация электронов составит

$$n_p(w_B) = n_{bc} = j_n / ev_m. \quad (\text{IV.5.19б})$$

Распределения электронов в базе и в коллекторе, соответствующие условию (IV.5.19б), представлены на



рис. IV.5.8. Для сравнения штриховыми линиями показаны те же распределения без учета ограничения скорости электронов. Из рис. IV.5.8 и соотношения (IV.5.196) следует, что величина коллекторного тока $I_C = -eD_B dn/dx(w_B)$ составляет

$$I_C = eS_E D_B [n_p(0) - n_{bc}] / w_B = eS_E n_{bc} v_m, \quad (\text{IV.5.20})$$

а суммарный заряд избыточных электронов в базе и коллекторном переходе

$$Q_N + \Delta Q_N + Q_{bc} = eS_E \left[\frac{w_B}{2} [n_p(0) - n_{bc}] + n_{bc}(w_b + l_c) \right]. \quad (\text{IV.5.20})$$

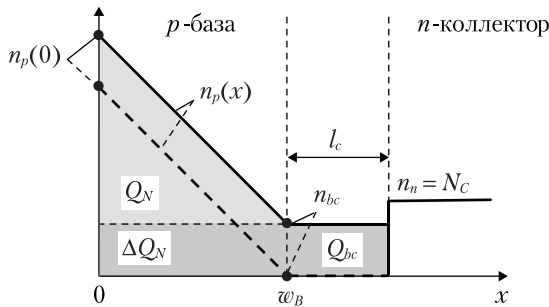


Рис. IV.5.8. Распределения электронов в базе и в коллекторе в нормальном режиме работы транзистора ($V_{be} > 0$, $V_{bc} < -3\phi_T$) с учетом ограничения скорости электронов: $n_{bc} = j_n / ev_m$.

Штриховые линии — без учета ограничения скорости электронов

Таким образом, время пролета электронов до коллектора через базу и коллекторный переход $T_N = (Q_N + \Delta Q_N + Q_{bc}) / I_C$ составляет

$$T_N = \frac{w_B^2}{2D_B} + \frac{w_B + l_c}{v_m}. \quad (\text{IV.5.21a})$$

Первое слагаемое в правой части (IV.5.21a) соответствует времени диффузии через базу без учета ограничения диффузионной скорости электронов, а второе — времени пролета через базу и коллекторный переход с максимальной скоростью v_m .

Логично учесть различие в предельных скоростях диффузии и дрейфа, положив

$$T_N = \frac{w_B^2}{2D_B} + \frac{w_B}{v_{Tn}} + \frac{l_c}{v_{Sn}}. \quad (\text{IV.5.21б})$$

Следует, однако, отметить, что соотношение (IV.5.21б) можно получить, пользуясь рисунком IV.5.8, только допустив разрыв функции $n(x)$ в точке $x = w_B$.

При очень малой длине базы ($w_B \ll 2D_B / v_m$)

$$T_N \xrightarrow{w_B v_m / 2D_B \ll 1} (w_B + l_c) / v_m. \quad (\text{IV.5.21в})$$

В соответствии с соотношением (IV.4.27а) пропорционально времени пролета T_N возрастает и диффузионная емкость C_{DN} , отражающая накопление заряда неосновных носителей в базе.

Менее важное следствие ограниченной скорости носителей в биполярном транзисторе состоит в некоторой деформации его входной характеристики $I_E(V_{be})$ в нормальном режиме. На рис. IV.5.8 показано, что при заданном значении напряжения база—эмиттер, т.е. заданном значении $n_p(0)$, эмиттерный ток (пропорциональный dn_p/dx при $x = 0$) снижается в $1 + D_B / w_B v_m$ раз. При $V_{be} > 3\phi_T$ это соответствует сдвигу входной характеристики $I_E(V_{be})$ в сторону меньших напряжений на величину

$$\Delta V_{be} = \phi_T \ln(1 + D_B / w_B v_m). \quad (\text{IV.5.22})$$

Отметим, что эффект ограничения скорости носителей проявляется тем сильнее, чем короче база. Согласно (IV.5.21) и (IV.5.22) характерным параметром, определяющим важность его учета, является отношение $D_B / w_B v_m$.

Еще один эффект связан с ограничением дрейфовой скорости электронов в коллекторе — *эффект Кирка* [8]. Эффект Кирка проявляется в планарных транзисторах, имеющих сравнительно слабо легированный n -коллектор, когда плотность коллекторного тока в нормальном режиме превышает критическое значение

$$j_{CS} = eN_c v_{Sn}. \quad (\text{IV.5.23})$$

При $j_c > j_{CS}$ n -коллектор не может пропустить столь большой ток, оставаясь электронейтральным, поскольку концентрация электронов $n_c = j_c / ev_{Sn}$, движущихся через коллектор с предельной дрейфовой скоростью v_{Sn} , превышает концентрацию примеси N_c (рис. IV.5.9, а). В результате помимо избыточного заряда в коллекторном переходе

$$Q_{bc} = eS_E n_c l_c = S_E l_c j_c / v_{Sn}$$



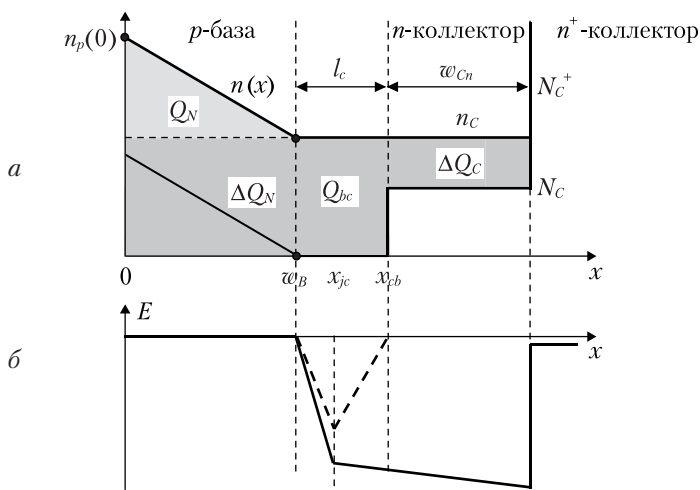


Рис. IV.5.9. Эффект Кирка:

a — распределение концентрации электронов в базе, коллекторном переходе и коллекторе;

б — распределение напряженности электрического поля.

Штриховые линии — без учета ограничения скорости электронов

в *n*-коллекторе также накапливается избыточный заряд

$$\Delta Q_C = S_E w_{Cn} (j_C - j_{CS}) / v_{Sn},$$

где w_{Cn} — длина *n*-коллектора (см. рис. IV.5.9, *a*). Нарушение электронейтральности *n*-коллектора можно трактовать как расширение ОПЗ коллекторного перехода до границы с *n*⁺-коллектором (рис. IV.5.9, *б*). Заметим, что расширение ОПЗ коллекторного перехода под действием эффекта Кирка вызвано не повышением запирающего напряжения на коллекторном переходе, а повышением плотности тока.

Как показывает рис. IV.5.9, следствием эффекта Кирка является заполнение коллекторного перехода и *n*-коллектора электронами, инжектированными из эмиттера, в результате чего резко возрастает диффузионная емкость C_{DN} и снижается быстродействие транзистора. Иногда эффект Кирка трактуют как расширение базы до границы с *n*⁺-коллектором.

Предотвращение эффекта Кирка накладывает дополнительные ограничения на режим работы транзистора при проектировании ИС.

IV.5.7. Специфика пробоя в биполярных транзисторах

Если любой из переходов транзистора использовать в качестве диода, то пробой этого перехода при обратном напряжении будет иметь такой же характер, как в диоде (см. раздел II.6). В планарных транзисторах база легируется достаточно сильно, особенно в поверхностном слое кристалла. Поэтому для эмиттерного перехода характерен туннельный пробой при напряжении от 1 до 4 В.

В коллекторном переходе обычно реализуется лавинный механизм пробоя, так как n -коллектор легирован сравнительно слабо. Вероятность теплового пробоя в кремниевых транзисторах мала ввиду малых обратных токов коллекторного перехода. Его можно наблюдать только при высокой температуре либо при очень плохом теплоотводе.

При использовании транзистора по прямому назначению эмиттерный переход редко подвергается опасности пробоя, хотя для некоторых применений такую возможность следует иметь в виду. Мы не рассматриваем здесь специфику работы транзистора в мощных силовых устройствах, однако даже в логических вентилях типа ТТЛ на эмиттерный переход входного многоэмиттерного транзистора подаются обратные напряжения в несколько вольт.

При работе транзистора в нормальном режиме для пробоя коллекторного перехода транзистора характерны некоторые особенности. Эти особенности проявляются в зависимости пробивного напряжения от тока эмиттера или базы, а также от схемы включения транзистора. Кроме того, для транзистора характерен специфический механизм пробоя — прокол базы, который уже рассмотрен в параграфе IV.5.2.

При включении транзистора по схеме ОБ ток коллектора в нормальном режиме определяется соотношением (IV.2.13а)

$$I_C = \alpha_N I_E + I_{C0}. \quad (IV.5.24a)$$

При наличии ударной ионизации электронно-дырочных пар в коллекторном переходе увеличивается любая составляющая коллекторного тока независимо от ее природы. Поэтому соотношение (IV.5.24а) можно записать в виде

$$I_C = \alpha_M(V_{bc})(\alpha_N I_E + I_{C0}), \quad (IV.5.24б)$$

где $\alpha_M(V_{bc})$ — коэффициент умножения тока в коллекторном переходе. Для лавинного механизма пробоя коэффициент умножения соответствует полуэмпирической формуле (II.6.6)



$$\alpha_M(V_{bc}) = 1/[1 - (-V_{bc}/V_{brI})^\nu]. \quad (IV.5.25)$$

При напряжении пробоя

$$\alpha_M(V_{brB}) \rightarrow \infty. \quad (IV.5.26)$$

Таким образом, в схеме ОБ напряжение лавинного пробоя равно напряжению пробоя коллекторного перехода: $V_{brB} = V_{brI}$. Вид выходных ВАХ при пробое представлен на рис. IV.5.10, а. Отметим, что в планарных транзисторах напряжение пробоя определяется свойствами коллектора, который легирован слабее, чем база.

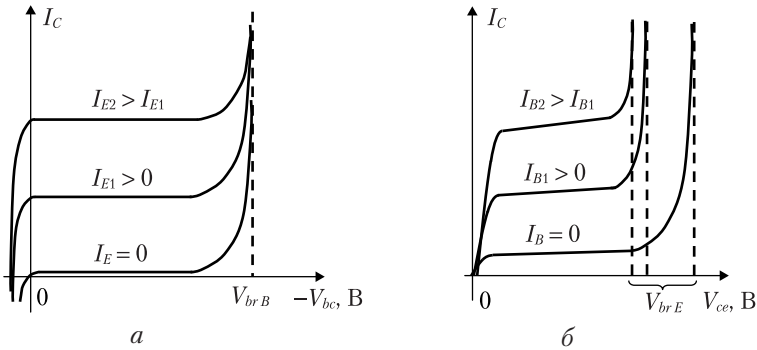


Рис. IV.5.10. Выходные ВАХ транзистора в режиме пробоя при включении ОБ (а) и ОЭ (б)

При включении транзистора по схеме ОЭ заданным электрическим параметром является ток базы. Подставляя в (IV.5.24б) $I_E = I_C + I_B$ и решая полученное уравнение относительно тока коллектора, получаем

$$I_C = \frac{\alpha_M(V_{bc})(\alpha_N I_B + I_{C0})}{1 - \alpha_M(V_{bc})\alpha_N}, \quad (IV.5.27)$$

где $V_{bc} = V_{be} - V_{ce} \approx -V_{ce}$. Соотношение (IV.5.27) показывает, что пробой происходит при условии

$$\alpha_M(V_{brE}) = 1/\alpha_N. \quad (IV.5.28)$$

Сравнение (IV.5.28) и (IV.5.26) показывает, что в схеме ОЭ напряжение пробоя существенно ниже, чем в схеме ОБ. Из (IV.5.25), (IV.5.26) и (IV.5.28) получаем

$$\frac{V_{brE}}{V_{brB}} = (1 - \alpha_N)^{1/\nu} \approx \beta_N^{-1/\nu}. \quad (IV.5.29)$$

Для коллекторного перехода кремниевого планарного n - p - n транзистора $v = 5$ (см. параграф II.6.2) Таким образом, при $\beta_N = 70$ напряжение пробоя в схеме ОЭ в 2,3 раза ниже, чем в схеме ОБ.

В заключение отметим, что коэффициент усиления тока базы β_N зависит от тока эмиттера (см. параграф IV.5.5) и, следовательно, от тока базы. Поэтому разным значениям тока базы в схеме ОЭ соответствуют различные напряжения пробоя, что показано на рис. IV.5.10, б.

IV.5.8. Диодное включение транзисторов

В интегральных схемах транзисторные структуры используются для реализации диодных элементов. Возможны шесть способов диодного включения транзистора, показанных на рис. IV.5.11. На рисунке схематично изображены структура планарного транзистора и распределения неосновных носителей заряда в квазинейтральных областях базы, эмиттера и коллектора для положительного напряжения анод—катод.

При одинаковых параметрах транзисторной структуры характеристики диодов различаются значениями тепловых токов, барьерной емкости, диффузионной емкости (при равных токах), сопротивления базы и площадью, занимаемой элементом на кристалле.

Сравнительные характеристики транзисторов в диодом включении сведены в табл. IV.5.1. Значения тепловых токов определялись из уравнений модели Эберса—Молла (IV.2.8), а диффузионная емкость оценивалась по суммарному заряду неосновных носителей в квазинейтральных областях. При этом принималось во внимание возможное накопление избыточного заряда электронов в пассивной базе (области II на рис. IV.3.2, пунктир на рис. IV.5.11). Площадь элемента на кристалле уменьшается при исключении из транзисторной структуры n^+ -эмиттера либо контактной n^+ -области коллектора.

Отметим, что минимальное значение диффузионной емкости диода достигается при объединении электродов базы и коллектора ($V_{bc} = 0$, рис. IV.5.11, е). В этом случае заряд накапливается только в активной области базы и в эмиттере. Поскольку эффективность эмиттера близка к 1, накоплением заряда в эмиттере можно пренебречь, и диффузионная емкость определяется временем пролета электронов через базу T_N , минимальным из всех времен пролета в транзисторе.

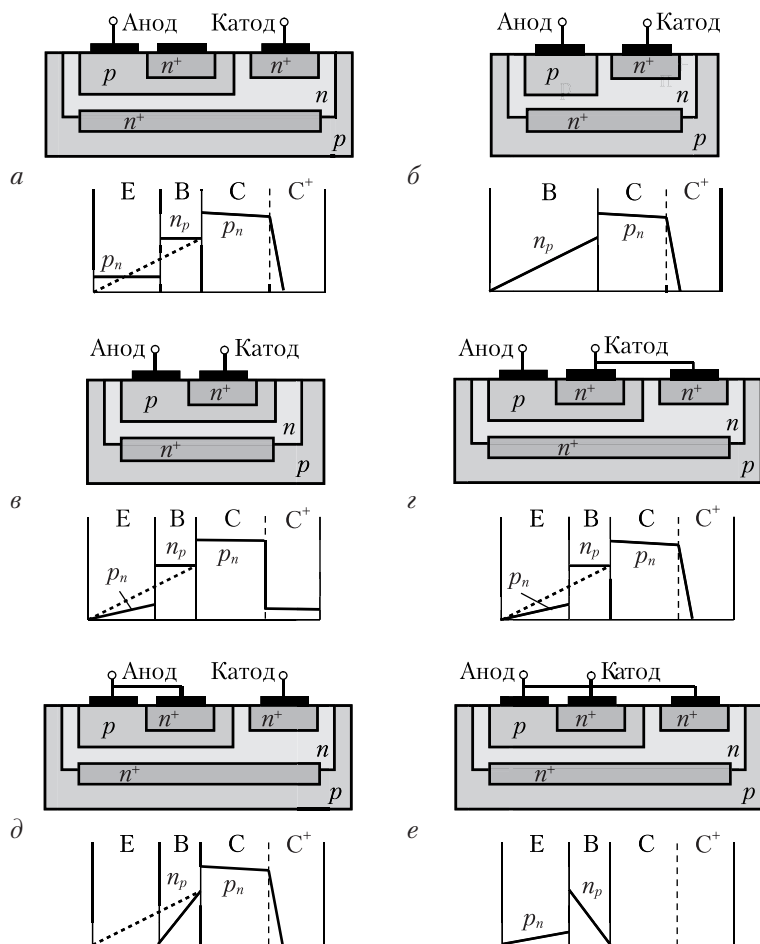


Рис. IV.5.11. Способы диодного включения планарного транзистора (структура и распределение неосновных носителей заряда при $V > 0$):

- a* — база-коллектор (эмиттер оборван);
- б* — база-коллектор (без эмиттера);
- в* — база-эмиттер (без коллекторного контакта);
- г* — база-эмиттер+коллектор;
- д* — база+эмиттер-коллектор;
- е* — база+коллектор-эмиттер.

Пунктир — распределение электронов в пассивной базе

Таблица IV.5.1

**Сравнительные характеристики
транзисторов в диодном включении**

Способ включения	Тепловой ток	Барьерная емкость	Диффузионная емкость	Сопротивление базы	Площадь
База — коллектор (эмиттер оборван)	$\dot{I}_{E0}$	C_c	Большая	$R_b + R_c$	Максимальная
База — коллектор (без эмиттера)	$< I_{2s}$	$< C_c$	Малая	$R_b + R_c$	Минимальная
База — эмиттер (без коллекторного контакта)	$< \dot{I}_{c0}$	C_E	Максимальная	R_b	Минимальная
База — эмиттер + коллектор	$I_{1s} + I_{2s}$	$C_E + C_c$	Большая	R_b	Максимальная
База+эмиттер — коллектор	I_{2s}	C_c	Большая	$R_b + R_c$	Максимальная
База+коллектор — эмиттер	I_{1s}	C_E	Минимальная	$R_b R_c$	Максимальная

Основные выводы

1. Напряжение на коллекторном переходе биполярного транзистора влияет на длину базы (эффект Эрли). Следствиями эффекта Эрли являются: а) конечное выходное сопротивление транзистора в нормальном режиме (сопротивление коллекторного перехода); б) внутренняя отрицательная обратная связь; в) прокол базы.

2. Сопротивление коллекторного перехода обратно пропорционально току эмиттера. В схеме ОЭ сопротивление коллекторного перехода в $1 + \beta_N$ раз меньше, чем в схеме ОБ. Степень проявления эффекта Эрли характеризуется напряжением Эрли.

3. Прокол базы состоит в уменьшении длины базы до нуля. В схеме ОЭ прокол базы является видом пробоя, а напряжение прокола близко к напряжению Эрли и равно произведению сопротивления коллекторного перехода на ток эмиттера.

4. В эквивалентной схеме Эберса—Молла эффект Эрли может быть учтен сопротивлением коллекторного перехода и генератором ЭДС обратной связи в цепи эмиттера.

5. Конечное сопротивление базы вызывает эффект оттеснения эмиттерного тока к контакту базы. Этот эффект проявляется, когда падение напряжения на сопротивлении активной базы сравнимо с температурным потенциалом.

6. Коэффициент усиления тока базы β_N зависит от тока эмиттера. При малом токе эмиттера коэффициент усиления снижается вслед-



ствие возрастания доли тока рекомбинации в эмиттерном переходе и снижения эффективности эмиттера. При высокой плотности тока эффективность эмиттера снижается вследствие изменения граничного условия на границе базы с эмиттерным переходом, что усугубляется эффектом оттеснения эмиттерного тока. Коэффициент переноса при высокой плотности тока снижается при оттеснении эмиттерного тока к периферии эмиттера вследствие увеличения среднего времени пролета электронов через базу и снижения их времени жизни.

7. Ограничение скорости носителей заряда в базе и в коллекторном переходе увеличивает время их пролета через базу и через коллекторный переход, что приводит к увеличению диффузионной емкости C_{DN} . Степень проявления этого эффекта возрастает с уменьшением длины базы. При высокой плотности коллекторного тока ограничение скорости носителей приводит к заполнению коллекторного перехода и n -коллектора электронами, инжектированными из эмиттера. Этот эффект (эффект Кирка) приводит к резкому увеличению диффузионной емкости C_{DN} и снижению быстродействия транзистора.

8. В планарных транзисторах напряжение пробоя определяется свойствами коллектора, который легирован слабее, чем база. В схеме ОЭ напряжение пробоя существенно ниже, чем в схеме ОБ.

9. В интегральных схемах транзисторные структуры используются для реализации диодных элементов. Минимальное значение диффузионной емкости диода достигается при объединении электродов базы и коллектора.

IV.6. Особенности дрейфовых планарных транзисторов

Особенностью современных транзисторов является неоднородное распределение легирующей примеси в электронейтральных областях, в первую очередь в базе. Неоднородность распределения результирующей примеси приводит к существованию «встроенного» (не связанного с внешними воздействиями) электрического поля. При этом механизм переноса неосновных носителей уже не может считаться чисто диффузионным. Ранее транзистор рассматривался в *бездрейфовом приближении*, которое соответствует усреднению концентраций результирующей примеси в пределах каждого квазинейтрального слоя. Особенности дрейфовых планарных транзисторов составляют предмет данной главы.



IV.6.1. Примесный профиль планарного дрейфового транзистора и встроенные электрические поля

Типичный примесный профиль планарного дрейфового транзистора изображен на рис. IV.6.1, *а*. Исходным материалом является подложка p -типа (p -S) с концентрацией акцепторов $N_s \approx 10^{15} \text{ см}^{-3}$. Скрытый n^+ -слой в коллекторе (n^+ -C) создается предельным легированием поверхности подложки донорами с низким коэффициентом диффузии (например, мышьяк). Затем на поверхности методом эпитаксии выращивается коллекторный слой n -типа (n -C) с концентрацией примеси $N_c \approx 10^{17} \text{ см}^{-3}$. Слой базы p -типа (p -B) формируется диффузией или ионной имплантацией акцепторной примеси (обычно бора) с поверхностной концентрацией $(5-10) \cdot 10^{18} \text{ см}^{-3}$. Теми же методами с использованием в качестве донорной примеси фосфора или мышьяка формируется эмиттерный n^+ -слой (n^+ -E) с поверхностной концентрацией доноров $N_E(0) > 10^{20} \text{ см}^{-3}$. Типичные значения глубины залегания металлургических переходов x_j составляют $x_{jE} \approx 0,5 \text{ мкм}$, $x_{jC} \approx 0,7 \text{ мкм}$, $x_{jS} \approx 3 \text{ мкм}$. Во время высокотемпературных процессов формирования базы и эмиттера скрытый n^+ -слой в коллекторе расплывается под действием диффузии до толщины 1–2 мкм. Штриховыми линиями на рис. IV.6.1, *а* показаны эффективные концентрационные профили в n^+ -слоях эмиттера N_{Eef} и коллектора N_{cef} , определяющие равновесные концентрации неосновных носителей — дырок (см. параграф I.3.1).

Практически всегда (как и на рис. IV.6.1, *а*) концентрация примеси в активной базе $N_B \ll N_v$ (N_v — эффективная плотность квантовых состояний в валентной зоне) и база может считаться невырожденной.

На рис. IV.6.1, *б* представлена эпюра напряженности *встроенного* электрического поля $E_0(x)$, возникающего вследствие неоднородного распределения примесей (координата x на рис. IV.6.1 смещена относительно координаты x'). Помимо полей в ОПЗ эмиттерного, коллекторного и изолирующего (коллектор-подложка) p - n переходов с равновесной шириной l_{e0} , l_{c0} и l_{s0} , соответственно, электрические поля формируются в квазинейтральных областях базы, эмиттера и n^+ -коллектора, легированных неоднородно. Равновесное поле может быть найдено из условия компенсации дрейфовой и диффузионной плотностей тока основных либо неосновных носителей

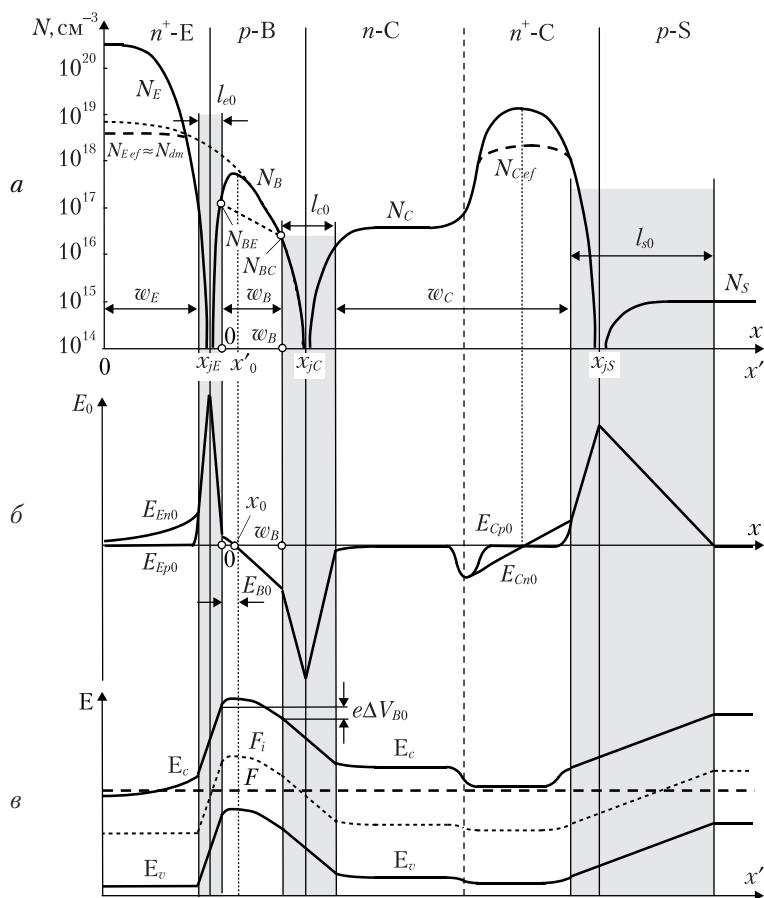


Рис. IV.6.1. Дрейфовый планарный транзистор в равновесном состоянии:

a — примесный профиль;

б — эпюра напряженности электрического поля;

в — энергетическая диаграмма

$$j_{p0} = ep_0\mu_p E_0 - eD_p dp_0 / dx = 0; \quad (\text{IV.6.1a})$$

$$j_{n0} = en_0\mu_n E_0 + eD_n dn_0 / dx = 0. \quad (\text{IV.6.1б})$$

Используя для *p*-базы условие (IV.6.1a), с учетом соотношения Эйнштейна

$$D_p = \Phi_T \mu_p \quad (\text{IV.6.2})$$

и условия квазинейтральности $p_{p0} \approx N_B$ получаем

$$E_{B0}(x) = \frac{\varphi_T}{p_{p0}(x)} \cdot \frac{dp_{p0}(x)}{dx}, \quad (\text{IV.6.3a})$$

или

$$E_{B0}(x) = \frac{\varphi_T}{N_B(x)} \cdot \frac{dN_B(x)}{dx}. \quad (\text{IV.6.3б})$$

Для невырожденной базы тот же результат с учетом закона действующих масс $p_0 n_0 = n_i^2$ следует из уравнения (IV.6.1б). Примесный профиль $N_B(x)$ обычно имеет максимум в точке $x' = x'_0$, $x = x_0$ (см. рис. IV.6.1, а), поэтому $E_{B0}(x_0) = 0$ (см. рис. IV.6.1, б). На участке $x > x_0$ поле $E_{B0} < 0$ направлено от коллектора к эмиттеру (см. рис. IV.6.1, б) и, следовательно, способствует переносу отрицательно заряженных электронов через базу в нормальном направлении — от эмиттера к коллектору (участок *ускоряющего поля*). На участке $x < x_0$ поле $E_{B0} > 0$ направлено в обратную сторону и препятствует переносу электронов в нормальном направлении (участок *тормозящего поля*).

Падение напряжения на базе от встроенного поля составляет $\Delta V_{B0} = - \int_0^{w_B} E_{B0}(x) dx$. С учетом (IV.6.3б) получаем

$$\Delta V_{B0} = \varphi_T \ln(N_{BE} / N_{BC}), \quad (\text{IV.6.4a})$$

или

$$\Delta V_{B0} = \eta \varphi_T, \quad (\text{IV.6.4б})$$

где $N_{BE} = N_B|_{x=0}$ и $N_{BC} = N_B|_{x=w_B}$ — концентрации примеси в базе на границах с эмиттерным и коллекторным переходами (см. рис. IV.6.1, а),

$$\eta = \ln(N_{BE} / N_{BC}) \quad (\text{IV.6.5})$$

— *фактор поля* в базе. Среднее значение поля в базе составляет

$$\bar{E}_{B0} = -\Delta V_{B0} / w_B = -\varphi_T \eta / w_B. \quad (\text{IV.6.6})$$

Как правило, $N_{BE} > N_{BC}$ (см. рис. IV.6.1, а). При этом $\eta > 0$, $\bar{E}_{B0} < 0$ и в среднем поле в базе является *ускоряющим*.

Для вырожденного n -эмиттера закон действующих масс имеет вид (I.3.13)

$$p_{n0} = n_i^2 / N_{Eef}, \quad (\text{IV.6.7})$$



причем вследствие предельно высокой степени легирования ($N_E \gg N_c$)

$$N_{Eef} \approx N_{dm} = N_c e^{-\Delta E_d / kT} \quad (\text{IV.6.8})$$

(соотношения (I.3.17), (I.3.18)) и

$$p_{n0} \approx n_i^2 / N_{dm}. \quad (\text{IV.6.9})$$

С учетом (IV.6.7) из (IV.6.1a) и (IV.6.1б) получаем

$$E_{Ep0}(x) = \frac{\Phi_T}{p_{n0}(x)} \cdot \frac{dp_{n0}(x)}{dx},$$

или
$$E_{Ep0}(x) = -\frac{\Phi_T}{N_{Eef}(x)} \cdot \frac{dN_{Eef}(x)}{dx}; \quad (\text{IV.6.10a})$$

$$E_{En0}(x) = -\frac{D_{nE}}{\mu_{nE} N_E(x)} \cdot \frac{dN_E(x)}{dx} \quad (\text{IV.6.10б})$$

(в (IV.6.10б) использование соотношения Эйнштейна (IV.6.2) для вырожденного электронного газа неправомерно). Условия равенства нулю дырочного (IV.6.1a) и электронного (IV.6.1б) токов дают для напряженности поля в эмиттере *различные* результаты. Это означает, что на дырки и на электроны действуют различные электрические поля: $E_{p0}(x)$ и $E_{n0}(x)$. Эти поля мы будем называть *квазиполями для дырок и электронов*.

Причина различия квазиполей для дырок и электронов состоит в деформации зонной структуры вырожденного полупроводника, что поясняется энергетической диаграммой на рис. IV.6.1, в. Энергия E_c соответствует потенциальной энергии электронов, поэтому напряженность квазиполя для электронов $E_{n0} = d(E_c / e) / dx$. Напряженность квазиполя для дырок составляет $E_{p0} = d(E_v / e) / dx$. Вследствие перекрытия примесной зоны с зоной проводимости дно объединенной зоны E_c расположено в примесной зоне, причем ширина запрещенной зоны $E_g = E_c - E_v$ уменьшается в направлении повышения степени легирования (на рис. IV.6.1, в — к поверхности кристалла). Таким образом, уровни E_c и E_v не параллельны, что соответствует различию полей E_{p0} и E_{n0} .

Концентрация дырок $p_{n0} = N_v e^{(E_v - F) / kT}$ определяется разностью между уровнем E_v и горизонтальным уровнем Ферми. Из (IV.6.9) следует, что она практически одинакова во всем эмиттере. Это означает, что *в эмиттере уровень E_v горизонтален и напряженность квазиполя для дырок равна нулю* (см. рис. IV.6.1, в).

В коллекторе напряженность квазиполя для дырок определяется соотношением (IV.6.10a) с заменой $N_{Eef} \rightarrow N_{Cef}$

$$E_{Cp0}(x) = -\frac{\Phi_T}{N_{Cef}(x)} \cdot \frac{dN_{Cef}(x)}{dx}; \quad (IV.6.11)$$

Она отлична от нуля только в переходной области между n - и n^+ -слоями, где коллектор невырожден и $N_{Cef}(x) \approx N_C(x)$ (см. рис. IV.6.1, в). В этой области квазиполе для дырок может быть весьма значительным, так как концентрация примеси сравнительно мала, а производная $dN_C(x)/dx$ велика. Это квазиполе является тормозящим для дырок, инжектированных из базы в коллектор. Оно препятствует проникновению дырок в n^+ -слой, о чем без детальных разъяснений говорилось в параграфе IV.4.7.

IV.6.2. Распределение избыточных носителей заряда в базе

В бездрейфовом приближении распределение неосновных носителей заряда в базе $n_p(x)$ находилось путем решения биполярного уравнения непрерывности в пренебрежении дрейфовым током (см. параграф IV.3.1). При учете дрейфового тока это уравнение решается в квадратурах только для случая, когда напряженность поля не зависит от координаты, что соответствует ее усреднению в соответствии с (IV.6.6). Более общий и более компактный результат может быть получен в пренебрежении рекомбинацией в базе. Погрешность определения функции $n_p(x)$ при этом незначительна, однако полученное решение не позволяет найти коэффициент переноса, который в пренебрежении рекомбинацией равен 1. Мы используем функцию $n_p(x)$ для определения теплового тока электронов и времени их пролета через базу, а коэффициент переноса определим приближенным, но достаточно точным соотношением (IV.3.6в).

В пренебрежении рекомбинацией в базе плотность электронного тока j_n не зависит от координаты x и функция $n_p(x)$ может быть найдена непосредственно из уравнения для тока

$$j_n = en_p \mu_B E_B + e D_B dn_p / dx, \quad (IV.6.12a)$$

которое является первым интегралом биполярного уравнения непрерывности (I.7.86) при $\tau_n = \infty$.



Вместо допущения о малости дрейфового тока электронов в теории идеализированного транзистора мы используем значительно более слабое допущение о малости плотности тока основных носителей, которое было обосновано в параграфе II.5.6:

$$j_p = e p_p \mu_B E_B - e D_B dp_p / dx = 0. \quad (\text{IV.6.12б})$$

Отсюда с учетом (IV.6.2) получаем

$$E_B(x) = \frac{\Phi_T}{p_p(x)} \cdot \frac{dp_p(x)}{dx}. \quad (\text{IV.6.13})$$

Подстановка (IV.6.13) в (IV.6.12а) позволяет представить дифференциальное уравнение (IV.6.12а) в виде уравнения в полных дифференциалах

$$-j_n dx = \frac{e D_B(x)}{p_p(x)} \cdot d[n_p(x) p_p(x)]. \quad (\text{IV.6.14})$$

Интегрируя уравнение (IV.6.14) от x до w_B (в системе координат рис. IV.6.1, б), получаем

$$n_p(x) p_p(x) - n_p(w_B) p_p(w_B) = -\frac{j_n}{e} \int_x^{w_B} \frac{p_p(x)}{D_B(x)} dx. \quad (\text{IV.6.15})$$

Произведение $n_p(w_B) p_p(w_B)$ определяется граничным условием (П.3.4а) на коллекторном переходе: $n_p(w_B) p_p(w_B) = n_i^2 e^{V_{bc}/\Phi_T}$. Для низкого уровня инжекции $p_p(x) = N_B(x)$ и

$$n_p(x) - \frac{n_i^2}{N_B(x)} e^{V_{bc}/\Phi_T} = -\frac{j_n}{e N_B(x)} \int_x^{w_B} \frac{N_B(x)}{D_B(x)} dx, \quad (\text{IV.6.16а})$$

где $n_i^2 / N_B(x) = n_{p0}(x)$. При допущении (IV.6.12б) ток коллектора имеет только электронную составляющую: $j_C = -j_n$ (знак «-» соответствует положительному направлению коллекторного тока против оси x на рис. IV.6.1). Таким образом, при $V_{bc} = 0$ из (IV.6.16а) получаем

$$\Delta n_p(x) = n_p(x) - n_{p0}(x) = \frac{j_C}{e N_B(x)} \int_x^{w_B} \frac{N_B(x)}{D_B(x)} dx. \quad (\text{IV.6.16б})$$

Соотношение (IV.6.16б) определяет в квадратуре распределение избыточных носителей заряда в базе для произвольного примесного профиля и учитывает, что коэффициент диффузии неосновных носителей в базе D_B зависит от концентрации примеси N_B и, следовательно, от координаты x .



Оно справедливо при $V_{bc} = 0$ и при низком уровне инжекции, когда *напряженность электрического поля мало отличается от равновесной*: $E_B(x) \approx E_{B0}(x)$.

В качестве *тестового примера* в дальнейшем будем рассматривать примесный профиль в виде простой функции

$$N_B(x) = N_{BE} e^{-\eta x / w_B}, \quad (\text{IV.6.17})$$

полагая $D_B = \text{const.}$ Из (IV.6. 3б) следует, что при этом напряженность поля

$$E_B = \Phi_T \eta / w_B \quad (\text{IV.6.18})$$

не зависит от координаты, а фактор поля η соответствует определению (IV.6.5), если $N_B(0) = N_{BE}$ и $N_B(w_B) = N_{BC}$. Таким образом, на рис. IV.6.1, а, выполненном в полулогарифмическом масштабе, функция $N_B(x)$ есть прямая, проходящая через точки N_{BE} и N_{BC} .

Подстановка (IV.6.17) в (IV.6.16б) дает

$$\Delta n_p(x) = \frac{j_c w_B}{e D_B \eta} [1 - e^{-\eta(1-x/w_B)}]. \quad (\text{IV.6.19})$$

Зависимости $\Delta n_p(x)$ для различных значений фактора поля при фиксированной плотности тока j_c представлены на рис. IV.6.2. Учитывая, что при постоянном поле плотность тока

$$-j_c = e n_p(x) \mu_B E_B + e D_B dn_p(x) / dx,$$

отметим следующие особенности функции $\Delta n_p(x) \approx n_p(x)$.

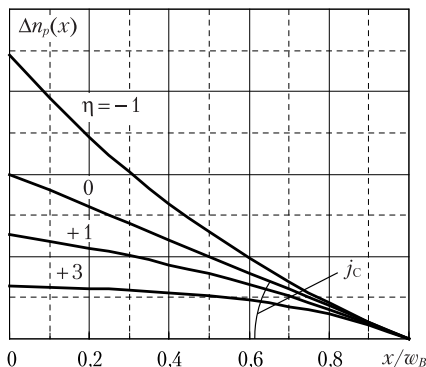


Рис. IV.6.2. Распределение избыточных носителей заряда в базе дрейфового транзистора для различных значений фактора поля η при фиксированной плотности тока j_c и $V_{bc} = 0$.

Примесный профиль $N_B(x) = N_{BE} e^{-\eta x / w_B}$

1. При $\eta = 0$ (бездрейфовый транзистор, $E_B = 0$) функция $n_p(x)$ линейна, что соответствует не зависящему от координаты градиенту концентрации электронов $dn_p(x)/dx$.

2. При $x = w_B$, где $n_p(x) \approx \Delta n_p(x) = 0$, ток является чисто диффузионным и определяется градиентом концентрации электронов $dn_p(x)/dx$, который для фиксированной плотности тока не зависит от фактора поля.

3. По мере приближения к эмиттерному переходу ($x = 0$) концентрация электронов возрастает. При $\eta > 0$ возрастает и доля дрейфового тока в полном токе, а доля диффузионного — снижается, что означает уменьшение $|dn_p(x)/dx|$. Таким образом, кривая $n_p(x)$ имеет выпуклый характер. При $\eta < 0$ кривая $n_p(x)$ вогнута (см. рис. IV.6.2).

4. При $\eta \gg 1$ ток является в основном дрейфовым. Везде, кроме области $x \approx w_B$, диффузионная составляющая мала, следовательно, $dn_p(x)/dx \approx 0$ (кривая $\eta = 3$ на рис. IV.6.2).

5. С увеличением фактора поля площадь под кривой $n_p(x)$ уменьшается. Это означает уменьшение заряда избыточных носителей в базе Q_N при фиксированном токе I_C и, следовательно, уменьшение времени пролета через базу

$$T_N = \frac{Q_N}{I_C} = \frac{e}{j_C} \int_0^{w_B} \Delta n_p(x) dx \quad (\text{IV.6.20})$$

под действием ускоряющего электроны поля. При $\eta < 0$ поле препятствует движению электронов к коллектору и время пролета возрастает (см. рис. IV.6.2).

IV.6.3. Время пролета неосновных носителей через базу и тепловые токи

В общем случае время пролета неосновных носителей через базу в нормальном направлении определяется соотношениями (IV.6.20) и (IV.6.16б)

$$T_N = \int_0^{w_B} \frac{dx}{N_B(x)} \int_x^{w_B} \frac{N_B(x)}{D_B(x)} dx. \quad (\text{IV.6.21a})$$

Аналогично для инверсного времени пролета через активную базу получаем

$$T_I = \int_0^{w_B} \frac{dx}{N_B(x)} \int_0^x \frac{N_B(x)}{D_B(x)} dx. \quad (\text{IV.6.21б})$$



Соотношения (IV.6.21) обобщают формулу (IV.3.7) $T_{N\text{dif}} = w_B^2 / 2D_B$, полученную в параграфе IV.3.1 в бездрейфовом приближении $\eta = 0$, когда неосновные носители переносятся через базу под действием диффузии. Для тестового примера (IV.6.17) при $D_B = \text{const}$ получаем

$$T_N = T_{N\text{dif}} \cdot 2(\eta - 1 + e^{-\eta}) / \eta^2. \quad (\text{IV.6.22})$$

При $\eta \geq 0$ хорошей аппроксимацией является функция

$$T_N / T_{N\text{dif}} \approx 2 / (2 + \eta). \quad (\text{IV.6.23})$$

Зависимости $T_N(\eta) / T_{N\text{dif}}$ (IV.6.22) и (IV.6.23) представлены на рис. IV.6.3. Рисунок показывает, что ускоряющее поле в базе снижает время пролета вдвое при $\eta \approx 2$. При $\eta < 1$ транзистор практически может считаться бездрейфовым; при $\eta > 2$ дрейфовый механизм переноса является основным. Соотношение (IV.6.5) показывает, что фактор поля слабо (логарифмически) зависит от отношения N_{BE} / N_{BC} . Практически в планарных транзисторах значение фактора поля редко превышает 2. Исключение составляют гетеропереходные биполярные транзисторы, которые будут рассмотрены в разделе VIII.3.

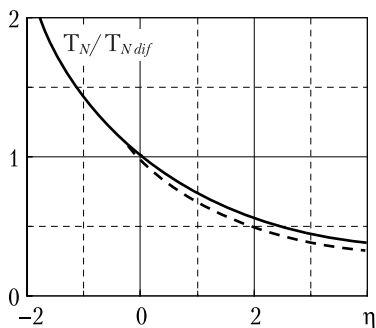


Рис. IV.6.3. Зависимость времени пролета неосновных носителей через базу от фактора поля.

Примесный профиль $N_B(x) = N_{BE} e^{-\eta x / w_B}$. Сплошная линия — формула (IV.6.22), штриховая — аппроксимация (IV.6.23)

Для расчета времени пролета носителей через активную базу в инверсном направлении в (IV.6.22) следует изменить знак фактора поля.

В пренебрежении рекомбинацией плотности электронного тока через эмиттерный и коллекторный переходы одина-

ковы: $j_{En} = j_c = -j_n$. Поэтому соотношение (IV.6.15) может быть использовано для определения ВАХ электронного тока I_{1n} эмиттерного диода D_1 в модели Эберса—Молла (см. рис. IV.2.2). При $x = 0$ из (IV.6.15) получаем

$$I_{En} = eS_E \frac{n_p(0)p_p(0) - n_p(w_B)p_p(w_B)}{\int_0^{w_B} \frac{N_B(x)}{D_B(x)} dx}. \quad (\text{IV.6.24})$$

Ток I_{En} равен электронному току эмиттера I_{1n} при нулевом напряжении база—коллектор V_{bc} (см. рис. IV.2.2). В (IV.6.24) произведения $n_p(0, w_B)p_p(0, w_B)$ определяются граничным условием (II.3.4а) на p - n переходах:

$$n_p(0)p_p(0) = n_i^2 e^{V_{be}/\Phi_T}; \quad n_p(w_B)p_p(w_B) = n_i^2 e^{V_{bc}/\Phi_T}.$$

Таким образом, при $V_{bc} = 0$ из (IV.6.24) следует

$$I_{1n} = \frac{eS_E n_i^2}{G_B} (e^{V_{be}/\Phi_T} - 1) = I_{1Sn} (e^{V_{be}/\Phi_T} - 1), \quad (\text{IV.6.25})$$

$$\text{где} \quad I_{1Sn} = eS_E n_i^2 / G_B \quad (\text{IV.6.26})$$

— электронный тепловой ток эмиттерного диода;

$$G_B = \int_0^{w_B} \frac{N_B(x)}{D_B(x)} dx \quad (\text{IV.6.27})$$

— число Гуммеля в базе.

При выводе соотношений (IV.6.25) — (IV.6.27) использованы граничные условия для произвольного уровня инжекции, однако сами эти соотношения справедливы только для *низкого уровня инжекции*, поскольку в (IV.6.15) и (IV.6.24) под интегралом вместо $p_p(x)$ использована функция $N_B(x)$.

Как показано в параграфе IV.6.1, в эмиттере квазиполе для дырок отсутствует, поэтому для дырочного тока эмиттерного диода D_1 справедливы результаты параграфа IV.3.2 для бездрейфового транзистора.

Полученные результаты позволяют сделать следующие выводы.

1. В дрейфовых транзисторах ВАХ эмиттерного диода и его тепловой ток определяются теми же соотношениями, что и в бездрейфовых (соотношения (IV.2.6а) и (IV.3.12а, б)).

2. Формула (IV.6.27) обобщает определение числа Гуммеля в базе (IV.3.14а) на случай произвольного примесного профиля.



IV.6.4. Коэффициент передачи эмиттерного тока

При известном времени пролета неосновных носителей через базу (IV.6.21а) коэффициент переноса определяется приближенным, но достаточно точным соотношением (IV.3.6в) и может быть представлен в виде

$$\kappa_N \approx 1 / (1 + T_N / \bar{\tau}_N) \approx 1 - T_N / \bar{\tau}_N. \quad (\text{IV.6.28})$$

Поскольку время жизни в дрейфовом транзисторе зависит от координаты x , его среднее значение с учетом (IV.6.16б) можно определить формулой

$$\begin{aligned} \bar{\tau}_N &= \int_0^{w_B} \Delta n_p(x) dx \bigg/ \int_0^{w_B} \frac{\Delta n_p(x)}{\tau_N(x)} dx = \\ &= \int_0^{w_B} \frac{dx}{N_B(x)} \int_x^{w_B} \frac{N_B(x)}{D_B(x)} dx \bigg/ \int_0^{w_B} \frac{dx}{\tau_N(x) N_B(x)} \int_x^{w_B} \frac{N_B(x)}{D_B(x)} dx, \end{aligned}$$

хотя на практике функция $\tau_N(x)$ с достаточной точностью неизвестна.

Эффективность эмиттера, как и в бездрейфовом приближении, определяется соотношением (II.4.43):

$$\gamma_N = \frac{1}{1 + G_B / G_E} \approx 1 - \frac{G_B}{G_E}.$$

Наконец, для определения инверсных параметров I_{2S} и $\alpha_I = \alpha_N I_{1S} / I_{2S}$ может быть использована методика параграфа IV.3.4 с учетом описанных особенностей неоднородной базы.

Частотные и переходные характеристики коэффициента передачи эмиттерного тока описываются соотношениями (IV.4.12а, б) и (IV.4.17а, б). Следует, однако, иметь в виду, что коэффициент ξ_α , определяющий фазовый сдвиг $\xi_\alpha \omega / \omega_\alpha$ в (IV.4.12б) или задержку коллекторного тока $\xi_\alpha \tau_\alpha$ в (IV.4.17б), в дрейфовом транзисторе должен быть больше, так как дрейф в электрическом поле характеризуется меньшим разбросом скоростей носителей, чем диффузия.

Оценить коэффициент ξ_α в дрейфовом транзисторе можно исходя из следующих соображений. При диффузионном механизме переноса электронов через базу (фактор поля $\eta = 0$) время задержки функции $\alpha(t)$ согласно (IV.4.13), (IV.3.7) и (IV.4.8) составляет

$$t_{dif} = \xi_{\alpha dif} \tau_\alpha \approx \xi_{\alpha dif} T_{N dif} = \xi_{\alpha dif} w_B^2 / 2D_B,$$



где $\xi_{\alpha dif} = (1 - 1/\sqrt{3})/2 \approx 0,21$. Это означает, что «наиболее быстрые электроны» за время $t_{d dif}$ диффундируют на расстояние

$$\Delta x_{dif} = \sqrt{2D_B t_{d dif} / \xi_{\alpha dif}} = w_B.$$

В дрейфовом транзисторе ($\eta \neq 0$) задержка фронта функции $\alpha(t)$ составляет $t_d = \xi_{\alpha} T_N$. За это время «наиболее быстрые электроны» проходят расстояние

$$\Delta x_{dif} = \sqrt{2D_B t_d / \xi_{\alpha dif}}$$

под действием диффузии, а также расстояние

$$\Delta x_{dr} = \mu_B \bar{E}_B t_d = \eta \varphi_T \mu_B t_d / w_B = \eta D_B t_d / w_B$$

под действием дрейфа в поле $\bar{E}_B = \varphi_T \eta / w_B$ (IV.6.6). Сумма $\Delta x_{dif} + \Delta x_{dr}$ должна составлять длину базы w_B

$$\sqrt{2D_B t_d / \xi_{\alpha dif}} + \eta D_B t_d / w_B = w_B.,$$

Отсюда с учетом приближенного соотношения $T_N / T_{N dif} \approx 2 / (2 + \eta)$ (IV.6.23) для $\eta \geq 0$ получаем

$$\xi_{\alpha} = \frac{t_d}{T_N} = \frac{2 + \eta}{\eta^2 \xi_{\alpha dif}} (1 + \eta \xi_{\alpha dif} - \sqrt{1 + 2\eta \xi_{\alpha dif}}). \quad (IV.6.29)$$

Зависимость (IV.6.29) относительной задержки переходной характеристики $\alpha(t)$ от фактора поля представлена на рис. IV.6.4.

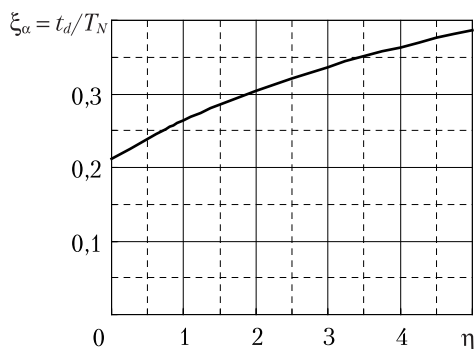


Рис. IV.6.4. Зависимость относительной задержки переходной характеристики $\alpha(t)$ от фактора поля

IV.6.5. Тиристорный эффект

Тиристорный эффект проявляется при взаимодействии трех близкорасположенных p - n переходов и состоит в том, что ВАХ четырехслойного полупроводникового прибора с чередующимся типом проводимости слоев имеет участок отрицательной дифференциальной проводимости.

На рис. IV.6.5, *a* показана четырехслойная тиристорная структура, на основе которой производится полупроводниковые приборы — *тиристоры*. Тиристоры могут иметь четыре электрода (катод C , две базы B_1 и B_2 и анод A), три электрода (один из электродов $B_{1,2}$ «оборван») или два электрода (оба электрода $B_{1,2}$ «оборваны»). При определенных условиях ВАХ тиристорной структуры (зависимость тока анода I_A от напряжения анод–катод V_{AC}) может иметь S-образный характер с участком отрицательной дифференциальной проводимости $G = \partial I_A / \partial V_{AC}$ (между точками 1 и 2 на рис. IV.6.6).

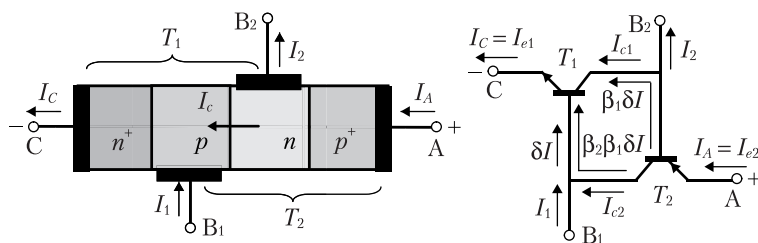


Рис. IV.6.5. Четырехслойная тиристорная структура (*a*) и ее двухтранзисторная модель (*b*)

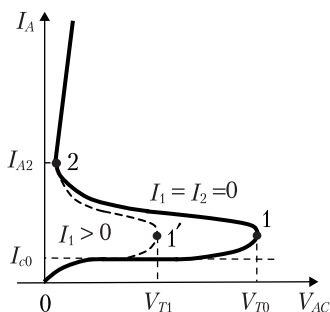


Рис. IV.6.6. ВАХ тиристорной структуры: при $I_1 = I_2 = 0$ (сплошная линия) и при $I_1 > 0, I_2 = 0$ (штриховая линия)

Тиристорный эффект удобно анализировать с помощью двухтранзисторной модели тиристорной структуры (рис. IV.6.5, б), включающей n - p - n транзистор T_1 и p - n - p транзистор T_2 . В реальной тиристорной структуре (рис. IV.6.5, а) база B_1 n - p - n транзистора T_1 совмещена с коллектором p - n - p транзистора T_2 и наоборот.

Отрицательная дифференциальная проводимость является следствием внутренней положительной обратной связи, которую в простейшем случае $I_1 = I_2 = 0$ можно проследить на двухтранзисторной модели (см. рис. IV.6.5, б). При $V_{AC} > 0$ оба транзистора работают в нормальном режиме (крайние эмиттерные переходы открыты, центральный коллекторный переход, общий для транзисторов T_1 и T_2 , закрыт). Поэтому флуктуация δI тока базы транзистора T_1 усиливается в коллекторной цепи T_1 в β_1 раз до значения $\beta_1 \delta I$, где β_1 — коэффициент усиления базового тока T_1 (см. рис. IV.6.5, б). При $I_2 = 0$ ток $\beta_1 \delta I$ является одновременно током базы транзистора T_2 и усиливается в коллекторной цепи T_2 в β_2 раз до значения $\beta_1 \beta_2 \delta I$. При $I_1 = 0$ ток $\beta_1 \beta_2 \delta I$ вновь поступает в цепь базы T_1 . Таким образом, коэффициент положительной обратной связи составляет $K_{FB} = \beta_1 \beta_2$.

Из общей теории обратной связи следует, что отрицательная дифференциальная проводимость возникает при условии

$$K_{FB} > 1, \text{ т.е. при } \beta_1 \beta_2 = \frac{\alpha_1}{1 - \alpha_1} \cdot \frac{\alpha_2}{1 - \alpha_2} > 1,$$

или

$$\alpha_1 + \alpha_2 > 1. \quad (\text{IV.6.30})$$

В общем случае произвольных токов $I_{1,2}$ значения токов анода и катода могут быть получены из уравнений Кирхгофа

$$I_A = I_{e2} = I_2 + I_{c1} + I_{c2}; \quad (\text{IV.6.31a})$$

$$I_C = I_{e1} = I_1 + I_{c1} + I_{c2}, \quad (\text{IV.6.31б})$$

а также уравнений для коллекторных токов транзисторов $T_{1,2}$ в нормальном режиме:

$$I_{c1} = \alpha_1 I_{e1} + I_{c01}; \quad (\text{IV.6.32a})$$

$$I_{c2} = \alpha_2 I_{e2} + I_{c02}, \quad (\text{IV.6.32б})$$

где $I_{c01, c02}$ — тепловые токи коллекторов T_1 и T_2 (см. параграф IV.2.4).

Учитывая, что транзисторы T_1 и T_2 имеют общий коллекторный переход ($I_c = I_{c1} + I_{c2}$, $I_{c0} = I_{c10} + I_{c20}$), из (IV.6.31а, б) и (IV.6.32а, б) получаем

$$I_A = I_2 + \alpha_1 I_A + \alpha_2 I_C + I_{c0}; \quad I_C = I_1 + \alpha_2 I_C + \alpha_1 I_A + I_{c0}.$$

Отсюда

$$I_A = \frac{\alpha_1 I_1 - (1 - \alpha_1) I_2 + I_{c0}}{1 - (\alpha_1 + \alpha_2)}; \quad (IV.6.33a)$$

$$I_C = \frac{\alpha_2 I_2 - (1 - \alpha_2) I_1 + I_{c0}}{1 - (\alpha_1 + \alpha_2)}. \quad (IV.6.33б)$$

Для частного случая $I_1 = I_2 = 0$

$$I_A = I_C = \frac{I_{c0}}{1 - (\alpha_1 + \alpha_2)}. \quad (IV.6.34)$$

Соотношения (IV.6.33а, б) и (IV.6.34) позволяют объяснить S-образный характер ВАХ тиристорной структуры на рис. IV.6.6. При нулевом напряжении анод—катод $I_A = I_C = 0$ в кремниевых структурах коэффициенты передачи эмиттерного тока $\alpha_{1,2} \ll 1$ вследствие определяющей роли токов рекомбинации—генерации в эмиттерных переходах (см. параграф IV.5.3). При увеличении напряжения анод—катод общий коллекторный переход транзисторов T_1 и T_2 расширяется и уменьшаются длины обеих баз (эффект Эрли). В результате увеличиваются коэффициенты передачи $\alpha_{1,2}$ и в соответствии с (IV.6.34) токи $I_A = I_C$. Увеличение эмиттерных токов транзисторов также повышает значения $\alpha_{1,2}$ (см. параграф IV.5.3), что приводит к ускоренному нарастанию токов $I_A = I_C$ и коэффициентов передачи $\alpha_{1,2}$. При некотором *пороговом напряжении* $V_{AC} = V_{T0}$ (точка 1 на рис. IV.6.6) выполняется условие $\alpha_1 + \alpha_2 = 1$ и значение дифференциальной проводимости становится бесконечным (см. рис. IV.6.6). При дальнейшем увеличении токов $I_A = I_C$ коэффициенты $\alpha_{1,2}$ возрастают и условие $\alpha_1 + \alpha_2 = 1$ выполняется при напряжении $V_{AC} < V_{T0}$, что и соответствует отрицательной дифференциальной проводимости (см. рис. IV.6.6).

При повышении токов $I_A = I_C$ до величины I_{A2} (точка 2 на рис. IV.6.6) напряжение V_{AC} снижается настолько, что общий коллекторный переход транзисторов отпирается — транзисторы T_1 и T_2 переходят в режим насыщения. При этом их усилительные свойства резко снижаются ($\alpha_1 + \alpha_2 < 1$) и дифференциальная проводимость вновь становится положительной.



Если задать положительное значение одного из токов базы (например, $I_1 > 0$), то эмиттерный ток транзистора T_1 увеличится на величину $I_1 > 0$. В результате возрастет значение коэффициента передачи α_1 , и условие $\alpha_1 + \alpha_2 = 1$ будет выполнено при меньшем значении порогового напряжения $V_{T1} < V_{T0}$ (штриховая ВАХ на рис. IV.6.6). Таким образом, одна или обе базы могут служить *управляющими электродами* тиристорной структуры.

Снижение порогового напряжения наблюдается также при быстром повышении напряжения V_{AC} — «эффект dV/dt ». В этом случае через барьерную емкость коллекторного перехода C_{bc} протекает дополнительный ток смещения $\Delta I_c = C_{bc} dV_{bc} / dt$. Этот ток увеличивает токи эмиттера транзисторов T_1 и T_2 : $\Delta I_{e1} = \Delta I_{e2} = \Delta I_c$. В цепях эмиттеров токи ΔI_{e1} и ΔI_{e2} уже не являются токами смещения — это токи инжекции носителей заряда через эмиттерные переходы. Поэтому увеличение токов эмиттера повышает коэффициенты передачи тока $\alpha_{1,2}$. В результате «эффект dV/dt » проявляется таким же образом, как дополнительный управляющий ток базы, и *приводит к снижению порогового напряжения*.

Тиристорный эффект может проявляться как паразитный эффект в четырехслойных транзисторных структурах (см. рис. IV.1.2) в случае отпираания изолирующего p - n перехода коллектор—подложка. Особенно опасен тиристорный эффект в ИС на комплементарных МДП-транзисторах, где паразитный тиристор образуется между областями истоков n -канальных и p -канальных МДПТ (рис. IV.6.7). Включение паразитного тиристора приводит к катастрофическим отказам, так как при этом замыкаются области истоков соседних комплементарных транзисторов — *эффект защелки*. Эффект защелки усугубляется «эффектом dV/dt » и представляет серьезную технологическую проблему.

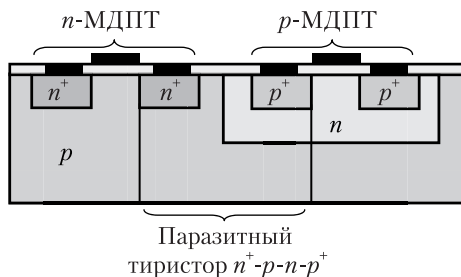


Рис. IV.6.7. Паразитный тиристор в структуре комплементарной пары МДПТ

Тиристоры используются как самостоятельные полупроводниковые приборы. Остаточное напряжение при включении тиристора достаточно мало. Этот фактор, а также возможность управления пороговым напряжением позволяет успешно использовать тиристоры в качестве мощных переключателей с высоким КПД в силовых системах (выпрямителях переменного тока, системах зажигания двигателей внутреннего сгорания).

Основные выводы

1. В базе дрейфового транзистора существует встроенное электрическое поле. Действие этого поля характеризуется фактором поля η , который определяется логарифмом отношения концентраций примеси в базе на границах с эмиттерным и коллекторным переходами.

2. В вырожденных областях эмиттера и коллектора напряженности квазиполей для дырок и электронов различны. В эмиттере напряженность квазиполя для дырок близка к нулю. В коллекторе квазиполе для дырок существует только в переходной области между n - и n^+ -слоями, препятствуя проникновению дырок в n^+ -слой.

3. Встроенное поле в базе сокращает время пролета неосновных носителей через базу в нормальном направлении.

4. ВАХ эмиттерного диода и его тепловой ток в дрейфовых транзисторах имеют тот же вид, что и в бездрейфовых.

5. Тепловые токи эмиттерного и коллекторного переходов, как и в бездрейфовых транзисторах, определяются числами Гуммеля. Для произвольного примесного профиля определение числа Гуммеля в базе обобщается соотношением (IV.6.27).

6. Частотные и переходные характеристики коэффициента передачи эмиттерного тока в дрейфовых транзисторах описываются теми же соотношениями, что и в бездрейфовых, однако коэффициент ξ_α , определяющий фазовый сдвиг или задержку коллекторного тока, возрастает с увеличением фактора поля.

7. В четырехслойных полупроводниковых структурах с чередующимся типом проводимости может проявляться тиристорный эффект, который состоит в S-образном характере ВАХ четырехслойной структуры с участком отрицательной дифференциальной проводимости. Тиристорный эффект проявляется как эффект защелки в ИС на комплементарных МДП-транзисторах и представляет серьезную технологическую проблему.



IV.7. Моделирование биполярных транзисторов

IV.7.1. Особенности моделирования биполярных транзисторов

Нелинейная эквивалентная схема биполярного транзистора для большого сигнала может быть синтезирована на основе модели Эберса—Молла путем добавления реактивных и резистивных элементов, не учтенных в идеализированной модели, а также элементов, учитывающих влияние подложки. Как и в модели МДПТ, в эквивалентной схеме биполярного транзистора выделяется часть, соответствующая «внутреннему» транзистору, с собственными внутренними электродами. В отличие от модели МДПТ, в расширенной модели Эберса—Молла нет проблемы описания зарядов, зависящих более чем от одного напряжения. Преимущества такой модели связаны с ее наглядностью. В качестве параметров модель содержит коэффициенты передачи тока, требования к которым неизбежно должны учитываться при проектировании электронных устройств. Эти же преимущества оборачиваются недостатками при расчетах, требующих высокой точности, так как феноменологический характер модели Эберса—Молла не позволяет адекватно учесть ряд существенных факторов.

Значительно большей мощностью обладает рассмотренная ниже модель Гуммеля—Пуна, основанная на математическом описании физических процессов в активных областях транзистора. Упрощенный вариант модели Гуммеля—Пуна соответствует расширенной модели Эберса—Молла. В модели Гуммеля—Пуна заряд в квазинейтральной области зависит от двух напряжений, однако изящный метод Гуммеля—Пуна решает задачу его математического описания. Модели высокого уровня описывают накопление заряда в коллекторной области методом Кулла (см. параграф IV.4.7).

Линейные малосигнальные эквивалентные схемы, как и для МДПТ, могут быть получены линеаризацией нелинейных элементов моделей для большого сигнала. Для оценочных расчетов небольших блоков электронных устройств иногда оказываются удобными специальные модели биполярного транзистора, включенного по схеме ОЭ.

Формальные линейные модели биполярного транзистора как линейного четырехполосника аналогичны соответствующим моделям МДПТ, однако используют другую систему параметров.



IV.7.2. Эквивалентные схемы для большого сигнала на основе модели Эберса—Молла

Эквивалентная схема транзистора, составленная на основе модели Эберса—Молла, изображена на рис. IV.7.1. Основная часть эквивалентной схемы включает диоды D_1 , D_2 , моделирующие эмиттерный и коллекторный переходы, и генераторы тока $\alpha_N I_1$, $\alpha_I I_2$, моделирующие транзисторный эффект. ВАХ диодов в идеализированной модели транзистора не включают токи рекомбинации—генерации в p - n переходах. Влияние эффекта Эрли модель не учитывает. Сопротивления R_B и R_C — это омические сопротивления тела базы и коллектора, которые считаются постоянными. Накопление заряда избыточных носителей в квазинейтральных областях базы и коллектора учитывается диффузионными емкостями C_{DN} и C_{DI} , зависящими от токов через соответствующие диоды (см. параграф IV.4.7). Емкости C_E , C_C и C_S — это барьерные емкости эмиттерного, коллекторного и изолирующего (подложка-коллектор) переходов, зависящие от напряжений на переходах.

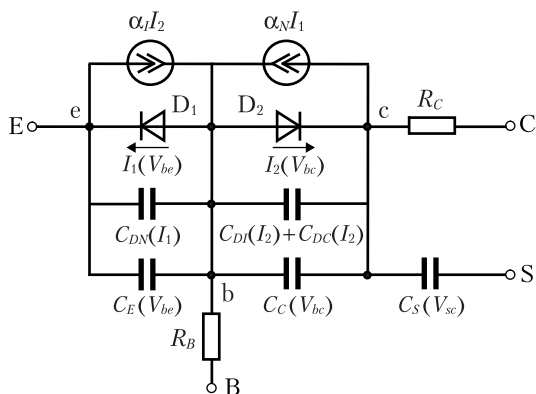


Рис. IV.7.1. Эквивалентная схема транзистора на основе модели Эберса—Молла

Модифицированный вариант эквивалентной схемы представлен на рис. IV.7.2. Главное отличие состоит в более корректном моделировании накопления заряда в n -коллекторе блоком 1, соответствующим модели Кулла (см. параграф IV.4.7, рис. IV.4.6). Кроме того, диод D_3 учитывает возможность отпираания изолирующего p - n перехода подложка-коллектор. При этом в p -подложку инжектируются электроны из n^+ -коллектора. Их заряд определяется вре-

менем жизни τ_s слаболегированной подложки $Q_s = \tau_s I_3$ и может быть весьма значительным, так как значение τ_s может составлять ~ 1 мкс.

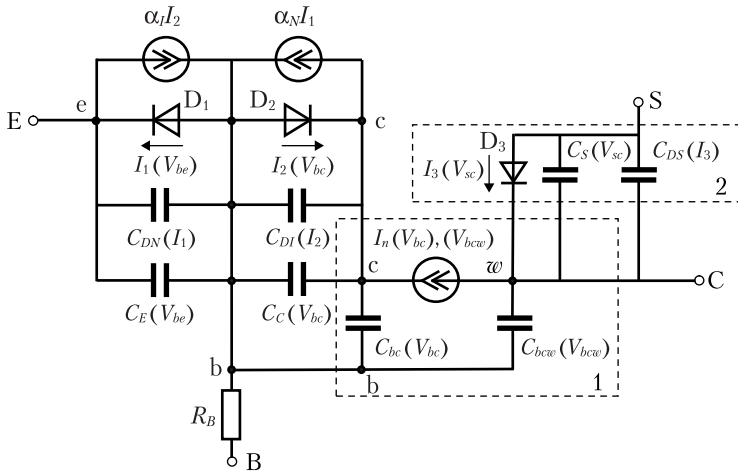


Рис. IV.7.2. Модифицированная эквивалентная схема транзистора на основе модели Эберса—Молла:

1 — модель n -коллектора Кулла;

2 — модель перехода коллектор-подложка

IV.7.3. Метод Гуммеля—Пуна

Метод Гуммеля—Пуна [9] состоит в представлении электронного тока I_n через активную базу в виде функции от напряжений на переходах и полного заряда Q_p основных носителей (дырок) в активной электронейтральной базе. На рис. IV.7.3 показаны распределения дырок в p -базе для равновесного (p_{p0}) и неравновесного (p_p) состояний. Последнее соответствует режиму насыщения ($V_{be} > 0$, $V_{bc} > 0$). Как и на рис. IV.6.1, a , координата x отсчитывается от поверхности кристалла, $x_{jE, jC}$ — координаты металлургических переходов эмиттер—база, коллектор—база, а $x_{10, 20}$ и $x_{1, 2}$ — равновесные и неравновесные границы нейтральной базы. Штриховая линия показывает распределение примеси в базе.

Задача решается при следующих допущениях:

1) рекомбинация в базе отсутствует и в стационарном состоянии электронный ток в базе не зависит от координаты x ;

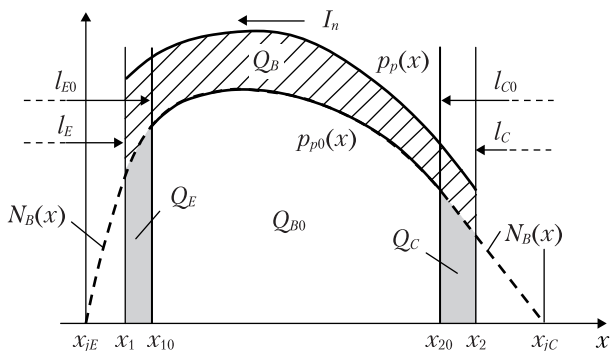


Рис. IV.7.3. Распределение дырок в активной базе для равновесного состояния и для режима насыщения ($V_{be} > 0$, $V_{bc} > 0$)

2) коэффициент диффузии электронов в базе D_B не зависит от координаты x ;

3) дырочный ток в базе мал: $I_n \gg I_p \approx 0$.

Отметим, что эти допущения являются значительно более слабыми, чем допущения идеализированной модели транзистора, и в реальных транзисторах выполнены с высокой точностью. В частности, сняты допущения о низком уровне инжекции, о постоянном значении длины базы w_B и об однородном легировании базы.

При допущениях 1 и 3 справедливо уравнение (IV.6.14)

$$j_n dx = -\frac{eD_B}{p_p(x)} \cdot d[n_p(x)p_p(x)], \quad (\text{IV.7.1})$$

где напряженность поля найдена из условия $j_p = 0$. Умножая это уравнение на eS_E^2 и интегрируя по всей базе (от x_1 до x_2 на рис. IV.7.3), получаем

$$I_n \left[eS_E \int_{x_1}^{x_2} p_p(x) dx \right] = (eS_E)^2 D_B [n_p(x_1)p_p(x_1) - n_p(x_2)p_p(x_2)]. \quad (\text{IV.7.2})$$

Выражение в квадратных скобках левой части (IV.7.2) есть *полный заряд дырок в базе*

$$Q_B = eS_E \int_{x_1}^{x_2} p_p(x) dx. \quad (\text{IV.7.3})$$

В правой части произведения $n_p(x_{1,2})p_p(x_{1,2})$ определяются граничными условиями (II.3.4а) на границах с ОПЗ

$$n_p(x_1)p_p(x_1) = n_i^2 e^{V_{be}/\Phi_T}; \quad n_p(x_2)p_p(x_2) = n_i^2 e^{V_{bc}/\Phi_T}$$

(справедливыми для *произвольного* уровня инжекции). Таким образом,

$$I_n = \frac{(en_i S_E)^2 D_B}{Q_B} (e^{V_{be}/\Phi_T} - e^{V_{bc}/\Phi_T}) \quad (\text{IV.7.4})$$

Уравнение (IV.7.4) называется *уравнением Гуммеля—Пуна*. Его можно записать в виде

$$I_n = I_{Sn} \frac{Q_{B0}}{Q_B} [(e^{V_{be}/\Phi_T} - 1) - (e^{V_{bc}/\Phi_T} - 1)],$$

$$\text{или} \quad I_n(V_{be}, V_{bc}) = I_{ne}(V_{be}, V_{bc}) - I_{nc}(V_{be}, V_{bc}), \quad (\text{IV.7.5})$$

$$\text{где} \quad I_{ne}(V_{be}, V_{bc}) = I_{Sn} \frac{Q_{B0}}{Q_B(V_{be}, V_{bc})} (e^{V_{be}/\Phi_T} - 1), \quad (\text{IV.7.6a})$$

$$I_{nc}(V_{be}, V_{bc}) = I_{Sn} \frac{Q_{B0}}{Q_B(V_{be}, V_{bc})} (e^{V_{bc}/\Phi_T} - 1), \quad (\text{IV.7.6б})$$

$$Q_{B0} = eS_E \int_{x_{10}}^{x_{20}} p_{p0}(x) dx = eS_E \int_{x_1}^{x_2} N_B(x) dx \quad (\text{IV.7.7})$$

— *равновесный* заряд дырок в базе (при $V_{be} = V_{bc} = 0$);

$$I_{Sn} = eS_E n_i^2 / G_{B0} \quad (\text{IV.7.8})$$

— электронный тепловой ток эмиттерного перехода;

$$G_{B0} = \frac{1}{D_B} \int_{x_{10}}^{x_{20}} N_B(x) dx \quad (\text{IV.7.9})$$

— *равновесное* число Гуммеля в базе.

Отметим, что соотношения (IV.7.8) и (IV.7.9) совпадают с (IV.6.26) и (IV.6.27), полученными в параграфе IV.6.3, однако относятся к *равновесному* состоянию. Потому величины I_{Sn} и G_{B0} , как и Q_{B0} , являются *параметрами* транзистора, т.е. числами, не зависящими от его электрического состояния. Заряды Q_B и Q_{B0} на рис. IV.7.3 соответствуют площадям под кривыми $p_p(x)$ и $p_{p0}(x)$.

Соотношения (IV.7.5) — (IV.7.6) показывают, что заряд Q_B и каждый из токов I_{ne} , I_{nc} являются функциями двух напряжений V_{be} , V_{bc} :

$$Q_B = Q_B(V_{be}, V_{bc}); \quad I_{ne} = I_{ne}(V_{be}, V_{bc}); \quad I_{nc} = I_{nc}(V_{be}, V_{bc}).$$



Для их определения кроме (IV.7.6а) и (IV.7.6б) необходимо еще одно уравнение. Оно будет получено из анализа структуры заряда Q_B , определенного в (IV.7.3). Как очевидно из рис. IV.7.3, заряд Q_B может быть представлен в виде суммы

$$Q_B = Q_{B0} + Q_E(V_{be}) + Q_C(V_{bc}) + \Delta Q_B(V_{be}, V_{bc}), \quad (\text{IV.7.10})$$

где значение равновесного заряда Q_{B0} определено в (IV.7.7),

$$Q_E(V_{be}) = eS_E \int_{x_1}^{x_{10}} N_B dx = \int_0^{V_{be}} C_E(V_{be}) dV_{be} \quad (\text{IV.7.11a})$$

$$\text{и} \quad Q_C(V_{bc}) = eS_E \int_{x_{20}}^{x_2} N_B dx = \int_0^{V_{bc}} C_C(V_{bc}) dV_{bc} \quad (\text{IV.7.11б})$$

— заряды в *барьерных емкостях* эмиттерного перехода C_E и коллекторного перехода C_C (затемненные площади на рис. IV.7.3);

$$\begin{aligned} \Delta Q_B(V_{be}, V_{bc}) &= eS_E \int_{x_1}^{x_2} [p_n(x) - p_{n0}(x)] dx = \\ &= Q_{DN}(V_{be}, V_{bc}) + Q_{DI}(V_{be}, V_{bc}) \end{aligned} \quad (\text{IV.7.12})$$

— заряд в *диффузионной емкости базы*, который согласно рисунку IV.4.6 является суммой нормальной и инверсной составляющих Q_{DBN} и Q_{DBI} , определенных в (IV.4.27):

$$Q_{DBN}(V_{be}, V_{bc}) = T_N I_{ne}(V_{be}, V_{bc}); \quad (\text{IV.7.13a})$$

$$Q_{DBI}(V_{be}, V_{bc}) = T_I I_{nc}(V_{be}, V_{bc}). \quad (\text{IV.7.13б})$$

В соотношение (IV.7.12) следует подставить (IV.7.13а, б) с учетом (IV.7.6а, б), а затем результаты (IV.7.12) и (IV.7.11а, б) собрать в (IV.7.10). В результате получаем квадратное уравнение относительно переменной Q_B , которое имеет решение

$$\begin{aligned} \frac{Q_B(V_{be}, V_{bc})}{Q_{B0}} &= \frac{1}{2} + \frac{Q_E + Q_C}{2Q_{B0}} + \\ &+ \sqrt{\left(\frac{1}{2} + \frac{Q_E + Q_C}{2Q_{B0}} \right) \frac{I_{Sn}}{I_{KN}} (e^{V_{be}/\Phi_T} - 1) + \frac{I_{Sn}}{I_{KI}} (e^{V_{bc}/\Phi_T} - 1)}, \end{aligned} \quad (\text{IV.7.14})$$

$$\text{где} \quad I_{KN} = Q_{B0} / T_N; \quad (\text{IV.7.15a})$$

$$I_{KI} = Q_{B0} / T_I. \quad (\text{IV.7.15б})$$



В итоге уравнения (IV.7.5), (IV.7.6а, б), (IV.7.14) и (7.15а, б) определяют *сквозной электронный ток* I_n через активную базу как функцию напряжений V_{be} и V_{bc} . Ток I_n назван здесь *сквозным*, так как при допущении 1) он проходит через базу без изменений.

Функция $I_n(V_{be}, V_{bc})$ содержит четыре параметра: Q_{B0} , I_{Sn} , T_N и T_I . Кроме этого, уравнения (IV.7.11а, б), определяющие функции $Q_E(V_{be})$ и $Q_C(V_{bc})$, содержат параметры ВФХ барьерных емкостей $C_E(V_{be})$ и $C_C(V_{bc})$, которые все равно должны быть заданы для их включения в эквивалентную схему.

Изящество метода Гуммеля—Пуна состоит в том, что зависимость $I_n(V_{be}, V_{bc})$ удалось определить без отыскания конкретного вида функции распределения концентрации носителей в базе $p_n(x)$. Здесь следует заметить, что, к сожалению, метод Гуммеля—Пуна не является математически безупречным. Строго говоря, значения времен пролета T_N и T_I *не являются параметрами*, т.е. числами, не зависящими от электрического состояния транзистора. Действительными параметрами могли бы служить значения T_{N0} и T_{I0} , соответствующие равновесному состоянию, которые и задаются при практическом моделировании.

Тем не менее метод Гуммеля—Пуна является значительно более мощным, чем феноменологический метод Эберса—Молла. Его преимущества состоят в следующем:

- 1) автоматический учет последствий эффекта Эрли;
- 2) автоматический учет последствий изменения уровня инжекции в базе;
- 3) возможность учета нелинейности сопротивления активной базы.

Эффект Эрли учитывается функциями $Q_E(V_{be})$ и $Q_C(V_{bc})$ в уравнении (IV.7.14) для полного заряда дырок в базе $Q_B(V_{be}, V_{bc})$. В моделях биполярного транзистора, которые используются в SPICE-подобных системах автоматизированного проектирования ИС, для ускорения вычислений и облегчения процедуры экспериментальной верификации параметров модели вместо уравнения (IV.7.14) используется приближенное соотношение

$$\frac{Q_B(V_{be}, V_{bc})}{Q_{B0}} \approx \frac{1/2 + \sqrt{1/4 + (I_{Sn}/I_{KN})(e^{V_{be}/\Phi_T} - 1) + (I_{Sn}/I_{KI})(e^{V_{bc}/\Phi_T} - 1)}}{1 - V_{bc}/V_{IRLN} - V_{be}/V_{IRLI}}, \quad (\text{IV.7.16a})$$



где V_{IRLN} и V_{IRLI} – нормальное и инверсное напряжения Эрли (см. параграф IV.5.2).

Пренебрежение эффектом Эрли соответствует нулевым зарядам Q_E и Q_C в (IV.7.14) или бесконечным значениям напряжений Эрли в (IV.7.16а):

$$\frac{Q_B(V_{be}, V_{bc})}{Q_{B0}} = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{I_{Sn}}{I_{KN}}(e^{V_{be}/\Phi_T} - 1) + \frac{I_{Sn}}{I_{KI}}(e^{V_{bc}/\Phi_T} - 1)}. \quad (IV.7.16б)$$

Механизм учета произвольного уровня инжекции заложен в параметрах I_{KN} и I_{KI} , определенных в (IV.7.15а, б). Чтобы продемонстрировать действие этого механизма, положим для простоты $V_{bc} = 0$ и пренебрежем эффектом Эрли. При этом в (IV.7.6б) $I_{nc} = 0$ и из (IV.7.6а), (IV.7.16б) получаем

$$I_{ne}(V_{be}) = I_{Sn} \frac{e^{V_{be}/\Phi_T} - 1}{1/2 + \sqrt{1/4 + (I_{Sn}/I_{KN})(e^{V_{be}/\Phi_T} - 1)}}.$$

При *низком уровне инжекции* $(I_{Sn}/I_{KN})(e^{V_{be}/\Phi_T} - 1) \ll 1/4$ и зависимость

$$I_{ne}(V_{be}) = I_{Sn}(e^{V_{be}/\Phi_T} - 1)$$

совпадает с классической ВАХ идеализированного диода.

При *высоком уровне инжекции* $(I_{Sn}/I_{KN})(e^{V_{be}/\Phi_T} - 1) \gg 1/4$, $e^{V_{be}/\Phi_T} \gg 1$ и ВАХ тока $I_{ne}(V_{be})$ имеет вид

$$I_{ne}(V_{be}) = \sqrt{I_{KN}I_{Sn}} \cdot e^{V_{be}/2\Phi_T},$$

характерный для высокого уровня инжекции.

Таким образом, параметры I_{KN} и I_{KI} характеризуют токи инжекции через эмиттерный и коллекторный переходы, при которых уровень инжекции близок к 1. Эти токи могут быть определены экспериментально. При $I_{KN, KI} \rightarrow \infty$ уравнения (IV.7.5), (IV.7.6а, б) и (IV.7.14) перестают учитывать уровень инжекции.

Дополнительным преимуществом метода Гуммеля—Пуна является возможность учета изменения сопротивления активной базы при повышении уровня инжекции. Сопротивление базы должно снижаться пропорционально повышению полного количества основных носителей (т.е. пропорционально отношению Q_{B0}/Q_B), что легко обеспечить в компьютерной модели транзистора.

Принципиальным недостатком метода Гуммеля—Пуна является невозможность учета эффекта оттеснения эмиттерного тока (см. параграф IV.5.4).



IV.7.4. Модель Гуммеля—Пуна

На основе уравнений (IV.7.5), (IV.7.6а, б), (IV.7.14) и (IV.7.15а, б) может быть синтезирована модель биполярного транзистора Гуммеля—Пуна [10] (рис. IV.7.4). Природа токов, фигурирующих в эквивалентной схеме, поясняется рис. IV.7.5.

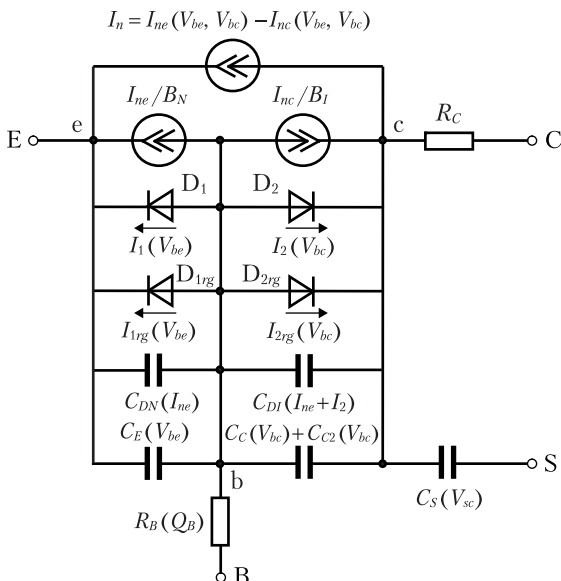


Рис. IV.7.4. Эквивалентная схема транзистора на основе модели Гуммеля—Пуна

Ядро модели составляет сквозной электронный ток $I_n(V_{be}, V_{bc})$, определенный соотношениями (IV.7.5), (IV.7.6а, б), (IV.7.14) и (IV.7.15а, б). На рис. IV.7.5 показаны его составляющие $I_{ne}(V_{be}, V_{bc})$ и $I_{nc}(V_{be}, V_{bc})$. Сквозной ток вычислен без учета рекомбинации в базе (допущение 1 в параграфе IV.7.3). В действительности электроны, движущиеся в нормальном (от эмиттере к коллектору) направлении и создающие ток I_{ne} , рекомбинируют с дырками в базе с вероятностью $1/B_N = T_N/\tau_N$, что приводит к протеканию в цепи база—эмиттер рекомбинационного тока I_{ne}/B_N (см. рис. IV.7.5). Такова же природа тока I_{nc}/B_I , где $1/B_I = T_I/\tau_I$. Эти составляющие тока базы учитывают отличие от 1 коэффициентов переноса $\kappa_N = 1/(1 + T_N/\tau_N)$ и $\kappa_I = 1/(1 + T_I/\tau_I)$. На эквивалентной схеме рис. IV.7.4 токи I_{ne}/B_N и I_{nc}/B_I учтены

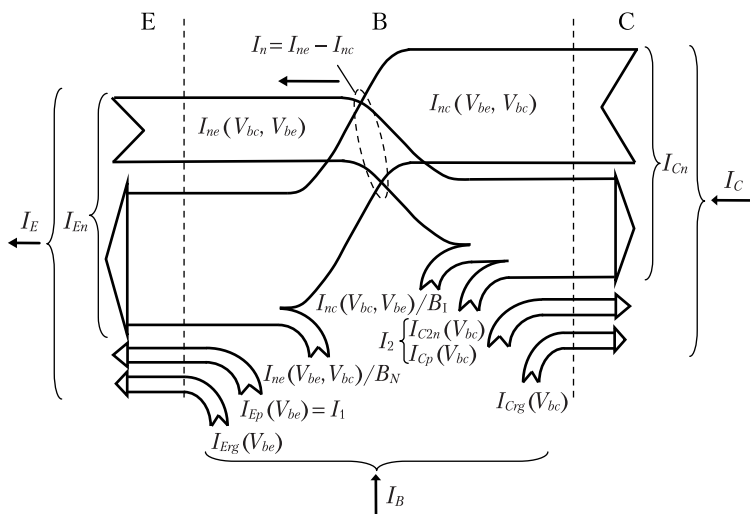


Рис. IV.7.5 Структура токов в транзисторе

специальными генераторами токов. Параметры B_N и B_I являются параметрами модели. Следует помнить, что, несмотря на сходство обозначений, *параметры B_N и B_I не есть нормальный и инверсный коэффициенты усиления тока базы*, так как в них не учтена эффективность эмиттера.

Ток $I_1(V_{be})$ диода D_1 на рис. IV.7.4 соответствует дырочной составляющей тока эмиттера $I_{Ep}(V_{be})$ на рис. IV.7.5, протекающей в цепи база—эмиттер. В цепи база—коллектор ситуация аналогична с той разницей, что в ток $I_2(V_{bc})$ диода D_2 на рис. IV.7.4 включена электронная часть тока коллектора $I_{C2n}(V_{bc})$, которая протекает через пассивную область коллекторного перехода (см. рис. IV.3.2) и не учтена в модели сквозного электронного тока I_n .

Диоды D_{1rg} и D_{2rg} на рис. IV.7.4 моделируют токи рекомбинации—генерации в эмиттерном и коллекторном переходах — $I_{Erg}(V_{be})$ и $I_{Crg}(V_{bc})$ на рис. IV.7.5, которые протекают в цепях база—эмиттер и база—коллектор, соответственно. Их ВАХ, отличны от ВАХ токов I_1 и I_2 .

Диффузионные емкости $C_{DN}(I_{ne})$ и $C_{DI}(I_{nc} + I_2)$ на рис. IV.7.4 моделируют накопление зарядов Q_{DN} и Q_{DI} в активной базе, определенных в (IV.7.13а, б), а также в пассивной базе и в n -коллекторе (см. параграф IV.4.7). При необходимости эквивалентная схема может быть дополнена моделью n -коллектора Кулла так же, как это сделано на рис. IV.7.2.

Как отмечалось выше, сопротивление базы R_B может быть представлено в виде суммы активной R_{B1} и пассивной R_{B2} составляющих. При этом легко может быть учтена зависимость сопротивления активной базы от уровня инжекции

$$R_B \approx k_R R_{B10} Q_{B0} / Q_B(V_{be}, V_{bc}) + R_{B2},$$

где коэффициент k_R учитывает различную подвижность избыточных дырок и электронов.

Недостатком модели Гуммеля—Пуна является ее сравнительная сложность и отсутствие наглядности. Практически она может быть использована только для компьютерных расчетов. В отличие от модели Эберса—Молла, модель не содержит в явном виде коэффициентов передачи тока; их значения могут быть определены только численно. Например, коэффициент усиления базового тока в нормальном режиме может быть рассчитан как $\beta_N = \partial I_C / \partial I_B$. При этом с достаточной точностью можно рассчитать зависимость $\beta_N(I_E)$, подобную приведенной на рис. IV.5.7. Как отмечено в параграфе IV.7.3, принципиальным недостатком метода Гуммеля—Пуна является неучет оттеснения эмиттерного тока. Однако даже этот эффект формально может быть отражен коррекцией токов I_{KN} и I_{KI} в (IV.7.15а, б).

В случае когда параметры модели Гуммеля—Пуна не установлены с достаточной точностью, для оценочных компьютерных расчетов ИС применяют упрощенный вариант модели, где по умолчанию полагается $I_{KN} = I_{KI} = \infty$, $V_{IRLN} = V_{IRLI} = \infty$, $I_{1rg} = I_{2rg} = 0$. При этом точность модели соответствует точности модели Эберса—Молла.

IV.7.5. Эквивалентные схемы для малого сигнала

Как и для МДПТ, для биполярного транзистора мало-сигнальные эквивалентные схемы обычно используются для моделирования его работы в нормальном режиме ($I_E \approx I_1 \gg I_{1S}$, $V_{bc} < -3\phi_T$) и могут быть получены из эквивалентной схемы для большого сигнала путем линеаризации ее элементов и исключения постоянных составляющих токов. Поскольку параметры линеаризованных элементов зависят только от режима работы (постоянные составляющие эмиттерного тока и напряжения база—коллектор), но не от амплитуды сигнала, использование в качестве исходной модели Гуммеля—Пуна не имеет смысла.



На рис. IV.7.6 представлена малосигнальная эквивалентная схема, полученная путем линеаризации элементов эквивалентной схемы Эберса—Молла. В качестве коэффициента передачи эмиттерного тока используется его дифференциальное значение $\alpha = \partial I_C / \partial I_E$. Дiode D_1 заменен дифференциальным сопротивлением эмиттерного перехода

$$r_e = \Phi_T / m I_E, \quad (\text{IV.7.17})$$

которое в малосигнальных схемах называют *сопротивлением эмиттера* (при низком уровне инжекции в базе $m = 1$, при высоком $m = 2$). Эффект Эрли учитывается дифференциальным сопротивлением закрытого коллекторного перехода r_c и генератором ЭДС обратной связи $\mu \dot{V}_{bc}$ (см. параграф IV.5.2).

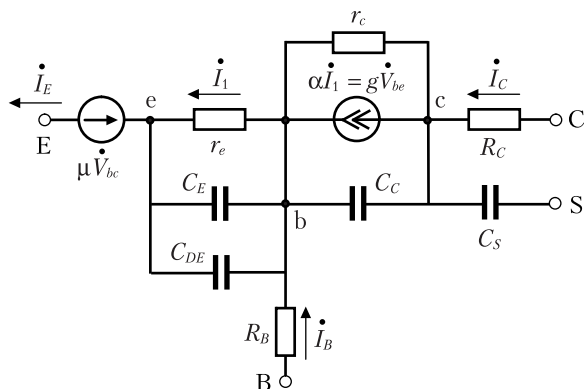


Рис. IV.7.6. Малосигнальная эквивалентная схема Джиаколетто

Управляемый генератор тока $\alpha_N I_1$ в эквивалентной схеме Эберса—Молла замещен линеаризованным генератором тока αI_1 . В компьютерных моделях, где для расчета токов обычно используется метод узловых потенциалов, току этого генератора присваивается значение $g \dot{V}_{be}$, где g — крутизна ВАХ биполярного транзистора. Поскольку $\dot{V}_{be} = I_e \Phi_T / I_E$, значения токов αI_1 и $g \dot{V}_{be}$ совпадают при условии

$$g = \alpha I_E / m \Phi_T. \quad (\text{IV.7.18})$$

Отметим, что при большом значении эмиттерного тока крутизна ВАХ биполярного транзистора может значительно превосходить крутизну ВАХ МДПТ.

Емкости C_E , C_C и C_S — это барьерные емкости эмиттерного, коллекторного и изолирующего p - n переходов. Частот-

ная зависимость коэффициента передачи эмиттерного тока учитывается диффузионной емкостью (IV.4.27а)

$$C_{DE} = \alpha_N T_N / r_e = \tau_\alpha / r_e,$$

что эквивалентно частотной характеристике (IV.4.12а)

$$\dot{\alpha}(\omega) \approx \frac{\alpha}{1 + i\omega\tau_\alpha}. \quad (\text{IV.7.19a})$$

Как отмечалось в параграфе II.6.4, в линейных устройствах часто предъявляются повышенные требования к точности моделирования АЧХ и особенно ФЧХ транзистора. Очевидный способ достижения этой цели состоит в использовании в управляемом генераторе тока $\alpha \dot{I}_1$ комплексного частотно-зависимого значения коэффициента передачи эмиттерного тока (IV.4.27б)

$$\dot{\alpha}(\omega) \approx \frac{\alpha e^{-i\xi_\alpha \omega \tau_\alpha}}{1 + i\omega\tau_\alpha (1 - \xi_\alpha)}. \quad (\text{IV.7.19б})$$

При этом *диффузионная емкость C_{DE} из эквивалентной схемы должна быть исключена*, иначе эффект накопления заряда в нейтральной базе будет учитываться дважды. Другой способ состоит в уменьшении диффузионной емкости C_{DE} до величины $(1 - \xi_\alpha)C_{DE}$ и использовании управляемого генератора *задержанного* тока $ge^{-i\xi_\alpha \omega \tau_\alpha} \dot{V}_{be}$. При компьютерном моделировании переходных процессов реализовать такой элемент значительно проще, чем осуществлять на каждом временном шаге обратное преобразование Фурье. В этом, кстати, состоит смысл замены комплексной частотной характеристики коэффициента переноса (IV.4.7б) аппроксимацией (IV.4.7в) в параграфе IV.4.2.

Уточненная малосигнальная схема биполярного транзистора, где реализован этот прием, приведена на рис. IV.7.7. Помимо введения генератора задержанного тока $g\dot{V}_{be}(t - \xi_\alpha \tau_\alpha)$ в этой схеме сопротивление базы разделено на активную и пассивную составляющие R_{B1} и R_{B2} , причем емкость C_C подключена к точке их соединения. Таким путем несколько повышается точность симуляции процесса ее заряда.

При моделировании линейных устройств коммуникационной и измерительной техники малосигнальные эквивалентные схемы дополняются генераторами шумового тока резистивных и активных элементов, как это делается при моделировании МДПТ.



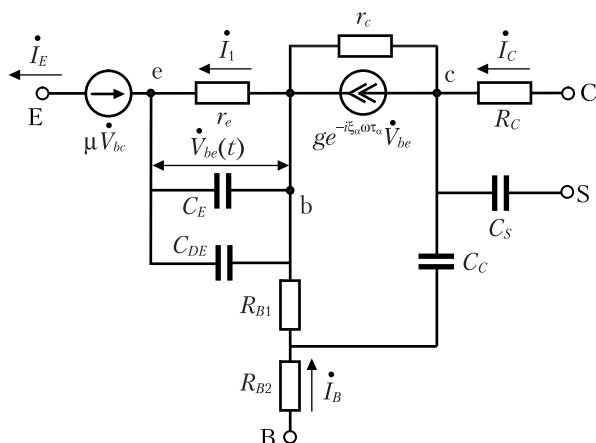


Рис. IV.7.7. Уточненная малосигнальная эквивалентная схема биполярного транзистора с управляемым генератором задержанного тока

Эквивалентные схемы на рис. IV.7.6, IV.7.7 пригодны для моделирования транзистора при произвольном способе его включения в схему электронного устройства. Однако для оценочных инженерных расчетов без применения вычислительных средств часто удобно использовать специальную эквивалентную схему для включения ОЭ. Такой анализ позволяет получить приближенные, но более общие результаты, которые отражают закономерности поведения электронных устройств и могут быть только проверены численными методами. При включении ОЭ ток эмиттера не интересует пользователя, поскольку не является ни входным, ни выходным током. Поэтому в эквивалентной схеме удобно иметь в качестве электрического аргумента ток базы либо его часть.

Упрощенная эквивалентная схема транзистора для включения ОЭ представлена на рис. IV.7.8. В управляемом генераторе тока $\beta \dot{I}'_B$ аргументом является часть тока базы $\dot{I}'_B$, не включающая ток смещения через барьерную емкость эмиттерного перехода C_E , который не связан с инжекцией электронов из эмиттера в базу. Диффузионная емкость базы отсутствует, так как в управляемом генераторе тока используется комплексное частотно-зависимое значение коэффициента усиления тока базы (IV.4.20а)

$$\dot{\beta}(\omega) \approx \frac{\beta}{1 + i\omega\tau_\beta}, \quad (\text{IV.7.20})$$

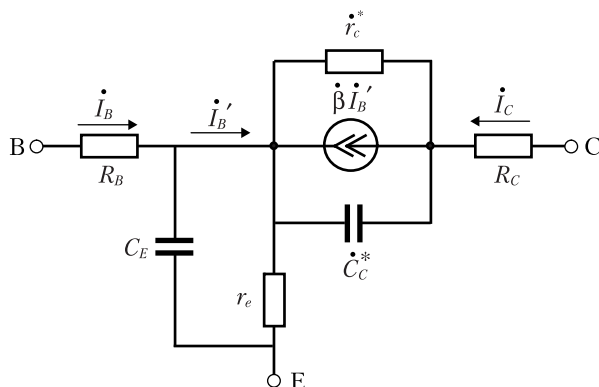


Рис. IV.7.8. Упрощенная малосигнальная эквивалентная схема биполярного транзистора для включения ОЭ

где согласно (IV.4.22) $\tau_\beta = (1 + \beta)\tau_\alpha$. Эффект Эрли учтен сопротивлением коллекторного перехода в схеме ОЭ (IV.5.5)

$$\dot{r}_c^*(\omega) = r_c / [1 + \dot{\beta}(\omega)]. \quad (\text{IV.7.21})$$

Отметим, что вследствие частотной зависимости коэффициента усиления базового тока сопротивление $\dot{r}_c^*$ также оказывается частотно-зависимой комплексной величиной. Емкость коллекторного перехода в схеме ОЭ $\dot{C}_c^*$ также является частотно-зависимой. Ее величина может быть найдена путем замены сопротивления r_c в (IV.7.21) на импеданс

$$\dot{Z}_c(\omega) \approx \frac{r_c}{1 + i\omega r_c C_c},$$

соответствующий параллельному соединению сопротивления r_c и емкости C_c (см. рис. IV.7.6):

$$\dot{Z}_c^*(\omega) = \frac{r_c}{(1 + \dot{\beta})(1 + i\omega r_c C_c)} = \frac{\dot{r}_c^*}{1 + i\omega \dot{r}_c^* \dot{C}_c^*}.$$

Импеданс $\dot{Z}_c^*(\omega)$ соответствует параллельному включению комплексного сопротивления $\dot{r}_c^*(\omega)$, определенного в (IV.7.21), и комплексной емкости

$$\dot{C}_c^*(\omega) = C_c [1 + \dot{\beta}(\omega)], \quad (\text{IV.7.22})$$

которая и включена в эквивалентную схему на рис. IV.7.8.

С помощью эквивалентной схемы на рис. IV.7.8 нетрудно показать, что *собственный коэффициент усиления биполярного транзистора по напряжению составляет*

$$K_V = \left| \frac{\dot{V}_{ce}}{V_{be}} \right|_{i_c=0} = \frac{\beta r_c^* - r_e}{R_B + r_e}.$$

Используя соотношение $r_c^* \approx V_{pinch}^* / I_E$ (IV.5.6б) и учитывая, что $r_e = \Phi_T / I_E \ll \beta r_c^*$, получаем

$$K_V \approx \frac{\beta V_{pinch}}{\Phi_T + I_E R_B}, \quad (IV.7.23)$$

где V_{pinch} — напряжение прокола базы.

IV.7.6. Формальные линейные модели

Формальный метод моделирования линейных четырех-полюсников описан в параграфе III.6.5. Для биполярного транзистора в области не слишком высоких частот наиболее удобна система h -параметров, так как в нормальном режиме транзистор имеет низкий входной импеданс (эмиттерный переход открыт) и высокий выходной импеданс (коллекторный переход закрыт). При этом во входную цепь удобно задавать ток (эмиттера или базы), а в выходную — напряжение (база-коллектор или коллектор-эмиттер).

В системе h -параметров уравнения связи аргументов и функций имеют вид

$$\begin{aligned} \dot{V}_1 &= \dot{h}_{11} \dot{I}_1 + \dot{h}_{12} \dot{V}_2; \\ \dot{I}_2 &= \dot{h}_{21} \dot{I}_1 + \dot{h}_{22} \dot{V}_2, \end{aligned}$$

а сами параметры имеют следующий смысл:

$$\begin{aligned} \dot{h}_{11} &= \dot{V}_1 / \dot{I}_1 \Big|_{\dot{V}_2=0} \quad \text{— входной импеданс при коротком замы-} \\ &\quad \text{кании (КЗ) на выходе;} \\ \dot{h}_{12} &= \dot{V}_1 / \dot{V}_2 \Big|_{\dot{I}_1=0} \quad \text{— обратный коэффициент передачи на-} \\ &\quad \text{пряжения при холостом ходе (ХХ) на} \\ &\quad \text{входе;} \\ \dot{h}_{21} &= \dot{I}_2 / \dot{I}_1 \Big|_{\dot{V}_2=0} \quad \text{— комплексный коэффициент передачи} \\ &\quad \text{тока при КЗ на выходе;} \\ \dot{h}_{22} &= \dot{I}_2 / \dot{V}_2 \Big|_{\dot{I}_1=0} \quad \text{— выходной адмиттанс при ХХ на входе;} \end{aligned}$$

В системе h -параметров измерения осуществляются при короткозамкнутом выходе и разомкнутом входе, что легко осуществить при малых значениях входного импеданса и выходного адмиттанса. Нетрудно видеть, что формальной



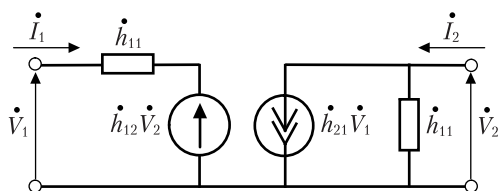


Рис. IV.7.9. Формальная эквивалентная схема линейного четырехполюсника в системе h -параметров

модели в системе h -параметров соответствует формальная эквивалентная схема, приведенная на рис. IV.7.9.

Для биполярного транзистора важнейшим является параметр h_{21} , имеющий смысл коэффициента передачи тока при КЗ на выходе. Для идеализированного транзистора на низкой частоте значения h_{21} совпадают с коэффициентом передачи эмиттерного тока для включения ОБ и с коэффициентом усиления тока базы для включения ОЭ. Из эквивалентных схем на рис. IV.7.6 и IV.7.8 нетрудно установить связь между h -параметрами и физическими параметрами транзистора. В частности, для параметра h_{21} в схемах включения ОБ и ОЭ

$$\dot{h}_{21 \text{ ОБ}} \approx \frac{\alpha}{1 + i\omega[\tau_\alpha + r_e C_E + (R_B + R_C)C_C]}; \quad (\text{IV.7.24a})$$

$$\dot{h}_{21 \text{ ОЭ}} \approx \frac{\beta}{1 + i\omega\tau_\beta + i\omega(1 + \beta)[r_e C_E + (r_e + R_C)C_C]}. \quad (\text{IV.7.24б})$$

Соотношения (IV.7.24) получены в пренебрежении элементами, моделирующими эффект Эрли, причем в знаменателях правой части отброшены члены, содержащие ω^2 и ω^3 . Учет сопротивлений r_c и r_c^* незначительно снижает низкочастотные значения коэффициентов передачи тока $h_{21 \text{ ОБ}}$ и $h_{21 \text{ ОЭ}}$.

Соотношение (IV.7.24б) позволяет оценить *предельную частоту транзистора* ω_T , на которой модуль параметра h_{21} в схеме ОЭ равен 1: $|\dot{h}_{21 \text{ ОЭ}}(\omega_T)| = 1$. Из (IV.7.24б) получаем

$$\begin{aligned} \omega_T &\approx 2\pi f_T \approx [\tau_\alpha + r_e C_E + (r_e + R_C)C_C]^{-1} = \\ &= [\alpha T_N + r_e C_E + (r_e + R_C)C_C]^{-1}. \end{aligned} \quad (\text{IV.7.25})$$

На предельной частоте транзистор теряет свои усиленные свойства. Соотношение (IV.7.24б) показывает, что

предельная частота транзистора определяется суммой времени пролета электронов через базу $\alpha_N T_N \approx T_N$, а также постоянных времени заряда емкости C_E через сопротивление r_e и емкости C_C через сопротивление $r_e + R_C$ (см. рис. IV.7.8). Согласно (IV.7.25) предельная частота не зависит от сопротивления базы R_B , так как по определению для ее измерения ток базы должен быть задан принудительно (от входного источника тока). В реальных схемах транзистор работает от источников сигнала с конечным выходным импедансом, поэтому сопротивление базы ограничивает быстродействие транзистора.

Следует отметить, что соотношение (IV.7.25) не относится к экстремальным режимам работы транзистора, например к режиму квазинасыщения, когда ограничение скорости носителей заряда в базе и в коллекторном переходе приводит к резкому снижению предельной частоты (см. параграф IV.5.6, эффект Кирка).

Основные выводы

1. Эквивалентная схема биполярного транзистора для большого сигнала может быть синтезирована на основе модели Эберса—Молла путем добавления барьерных и диффузионных емкостей, резистивных элементов, соответствующих сопротивлениям тела базы и коллектора, а также элементов, моделирующих изолирующий p - n переход подложка—коллектор. Эквивалентная схема может быть дополнена элементами, учитывающими накопление заряда в n -коллекторе в режиме квазинасыщения. Недостатком модели является ее феноменологический характер, не позволяющий учитывать ряд физических эффектов в транзисторе.

2. Наиболее точной является нелинейная модель Гуммеля—Пуна, основанная на представлении сквозного электронного тока через активную базу в виде функции от напряжений на переходах и полного заряда Q_p основных носителей в базе. Модель Гуммеля—Пуна обеспечивает автоматический учет последствий эффекта Эрли и изменения уровня инжекции в базе, а также возможность учета нелинейности сопротивления активной базы. Недостатками модели Гуммеля—Пуна являются ее сравнительная сложность и отсутствие наглядности, что ограничивает ее применение численными расчетами. Кроме того, эффект оттеснения эмиттерного тока может быть учтен только косвенными методами. Упрощенный вариант модели Гуммеля—Пуна соответствует точности модели Эберса—Молла.



3. Малосигнальные эквивалентные схемы для нормального режима работы транзистора могут быть получены из эквивалентной схемы для большого сигнала путем линеаризации ее элементов и исключения постоянных составляющих токов. Частотная зависимость коэффициента передачи эмиттерного тока учитывается диффузионной емкостью базы.

4. Повышение точности воспроизведения частотной характеристики коэффициента передачи эмиттерного тока достигается использованием управляемого генератора задержанного тока.

5. Для оценочных инженерных расчетов без применения вычислительных средств целесообразно использовать специальную эквивалентную схему для включения ОЭ. В этой схеме барьерная емкость и сопротивление коллекторного перехода являются комплексными функциями частоты.

6. Для формального моделирования биполярного транзистора как линейного четырехполюсника наиболее удобна система h -параметров. Частотные свойства транзистора характеризуются предельной частотой, на которой модуль параметра h_{21} в схеме ОЭ равен 1.

IV.8. Структуры интегральных биполярных транзисторов

Типовые структуры интегральных биполярных n - p - n транзисторов схематично показаны на рис. IV.8.1.

В структуре на рис. IV.8.1, *а* боковая изоляция осуществляется p - n переходами между коллекторной областью и глубокой диффузионной p^+ -областью (областью разделительной диффузии). Описание структуры было приведено ранее в параграфе IV.6.1. Структура транзистора с разделительной диффузией технологически наиболее проста, но не обеспечивает высокой плотности компоновки элементов на кристалле ИС.

Более компактна структура с боковой диэлектрической изоляцией (рис. IV.8.1, *б*), где для взаимной изоляции используются глубокие «щели», заполненные диэлектриком (SiO_2).

В базе транзисторов с боковой диэлектрической изоляцией скорость рекомбинации на границе с SiO_2 повышена, однако это не сказывается существенно на коэффициенте усиления, поскольку его значение в основном определяется эффективностью эмиттера (см. параграф IV.3.3).

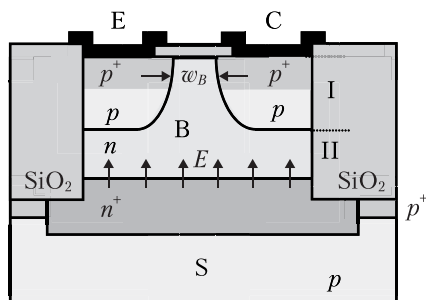


Рис. IV.8.2. Латеральный p - n - p транзистор, выполненный по технологии с боковой диэлектрической изоляцией.

n^+ -контакт к базе находится за плоскостью чертежа.

I — активная область; II — пассивная область

и на рис. IV.8.2 находится за плоскостью чертежа. Активная область I находится в пределах толщины p -слоя.

Латеральные p - n - p транзисторы значительно уступают по качеству n - p - n транзисторам вследствие нескольких причин.

1. Длина n -базы обычно в несколько раз превышает длину p -базы n - p - n транзисторов, так как расстояние между p -областями эмиттера и коллектора зависит не только от боковой диффузии примесей при формировании p -слоев, но и от топологических размеров на фотошаблонах.

2. Расстояние между эмиттером и коллектором минимально вблизи поверхности n -слоя, где скорость рекомбинации максимальна.

3. Заметная часть дырочного тока инжектируется из p -эмиттера в пассивную часть базы (под эмиттером). Обычно коэффициент усиления базового тока в латеральных p - n - p транзисторах не превышает 10–15. Поддержанию усилительных свойств на приемлемом уровне способствует встроенное электрическое поле в переходной области между n - и n^+ -слоями (стрелки на рис. IV.8.2), которое препятствует потоку дырок в сторону скрытого n^+ -слоя (в область пассивной базы II на рис. IV.8.2). Вследствие большей длины базы латеральные p - n - p транзисторы значительно уступают обычным n - p - n также и по частотным свойствам. В некоторых случаях их полезным свойством является симметрия — нормальный и инверсный коэффициенты передачи токов практически одинаковы.

Технология изготовления биполярных ИС непрерывно совершенствуется. Основными задачами являются

повышение степени интеграции и быстродействия ИС. Эти задачи взаимосвязаны, так как сводятся к созданию транзисторной структуры, занимающей на кристалле минимальную площадь при фиксированном значении минимального топологического размера. При этом минимизируются площади (и барьерные емкости) эмиттерного, коллекторного и изолирующего p - n переходов, что обеспечивает повышение быстродействия. Кроме того, должны быть приняты меры к снижению сопротивлений тела базы и коллектора.

В первых структурах задачи решались с помощью совершенствования методов взаимной изоляции транзисторов p - n переходом. Дальнейший прогресс был достигнут за счет применения изоляции термически выращенным оксидом в сочетании с совершенствованием методов литографии. На этой основе разработаны технологии LOCOS [11], ISOPLANAR [12], ISOPLANAR-S [12], ISAC [13] и др., использованные для создания БИС.

Для оптимизации транзисторных структур применяются следующие основные конструктивно-технологические приемы:

- 1) использование методов самосовмещения и самоформирования;
- 2) использование поликремния для уменьшения толщины эмиттера и дополнительного легирования пассивной базы;
- 3) сближение омических контактов к базе и эмиттеру;
- 4) использование щелевой диэлектрической изоляции;
- 5) создание «пристеночных» p - n переходов для уменьшения торцевых составляющих емкостей переходов.

Применение щелевой изоляции позволяет сократить размеры транзисторной структуры при заданном минимальном размере литографии. Примером является технология VIST [14] самосовмещенной транзисторной структуры с вертикальной изоляцией (рис. IV.8.3). Ее недостатком является невозможность уменьшения зазора между контактами до минимального топологического размера.

Этот недостаток устранен в серии суперсовмещенных технологических процессов SST [15–17], которые различаются способами формирования омических контактов. Структура транзистора, выполненного по технологии SST-3, в упрощенном виде представлена на рис. IV.8.4. В структурах SST расстояние между контактами эмиттера и базы уменьшено до толщины слоя диэлектрика.

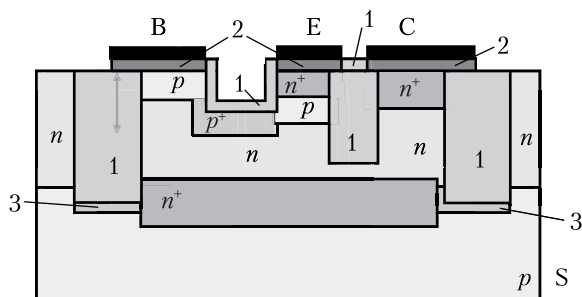


Рис. IV.8.3. Самосовмещенная транзисторная структура с вертикальной изоляцией, выполненная по технологии VIST:
1 — SiO_2 ; 2 — поликремний; 3 — p^+ -противоканальный слой

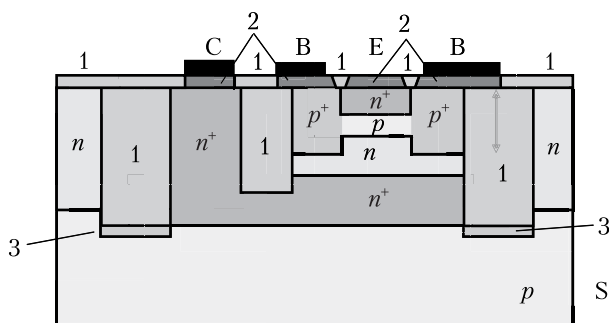


Рис. IV.8.4. Самосовмещенная транзисторная структура с вертикальной изоляцией, выполненная по технологии SST:
1 — SiO_2 ; 2 — поликремний; 3 — p^+ -противоканальный слой

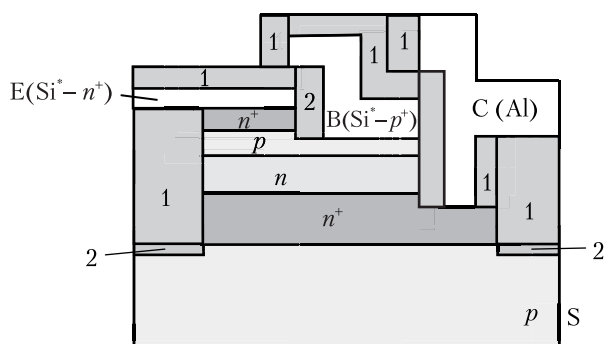


Рис. IV.8.5. Самосовмещенная транзисторная структура, выполненная по технологии СТЕПЛАНАР:
1 — SiO_2 ; 2 — p^+ -противоканальный слой

Рекордно малая площадь транзистора на кристалле достигнута с применением технологии СТЕПЛАНАР за счет разнесения контактов к различным областям транзисторной структуры на разные уровни [18] (рис. IV.8.5).

Особую группу составляют биполярные транзисторы, предназначенные для работы в коммуникационных устройствах СВЧ-диапазона (выше 1 ГГц). Транзисторы СВЧ имеют две основные области применения: 1) прием и усиление гармонических сигналов (несущей частоты); 2) генерация мощных гармонических колебаний в передающих системах.

Независимо от назначения транзисторы СВЧ должны иметь высокую предельную частоту и хорошие шумовые характеристики. В приемных системах требование к малому собственному шуму очевидно; в системах передачи собственный шум нарушает чистоту спектра генерируемых сигналов. Для снижения собственного шума транзисторы должны иметь малые сопротивления тела базы R_B и тела коллектора R_C .

Мощные транзисторы передающих устройств работают при высокой плотности тока, поэтому при их конструировании должны быть приняты меры к нейтрализации эффекта оттеснения эмиттерного тока, который снижает усилительные и частотные свойства транзисторов (см. параграф IV.5.4). Главным образом эти меры должны обеспечивать снижение сопротивления активной базы.

В приемных устройствах СВЧ сигналы передаются по волноводам (коаксиальным кабелям, полосковым линиям или полым металлическим трубам), которые имеют стандартное волновое сопротивление 50 Ом. Входные и выходные импедансы усилительных устройств должны быть согласованы с волновыми сопротивлениями трактов передачи сигналов во избежание отражений сигналов и паразитных резонансов. Это приводит к необходимости работы усилительных транзисторов в приемных устройствах также в режиме больших токов. Действительно, входное сопротивление биполярного транзистора в схеме ОЭ составляет $R_{in} = R_B + (1 + \beta)r_E = R_B + (1 + \beta)\varphi_T / I_E$ и его необходимая величина может быть достигнута только при малом значении сопротивления базы и большом токе.

Основные требования к параметрам структуры и конструкции транзисторов СВЧ приведены в табл. IV.8.1.

Таблица IV.8.1

**Основные требования к параметрам структуры
и конструкции транзисторов СВЧ**

№	Требования к параметрам и конструкции	Цель
1	Малая толщина базы w_B	Повышение предельной частоты
2	Малые барьерные емкости эмиттерного и коллекторного переходов C_E, C_C	
3	Низкое сопротивление базы R_B	Повышение предельной частоты. Снижение собственного шума. Нейтрализация эффекта оттеснения эмиттерного тока
4	Низкое сопротивление тела коллектора R_C	
5	Непланарная конструкция	Снижение температуры перегрева. Повышение предельной частоты (снижение R_C)

Как очевидно из таблицы, часть поставленных целей обеспечивается отказом от планарной конструкции, что во многих случаях вынуждает выполнять устройства СВЧ в виде многокристалльных сборок.

Поставленные цели частично достигаются в полосковой конструкции транзистора, показанной на рис. IV.8.6, *a*.

При заданном значении рабочего тока снижение плотности эмиттерного тока достигается увеличением площади эмиттерного перехода $S_E = b_E w_E$ (см. рис. IV.8.6, *a*). Для нейтрализации эффекта оттеснения эмиттерного тока следует снижать сопротивление активной базы $R_{B1} \sim b_E / w_E$. Таким образом, ширина эмиттера b_E в направлении протекания тока базы должна быть минимально возможной. При этом снижение плотности эмиттерного тока до допустимой величины достигается увеличением длины полоски эмиттера w_E . Отметим, что при выбранном значении w_E ширина эмиттера b_E может быть и увеличена, однако это приведет к увеличению барьерных емкостей эмиттерного и коллекторного переходов.

При большой мощности размер w_E может достигать ≈ 1 мм. При этом становится существенным сопротивление тела эмиттера R_E и возрастает размер кристалла. Избежать этого можно использованием встречно-штырьевой или многоэмиттерной конструкции (рис. IV.8.6, *б, в*).



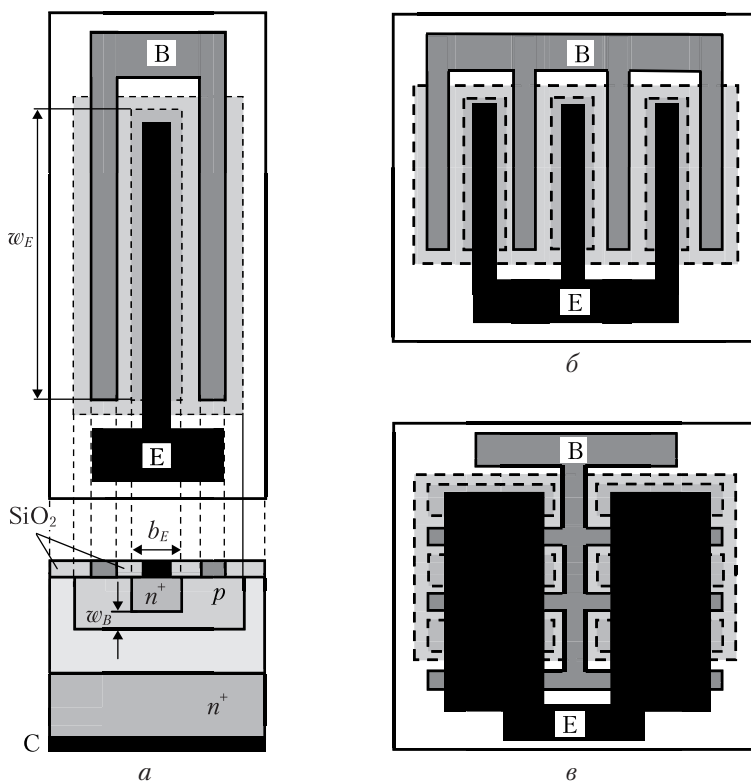


Рис. IV.8.6. Конструкции мощных транзисторов и транзисторов СВЧ:

a — полосковая; *б* — встречно-штырьевая; *в* — многоэмиттерная

В транзисторах, предназначенных для работы на частоте более 10 ГГц, длина базы w_B составляет менее 0,1 мкм, а ширина эмиттера b_E менее 0,3 мкм.

Описанные структуры могут быть реализованы и в планарном варианте (рис. IV.8.7). Недостатком планарного СВЧ-транзистора является повышенное сопротивление тела коллектора R_c .

Конструктивно транзисторы СВЧ оформляются в специальные корпуса, предназначенные для монтажа на полосковые линии. Для этого они снабжаются «балочными» выводами толщиной ~ 10 мкм, которые формируются осаждением металла (обычно золота) и служат держателями кристалла в корпусе или в микросборке.

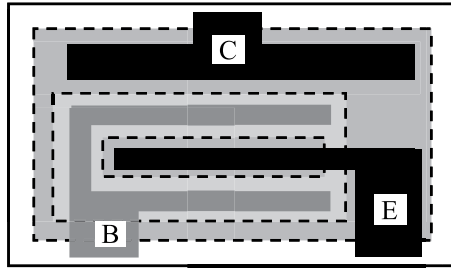


Рис. IV.8.7. Планарная полосковая конструкция транзистора СВЧ

Описанные приемы применяются и при конструировании биполярных транзисторов, предназначенных для работы в мощных силовых устройствах. Такие транзисторы обычно работают на низкой частоте, поэтому требования минимальных размеров b_E и w_B для них необязательны.

Литература

1. *Bardeen, J.* A semiconductor triod / J. Bardeen, W. H. Brattein // Phys. Rev. — 1948. — Vol. 74. — P. 230.
2. *Shockley, W.* The theory of p - n junctions in semiconductors and p - n junction transistors / W. Shockley // Bell Syst. Techn. J. 1949. — Vol. 28. — P. 435.
3. *Ebers, J. J.* Large-signal behavior of junction transistors / J. J. Ebers, J. L. Moll // Proc. IRE. — 1954. — Vol. 42. — P. 1761–1772.
4. *Kull, G. M.* Unified circuit model for bipolar transistor including quasi-saturation effects / G. M. Kull [and other] // IEEE Trans. — 1985. — Vol. — ED-32. — N_o 6. — P. 1103–1113.
5. *Ройзин, Н. М.* Накопление избыточных носителей заряда в дрейфовом транзисторе / Н. М. Ройзин, М. И. Маркович // Полупроводниковые приборы и их применение под ред. Я. А. Федотова. — М.: Сов. радио, 1963. Вып. 10. — С. 167–200.
6. *Старосельский, В. И.* Моделирование коллекторной области биполярного транзистора // Микроэлектроника. — 1998. — Т. 27. — № 2. — С. 125–133.
7. *Early, J. M.* Effects of space-charge layer widening in junction transistors / J. M. Early // Proc. IRE. — 1952. — Vol. 40. — P. 1401.
8. *Kirk, C. T.* A theory of transistor cutoff frequency (f_r) fall-off at high current density / C. T. Kirk // IEEE Trans. — 1962. — Vol. ED-9. — P. 164.

9. *Gummel, H. K.* Poon H.C. An integral charge control model of bipolar transistors / H. K. Gummel // Bell Syst. Techn. J. — 1970. — Vol. 49. — P. 827.
10. *Зи, С. М.* Физика полупроводниковых приборов. Кн.1 / С. М. Зи ; пер. с англ. — М. : Мир, 1984. — 456 с.
11. *Лохстрох, Я.* Приборы и схемы для СБИС на биполярных транзисторах / Я. Лохстрох // ТИИЭР. — 1981. Т. 69. — № 7. — С. 47—62.
12. *Peltre, P.* Isolation method strinks bipolar cells for fast, dens memories / P. Peltre // Electronics. — 1981. — Vol. 44. — № 5. — P. 52—55.
13. *Akasaka, J.* Higher speed bipolar process with fall implantation and self-aligned contact structure / J. Akasaka // IEEE IEDM. Dig. Techn. Papers. — 1978. — P. 189—192.
14. *Takemoto, T.* A vertically isolated self-aligned transistor/ T. Takemoto [and other] // VIST. IEEE Trass. — 1982. — Vol. ED-29. — № 11. — P. 1761—1775.
15. *Sasaka, H.* Gigabit logic bipolar technology advances super self-aligned process technology / H. Sasaka [and other] // Electron. Lett. — 1985. — Vol. 19. — P. 283—289.
16. *Sasaka, H.* Prospects of SST technology for high-speed LSI/ H. Sasaka [and other] // IEEE IEDM. — 1985. — P. 18—21.
17. *Masak, S.* An 86 K components bipolar VLSI master-slice with 290 ps loaded date delay / S. Masak [and other] // IEEE J. — 1987. — Vol. SC-22. — № 1. — P. 41—46.
18. *Вернер, В. Д.* Способ изготовления интегральной схемы «Сте-планар» / В. Д. Вернер [и др.]. — А.с. СССР № 1326132.



ГЛАВА V

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРИНЦИПОВ ДЕЙСТВИЯ И СВОЙСТВ БИПОЛЯРНЫХ И ПОЛЕВЫХ ТРАНЗИСТОРОВ

МДП- и биполярные транзисторы являются основными универсальными активными элементами современной микроэлектроники, на основе которых возможно выполнение произвольных операций обработки информации. Поэтому представляет интерес провести сравнительный анализ их принципов действия и основных свойств, определяющих преимущества и недостатки этих приборов в электронных устройствах.

В этом разделе мы рассмотрим сходство функционального устройства приборов, сходства и различия в принципах их действия и электрических свойствах, а также сравним технико-экономические аспекты их применения.

Вначале мы систематизируем сходство и различие в принципах действия МДПТ и БТ, а также в их электрических свойствах, рассмотрев по отдельности идеализированные модели приборов и неучтенные в них эффекты, а затем попытаемся объяснить различия в электрических характеристиках физическими причинами.

Результатом такого анализа должны стать общие выводы о целесообразности применения того или иного прибора в микроэлектронных изделиях различного типа.

V.1. Сходства и различия в устройстве и принципе действия МДП- и биполярных транзисторов

Оба рассматриваемых прибора состоят из элементов одинакового функционального назначения (см. параграф III.2.2, IV.1.1). Они содержат *источник* и *приемник* носителей заряда, *область переноса заряда*, через которую протекает



основной ток, и *управляющий электрод*, потенциал которого регулирует величину этого тока. В ИС оба прибора размещены на *подложке*, хотя для биполярного транзистора подложка является обязательным элементом только в планарной конструкции (табл. V.1.1). Общие признаки в устройстве приборов отражены в их названии «транзистор» (transfer resistor) — переходное сопротивление (области переноса заряда).

Таблица V.1.1

Элементы структуры МДП- и биполярных транзисторов

№	Элемент прибора	МДПТ	БТ	Примечания
1	Источник носителей заряда	Исток	Эмиттер	
2	Приемник носителей заряда	Сток	Коллектор	
3	Область переноса носителей	Канал, длина L	База, длина w_B	
4	Управляющий электрод	Затвор	Контакт к базе	
5	Конструктивное основание	Подложка	Подложка	В БТ не является обязательным
6	Барьерные переходы	Подложка-исток, подложка-сток	Эмиттер-база, коллектор-база	

В устройстве и принципе действия двух типов транзисторов имеются принципиальные различия, которые связаны со свойствами области переноса носителей заряда и иллюстрируются табл. V.1.2.

Таблица V.1.2

Принципиальные различия в устройстве и принципе действия МДП- и биполярных транзисторов

№	Свойства области переноса носителей заряда	МДПТ	БТ
1	Область переноса носителей	Канал	База
2	Тип носителей заряда в области переноса [*])	Основные (<i>униполярный</i> транзистор)	Неосновные (<i>биполярный</i> транзистор)



Продолжение таблицы V.1.2

№	Свойства области переноса носителей заряда	МДПТ	БТ
3	Контакт к области переноса носителей	Емкостный	Гальванический
4	Основной механизм переноса	Дрейф	Диффузия
5	Способ воздействия на область переноса	Эффект поля	Изменение высоты барьера между источником носителей заряда и областью переноса
6	Направление воздействия на область переноса	Перпендикулярное току	Параллельное току
7	Зона воздействия	Вся длина канала	Граница с эмиттерным переходом
8	Расположение подложки	Параллельна плоскости протекания тока. Примыкает к каналу	Перпендикулярна плоскости протекания тока. Отделена от базы коллекторным слоем
9	Расположение управляющего электрода	Затвор параллелен плоскости протекания тока	База перпендикулярна плоскости протекания тока
10	Симметрия источника и приемника тока	Есть	Нет

^{*)}Примечание: основными считаются носители с большей концентрацией.

V.2. Основные электрические свойства МДП- и биполярных транзисторов в рамках классических моделей

Основные электрические свойства МДП- и биполярных транзисторов в рамках классических моделей представлены в табл. V.2.1 (см. разделы III.3, IV.2 — IV.4). Для определенности полагается, что МДПТ имеет n -канал, а БТ — тип проводимости n - p - n .



Таблица V.2.1

Электрические свойства МДП- и биполярных транзисторов (классические модели)

№	Электрические свойства	Общие характеристики	МДПТ	БТ
1	Основная схема включения	Общий источник носителей заряда	Общий исток (ОИ)	Общий эмиттер (ОЭ)
2	Оптимальный режим работы		$V > V_p$; $V_{ds} > V_{gs}$ (пологая область ВАХ)	$V_{be} > 0$; $V_{bc} < 0$ (нормальный режим)
3	Вид управляющего сигнала	Напряжение V_{in}	Напряжение затвор-исток V_{gs}	Напряжение база-эмиттер V_{be}
4	Вид выходного сигнала	Ток I_{out}	Ток стока I_D	Ток коллектора I_C
5	Отношение входного сопротивления к выходному		$R_{in}/R_{out} = \infty$	$R_{in}/R_{out} \ll 1$
6	Параметр, характеризующий усилительные свойства	Крутизна ВАХ $g = \partial I_{out} / \partial V_{in}$	Крутизна ВАХ $g = \partial I_D / \partial V_{gs}$	Коэффициент усиления тока $\beta = \partial I_C / \partial I_B$
7	Вид проходной ВАХ $I_{out}(V_{in})$		Квадратичная	Экспоненциальная
8	Влияние потенциала подложки		Изменение порогового напряжения V_t Деформация ВАХ	Нет
9	Факторы, определяющие быстродействие	Время пролета носителей заряда через область переноса	Время пролета через канал $T_c = 2L^2 / 3\mu V_{gs}$	Время пролета через базу $T_N = w_B^2 / 2D_B$ Постоянная времени накопления заряда в коллекторе τ_l
10	Параметр, определяющий усилительные свойства и быстродействие	Длина области переноса	Длина канала L	Длина базы w_B



В МДПТ изменение потенциала подложки приводит к изменению порогового напряжения V_t и деформации ВАХ (см. параграф III.2.4, III.4.1). В БТ подложка не оказывает на ВАХ никакого воздействия (п. 8 табл. V.2.1). Причина этого различия состоит в том, что в МДПТ подложка примыкает к каналу и расположена параллельно плоскости протекания тока (п. 8 табл. V.1.2, рис. III.8.1 — III.8.3). В БТ подложка отделена от базы коллекторным слоем и расположена перпендикулярно плоскости протекания тока (см. п. 8 табл. V.1.2, рис. IV.8.1).

Отметим также, что в МДПТ расположение затвора параллельно плоскости протекания тока вызывает трудности разделения заряда в канале на составляющие Q_{gs} , Q_{gd} и представления емкости затвор—канал тензором 2-го ранга (см. параграф III.6.1). В БТ эта проблема практически отсутствует, так как заряды избыточных носителей Q_{BN} и Q_{BI} , инжектированные в базу через эмиттерный и коллекторный переходы (см. параграф IV.4.7), слабо зависят от напряжений на противоположных переходах (только вследствие эффекта Эрли).

V.3. Эффекты, ограничивающие характеристики реальных МДП- и биполярных транзисторов

Главным параметром, определяющим усилительные свойства и быстродействие транзисторов, является длина области переноса (канала или базы, п. 10 табл. V.2.1). При ее уменьшении в транзисторах обоих типов проявляется действие принципиально неустранимых физических эффектов, ограничивающих предельно достижимые характеристики (см. разд. III.5, IV.5). Эти эффекты и их воздействия на характеристики приборов указаны в табл. V.3.1.

V.4. Важнейшие характеристики МДП- и биполярных транзисторов

Прежде чем обсуждать предельно достижимые характеристики транзисторов, следует установить, какие именно характеристики важны для электронных устройств. Разумеется, можно выделить множество характеристик транзисто-



Таблица V.3.1

Эффекты, ограничивающие характеристики реальных МДП- и биполярных транзисторов

№	Физический эффект	МДПТ		БТ	
		Форма проявления	Воздействие на характеристики	Форма проявления	Воздействие на характеристики
1	Модуляция ширины барьерных переходов выходным напряжением	Эффект смятия канала	Вид пробоя	Эффект Эрли	Конечная выходная проводимость.
2	Ограничение скорости носителей заряда	В канале	Ограничение тока насыщения. Вырождение проходной ВАХ в линейную. Ограничение предельной крутизны. Снижение предельной частоты	Прокол базы В базе и в коллекторном переходе	Вид пробоя Снижение предельной частоты
		Модуляция эффективной длины канала	Конечная выходная проводимость. Ограничение предельной частоты	В коллекторе (эффект Кирка)	Резкое снижение предельной частоты
3	Избыточное накопление заряда в инверсном режиме	—	—	Накопление заряда в коллекторе в режиме насыщения	Снижение скорости переключения



ров, влияющих на параметры устройств различного типа. Не претендуя на полноту картины и не устанавливая порядок важности, выделим те характеристики приборов, которые имеют принципиальное значение:

- коэффициент усиления K ;
- крутизна ВАХ g ;
- напряжение насыщения выходных ВАХ $V_{out,S}$;
- входное сопротивление R_{in} ;
- предельная частота ω_T ;
- постоянная времени накопления заряда в приемнике тока $\tilde{\tau}_I$.

Строго говоря, под *коэффициентом усиления* следует понимать коэффициент усиления *мощности* сигнала. Так делается в технике СВЧ, где сигналы (электромагнитные колебания) описываются общими уравнениями электродинамики, блоки устройств зачастую имеют распределенный характер, волновые сопротивления трактов сигнала (в том числе вакуума $\rho_0 = \sqrt{\mu\mu_0 / \epsilon\epsilon_0} = 377 \text{ Ом}$) имеют ограниченный диапазон ($\sim 10 \text{ Ом} < \rho < \rho_0$), а понятия напряжений и токов существуют далеко не всегда (электрические поля могут не быть потенциальными). Мы, однако, ограничимся областью «электротехники», где сигналы характеризуются напряжениями и токами. Более того, несмотря на дуальность уравнений теории электрических цепей (симметрию понятий напряжение—ток, узел—ветвь схемы, последовательное—параллельное соединение, импеданс—адмиттанс, емкость—индуктивность), современная электроника (и микроэлектроника) за редкими исключениями использует «потенциальный» подход. Источники питания представляют собой источники ЭДС; элементы схем подключаются к ним параллельно, а под сигналами чаще всего понимаются напряжения. Реактивности, определяющие инерционность устройств и их элементов, имеют обычно емкостный (а не индуктивный) характер. Такая ситуация обусловлена соотношением между входными R_{in} и выходными R_{out} сопротивлениями элементов устройств и волновыми сопротивлениями ρ трактов сигнала (линий связи на кристалле ИС, проводов или кабелей). Как правило, $R_{in}, R_{out} \gg \rho$; это означает, что электрическая энергия сигнала значительно превышает магнитную.

Для электронного прибора коэффициент усиления следует определять в зависимости от соотношения между R_{in} и R_{out} . В МДПТ $R_{in} / R_{out} \rightarrow \infty$ (см. табл. V.2.1), и его целесооб-



разно характеризовать собственным *коэффициентом усиления напряжения* K_v . В БТ $R_{in} / R_{out} \ll 1$, и его целесообразно характеризовать собственным *коэффициентом усиления тока* K_I . Коэффициент усиления тока в МДПТ $K_I \rightarrow \infty$; значение коэффициента усиления напряжения в БТ (соотношение (IV.7.21) также очень велико (обычно $K_v > 100$). В электронных устройствах значения этих коэффициентов усиления маскируются сопротивлениями пассивных нагрузочных или согласующих элементов и практического значения не имеют.

Исключение составляют измерительные устройства, где входное сопротивление должно быть либо возможно большим (при измерении напряжений), либо возможно меньшим (при измерении токов). Поэтому в перечень наиболее важных характеристик транзисторов включено *входное сопротивление*.

Как МДПТ, так и БТ, обладают высокими выходными сопротивлениями (выходным сигналом является ток). Поэтому в практических электронных устройствах их усилительные свойства характеризуются *крутизной ВАХ* g (см. табл. V.2.1), также включенной в перечень. Максимальная крутизна достигается при некотором минимальном значении выходного напряжения — *напряжении насыщения*. Его величина ограничивает возможности снижения питающего напряжения и энергопотребления.

К наиболее важным характеристикам естественно отнести *предельную частоту* ω_T , а кроме того, постоянную времени τ_I , характеризующую время накопления заряда в коллекторе БТ в режиме насыщения.

V.5. Анализ важнейших характеристик МДП- и биполярных транзисторов

Соотношения, позволяющие оценить предельные значения этих характеристик, приведены в табл. V.5.1. В формулах учтены только принципиально неустранимые факторы, ограничивающие предельно достижимые значения характеристик. Так, например, в соотношениях для предельной частоты полагаются нулевыми барьерные емкости p - n переходов.



Таблица V.5.1

Предельные характеристики МДП- и биполярных транзисторов

№ п/п	Характеристика	МДПТ		БТ	
		Значение	Ссылка	Значение	Ссылка
1	Крутизна ВАХ g	$g_{\max} = Z\alpha_S C_{s,G} = I_{D0} / V_{gsr}$	(III.5.14), (III.5.26a)	$g \approx I_E / m\phi_T$	(IV.7.18)
2	Коэффициент усиления K	$K_{v0} = \text{ch}(L/a)$	(III.5.32б), (III.5.24)	$K_I \approx \beta$	
3	Напряжение насыщения выходных ВАХ $V_{out,s}$	$V_{ds,s} = E_S L \cdot (a/L) \cdot \text{sh}(L/a)$	(IV.5.25)	$V_{as} \approx 3\phi_T \ln(1+1/\beta)$	(III.2.20)
4	Входное сопротивление R_m	$R_{in} \approx \rho_d LZ / d$	—	$R_m \approx (1+\beta)m\phi_T / I_E$	Рис. IV.7.8
5	Предельная частота ω_T	$\omega_{T\max} = v_m / L$	(III.5.19)	$\omega_{T\max} \approx v_m / (w_B + l_c)$	(IV.7.25), (IV.5.21в)
6	Постоянная времени накопления заряда в приемнике тока $\tilde{\tau}_I$	—	—	$\tilde{\tau}_I \approx (1-\gamma_I)\tau_c$	(IV.4.28)

¹ Обозначения для МДПТ: Z — ширина канала; $C_{s,G}$ и ρ_d — удельная емкость и удельное сопротивление диэлектрика; v_m — максимальная скорость носителей заряда; $a = \sqrt{d_{ef} l}$; d , d_{ef} — истинная и эффективная толщина диэлектрика; l — ширина ОПЗ под каналом, а для БТ: l_c — ширина коллекторного перехода, $1 \leq m < 2$.



Рассмотрим подробнее свойства характеристик, приведенных в табл. V.5.1, для обоих типов транзисторов.

1. Крутизна ВАХ

В МДПТ крутизна ВАХ пропорциональна ширине канала, в то время как в БТ она не зависит от площади эмиттера и пропорциональна току эмиттера. Эти различия обусловлены принципиально разными способами воздействия управляющего напряжения на области переноса заряда — канал в МДПТ и базу в БТ (см. табл. V.1.2).

В МДПТ управляющий электрод (затвор) имеет *емкостную* связь с каналом и воздействует на *проводимость* канала (поверхностную плотность носителей заряда) *в направлении, перпендикулярном току* в цепи сток—исток *по всей длине канала* (п. 3, 5—7 табл. V.1.2). В результате поверхностная плотность заряда в канале *линейно* зависит от напряжения затвор—канал (см. параграф III.3.2). При этом в идеализированной модели *проходная ВАХ квадратична* (п. 7 табл. V.2.1), а ограничение дрейфовой скорости носителей в канале вырождает ее в линейную (п. 2 табл. V.3.1). Таким образом, максимально достижимая *крутизна пропорциональна ширине канала Z и не зависит от тока* (п. 1 табл. V.5.1).

В БТ управляющий электрод (база) *гальванически* связан с базой. Он воздействует на *высоту потенциального барьера* между источником носителей заряда (эмиттером) и областью переноса (базой) *в направлении, параллельном току* в цепи коллектор—эмиттер (п. 3, 5, 6 табл. V.2.1). Воздействие изменяет концентрацию неосновных носителей *на границе с эмиттерным переходом* (п. 7 табл. V.2.1) пропорционально $\exp(V_{be} / m\varphi_T)$, где $m = 1$ для низкого уровня инжекции в базе, $m = 2$ для высокого. Чрезвычайно сильное (экспоненциальное) изменение свойств области переноса под действием входного сигнала обусловлено тем, что в БТ ток через область переноса (базу) создают *неосновные* носители, хотя основные присутствуют в базе в большем количестве (*биполярный* транзистор). В МДПТ ток через канал создается *основными* носителями, число которых неизмеримо больше, чем неосновных (*униполярный* транзистор, п. 2 табл. V.2.1), и не может изменяться столь значительно.

В результате *проходная ВАХ БТ имеет экспоненциальный характер* (п. 7 табл. V.2.1). Разумеется, при заданном напряжении V_{be} крутизна пропорциональна площади эмиттера, но при заданном значении эмиттерного тока I_E *крутизна про-*



порциональна I_E (п. 1 табл. V.5.1) и не зависит от площади эмиттера.

Из формул п. 1 табл. V.5.1 следует, что для МДПТ и БТ достижимые значения крутизны одинаковы при условии

$$I_D = I_E V_{gst} / m\phi_T. \quad (V.5.1)$$

Практически МДПТ работают при напряжении $V_{gst} = V_{gs} - V_t > 1$ В. Это означает, что при равной крутизне ток стока МДПТ в 20—40 раз превышает ток эмиттера БТ.

Интересно отметить, что энергетические диаграммы МДПТ и БТ в продольном сечении весьма схожи (рис. V.5.1). Их сходство особенно показательно, когда МДПТ работает в подпороговом режиме (см. параграф III.4.2). В этом случае исчезают некоторые различия в принципе действия МДПТ и БТ, отраженные в табл. V.1.2. В частности, в МДПТ основным механизмом переноса носителей через канал является *диффузия* (а не дрейф, п. 4 табл. V.1.2). Входное воздействие (V_{gs}),

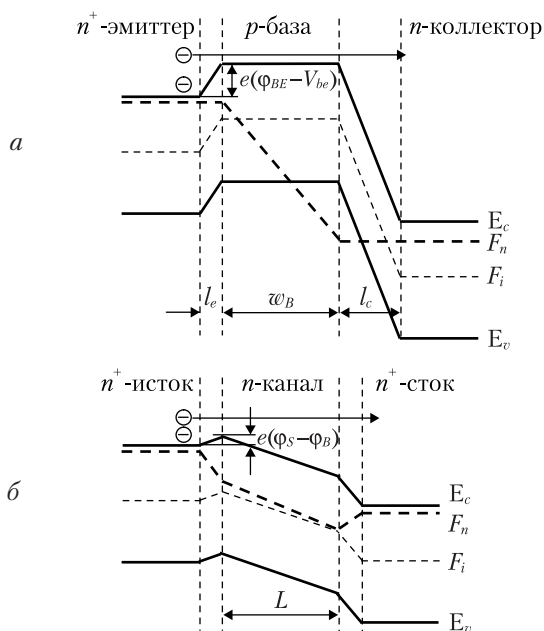


Рис. V.5.1. Энергетические диаграммы биполярного транзистора в нормальном режиме (а) и МДП транзистора в подпороговом режиме (б)

как и в БТ, сводится к изменению высоты барьера $e(\varphi_s - \varphi_B)$ между истоком и каналом (рис. III.4.2, V.9.1; п. 5 табл. V.1.2), а зоной воздействия, как и в БТ, является граница канала со стоком (п. 6 табл. V.1.2). В результате проходная ВАХ МДПТ в подпороговом режиме имеет, как и в БТ, *экспоненциальный характер* (п. 7 табл. V.2.1) — соотношение (III.4.18).

2. Коэффициент усиления

Как отмечено выше, МДПТ целесообразно характеризовать коэффициентом усиления напряжения K_V , а БТ — коэффициентом усиления тока K_I .

В короткоканальных КМДП физической причиной снижения коэффициента усиления K_V при уменьшении длины канала является модуляция эффективной длины канала напряжением сток—исток (см. параграф III.5.7), связанная с ограничением дрейфовой скорости носителей.

Отметим, что коэффициент усиления определяется отношением длины канала L к параметру $a = \sqrt{d_{ef}l}$, где d_{ef} — эффективная толщина диэлектрика; l — ширина ОПЗ (п. 2 табл. V.5.1). Для обеспечения приемлемого значения $K_V \geq 5$ отношение L/a должно быть не менее 2,3. Таким образом, при достижимых значениях $d_{ef} \approx l \approx 0,05$ мкм длина канала не должна превышать $\approx 0,115$ мкм.

В БТ значение $K_I \approx \beta$, рассчитанное по формулам раздела IV.3, обычно составляет более 50 и снижается при слишком малой либо слишком большой плотности эмиттерного тока (см. параграф IV.5.3). Тем не менее для конкретных применений площадь эмиттерного перехода можно выбрать таким образом, что в диапазоне рабочих токов значение β остается большим 20.

Следует отметить, что в МДПТ ограничение скорости носителей приводит к ограничению тока насыщения при заданном входном напряжении V_{gs} . В БТ этот эффект отсутствует (п. 2 табл. V.3.1). Физическая причина такого различия состоит в том, что в МДПТ управляющее напряжение V_{gs} воздействует на всю длину области переноса (канала); в БТ управляющее напряжение V_{be} изменяет концентрацию носителей в области переноса (базы) лишь на границе с эмиттерным переходом (п. 7 табл. V.2.1).

3. Напряжение насыщения выходных ВАХ

В короткоканальных МДПТ напряжение насыщения выходных ВАХ V_{ds} определяется напряженностью электрического поля E_s , соответствующей насыщению дрей-



фовой скорости носителей заряда в канале (п. 3 табл. V.5.1). Для n -канального МДПТ $E_s \approx 2 \cdot 10^4$ В/см, для p -канального $E_s \approx 4 \cdot 10^4$ В/см (см. раздел I.4). Таким образом, для принятых выше значений $d_{ef} \approx l \approx 0,05$ мкм при $K_V \geq 5$ ($L/a \geq 2,3$) получаем

$$V_{ds} \geq (0,22 - 0,44) \text{ В.}$$

В БТ напряжение V_{ceS} зависит только от инверсного коэффициента передачи тока базы β_I (п. 3 табл. IV.5.1). Для планарных n - p - n транзисторов со скрытым n^+ -слоем в коллекторе типичные значения составляют $\beta_I = 0,5 - 1$ (см. параграф IV.3.4), и $V_{ceS} = 60 - 100$ мВ.

В симметричных ($\beta_I \approx \beta_N$) латеральных p - n - p транзисторах напряжение насыщения очень мало: при $\beta_I > 8$ его значение не превышает 10 мВ.

4. Входное сопротивление

В МДПТ входной электрод (затвор) отделен от канала диэлектриком. Поэтому их входное сопротивление весьма велико (обычно $R_{in} > 10$ ГОм) и они могут считаться идеальными *приемниками напряжения* ($R_{in} \rightarrow \infty$ и входное напряжение V_{gs} не зависит от выходного сопротивления источника сигнала).

Входное сопротивление БТ в схеме включения ОЭ может быть найдено из эквивалентной схемы на рис. IV.7.8. При $R_B \rightarrow 0$

$$R_{in} \approx (1 + \beta) m \phi_T / I_E.$$

При $I_E = 0,1$ мА значение R_{in} лежит в диапазоне (1–2) кОм. В схеме включения ОБ величина входного сопротивления в $1 + \beta$ раз меньше, однако БТ все же не могут считаться идеальными приемниками тока.

5. Предельная частота

Как в МДП, так и в биполярных транзисторах предельно достижимое значение предельной частоты ограничивается максимальной скоростью носителей заряда $v_m \approx 10^7$ см/с (см. параграф III.5.9, IV.5.6) и достигается предельно возможным уменьшением длины области переноса заряда (п. 5 табл. V.5.1). В МДПТ возможность уменьшения длины канала ограничивается снижением коэффициента усиления напряжения K_V . В БТ уменьшение длины базы ведет к снижению напряжения прокола и повышению сопротивления

активной базы (которое в этом разделе не учитывается). По-видимому, для обоих типов кремниевых транзисторов минимальная длина области переноса (длина канала L для МДПТ, длина базы w_b для БТ) составляет менее 0,1 мкм. При этом предельно достижимое значение предельной частоты МДПТ составляет $f_{T\max} = \omega_{T\max} / 2\pi \approx > 150$ ГГц. В БТ носители заряда должны также преодолеть ОПЗ коллекторного перехода, ширина которой l_c может быть уменьшена примерно до длины базы.

Таким образом, теоретически достижимое значение предельной частоты в БТ примерно вдвое меньше, чем в МДПТ. В реальных транзисторах предельная частота ограничена также паразитными емкостями ОПЗ и емкостями перекрытия затвора с контактными областями стока и истока (в МДПТ, см. рис. III.6.1), а также сопротивлениями электронейтральных областей. В настоящее время максимальные значения предельной частоты для обоих типов транзисторов примерно одинаковы.

6. Постоянная времени накопления заряда в приемнике тока

В параграфе IV.4.7 показано, что в коллекторе планарных БТ в режиме насыщения накапливается заряд избыточных дырок. Его величина пропорциональна времени жизни в слаболегированном коллекторе τ_c и может быть весьма значительна (п. 6 табл. V.5.1). Эффект накопления заряда в коллекторе значительно затягивает переходный процесс выключения транзистора, что резко снижает быстродействие логических вентилях. Для нейтрализации этого эффекта в быстродействующих цифровых устройствах применяются специальные меры, предотвращающие попадание транзистора в режим насыщения (см. параграф VI.4).

Помимо принципиальных различий в устройстве и принципе действия МДПТ и БТ, их возможности определяются технологическими факторами. Как очевидно из рис. III.8.1—III.8.3, IV.8.1—IV.8.5, структура МДПТ существенно более проста, чем структура БТ. Поэтому изготовление ИС на основе МДП-транзисторов обходится дешевле. Чрезвычайную важность представляет возможность изготовления по МДП технологии высококачественных комплементарных транзисторных пар. В биполярной технологии такие возможности ограничены.

Результаты проведенного анализа сведены в табл. V.5.2.



Таблица V.5.2

**Сравнительные характеристики
МДП- и биполярной технологии**

№	Важнейшие характеристики	Оптимальная технология	Примечания
1	Крутизна ВАХ транзистора	Биполярная	Не зависит от площади эмиттера
2	Коэффициент усиления транзистора	Биполярная	
3	Напряжение насыщения выходных ВАХ транзистора	Биполярная	
4	Входное сопротивление транзистора	МДП	
5	Предельная частота транзистора	МДП, биполярная	
6	Возможность создания высококачественных комплементарных пар	МДП	
7	Экономичность	МДП	

**V.6. Области применения
МДП- и биполярных транзисторов**

Сравнительный анализ свойств и предельных возможностей МДП и биполярных транзисторов, а также технологических возможностей их изготовления показывает, что ни одному из этих приборов нельзя отдать абсолютного предпочтения.

Большую часть микроэлектронной продукции составляют цифровые СБИС для вычислительной техники. В настоящее время практически вся эта продукция изготавливается по технологии МДП. Решающую роль в такой ситуации играет возможность широкого применения комплементарных логических вентилях, практически не потребляющих энергии в статическом состоянии, и меньшая стоимость производства. Применение таких логических элементов особенно выгодно, когда частота их переключения значительно меньше тактовой частоты работы устройства. Применение технологии МДП обеспечивает максимальную степень интеграции.

Несмотря на примерно одинаковое собственное быстродействие, максимальная скорость работы цифровых блоков достигнута на биполярных вентилях. Преимущество БТ проявляется в их высоких усилительных свойствах, возможности достижения высокой крутизны при малой площади, а также



низком напряжении насыщения (п. 1—3 табл. V.5.2). Эти преимущества в максимальной степени реализуются в логических элементах ЭСЛ (эмиттерно-связанная логика), работающих по принципу переключения тока. Элементы ЭСЛ переключаются малыми логическими перепадами напряжения (0,5 В и менее), что обеспечивает снижение энергии, затрачиваемой на перезаряд емкостей линий связи. Малое напряжение насыщения БТ позволяет при допустимых напряжениях питания применять многоярусные ЭСЛ- вентили, которые обладают расширенными функциональными возможностями. Кроме того, в ЭСЛ- вентилях транзисторы работают только в нормальном режиме, что исключает накопление заряда в коллекторе (п. 6 табл. V.5.1). Основным недостатком биполярной логики является значительное энергопотребление, исключающее ее применение в СБИС.

Важным преимуществом БТ является возможность получения высокой крутизны и выходного тока без существенного увеличения площади прибора (п. 1 табл. V.5.1). В этом отношении БТ являются идеальными драйверами, способными работать на значительную активную или емкостную нагрузку. В частности, БТ широко применяются в интерфейсных выходных каскадах устройств высокого быстродействия, где передача сигналов между блоками осуществляется по согласованным низкоомным волноводам.

В широкополосных аналоговых и аналого-цифровых устройствах обработки информации БТ имеют определенные преимущества перед МДПТ, которые также связаны с высокими усилительными свойствами БТ. В качестве примера можно привести аналого-цифровые преобразователи (АЦП), осуществляющие преобразование аналогового сигнала в цифровую форму. Основой таких устройств являются компараторы напряжения, которые должны осуществлять сравнение сигнала с эталонным напряжением.

Проведенный сравнительный анализ принципов действия и свойств биполярных и МДП- транзисторов показывает, что у каждого из этих приборов имеется своя область применения, где они обладают существенными преимуществами. В сложных аналого-цифровых устройствах часто применяются оба типа транзисторов — цифровая часть выполняется на комплементарных МДП- вентилях, а аналоговая — на БТ. Повышение степени интеграции таких устройств достигается применением Би-КМОП- технологии, позволяющей создавать на одном кристалле комплементарные пары МДПТ и биполярные транзисторы.



Глава VI

КОНТАКТЫ МЕТАЛЛ—ПОЛУПРОВОДНИК

VI.1. Контакты металл—полупроводник

VI.1.1. Типы и основные характеристики контактов металл—полупроводник

Поскольку соединения электродов прибора с другими элементами электронного устройства осуществляются металлическими проводниками, контакты полупроводниковой области с металлом содержит любой полупроводниковый прибор. В идеале падение напряжения на таких контактах должно быть нулевым, однако реальные контакты полупроводниковых областей с металлическими выводами имеют конечное сопротивление, хотя их ВАХ близка к линейной («омические» контакты). При определенных условиях контакт металл—полупроводник может обладать нелинейной ВАХ («выпрямляющие» контакты). Такие контакты не могут быть использованы для соединения электродов прибора с металлическими проводниками, однако их свойства лежат в основе целого ряда весьма полезных полупроводниковых приборов.

Таким образом, мы будем различать два типа контактов:

- *невыпрямляющие (омические);*
- *выпрямляющие или барьерные.*

Механизм выпрямления на контакте металл—полупроводник впервые был объяснен У. Шоттки [1], поэтому выпрямляющие барьерные контакты также называют *контактами Шоттки*. Теоретическая модель барьерного контакта металла к тонким слоям полупроводника предложена Н. Моттом [2]. Основополагающие теории и обзор экспериментальных исследований в этой области изложены в работе [3].

Тип контакта зависит от свойств металла, полупроводника и поверхности их раздела.

Контактные свойства металла *характеризуются термодинамической работой выхода* Φ_M , которая определяется как



разность между потенциальной энергией электрона в вакууме E_0 и энергией Ферми F_M в металле. Контактные свойства полупроводника характеризуются *истинной работой выхода* Ψ (*электронным сродством*), которая определяется как разность между энергией E_0 и энергией дна зоны проводимости E_c в полупроводнике. Для полупроводника также применимо понятие термодинамической работы выхода (Φ_n для полупроводника n -типа, Φ_p — для p -типа), однако значения Φ_n или Φ_p зависят от степени его легирования.

Свойства поверхности раздела металл—полупроводник (подобно границе раздела диэлектрик—полупроводник в структурах МДП) характеризуются *плотностью поверхностного заряда* Q_{ss} , или плотностью *поверхностных состояний* N_{ss} . Заряд каждого поверхностного состояния, расположенного ниже уровня Ферми, считается равным заряду электрона. Таким образом

$$Q_{ss} = -eN_{ss}. \quad (\text{VI.1.1})$$

Контакт металл—полупроводник называется *идеальным*, если $N_{ss} = 0$. Тип идеального контакта металл—полупроводник определяется соотношением между термодинамическими работами выхода (табл. VI.1.1).

Таблица VI.1.1

**Зависимость типа контакта металл—полупроводник
от соотношения между термодинамическими
работами выхода**

Тип контакта		Омический	Барьерный
Тип проводимости полупроводника	n	$\Phi_M < \Phi_n$	$\Phi_M > \Phi_n$
	p	$\Phi_M > \Phi_p$	$\Phi_M < \Phi_p$

Отметим, что большинство металлов с наиболее распространенными полупроводниковыми материалами n -типа образуют барьерные контакты, а с материалами p -типа — омические.

Вначале будут рассмотрены свойства барьерных контактов (контактов Шоттки). Свойствам омических контактов посвящен раздел VI.3.

VI.1.2. Параметры барьерных контактов

Энергетическая диаграмма *идеального* ($N_{ss} = 0$) контакта Шоттки металл—кремний n -типа в состоянии равновесия представлена на рис. VI.1.1, а. Полупроводниковый слой



называется *базой*. Благодаря выполнению условия $\Phi_M > \Phi_n$ энергетические зоны полупроводника в районе контакта изгибаются вверх, что приводит к образованию обедненной носителями заряда ОПЗ шириной l_0 . ОПЗ представляет собой потенциальный барьер как для электронов металла, так и для электронов полупроводника.

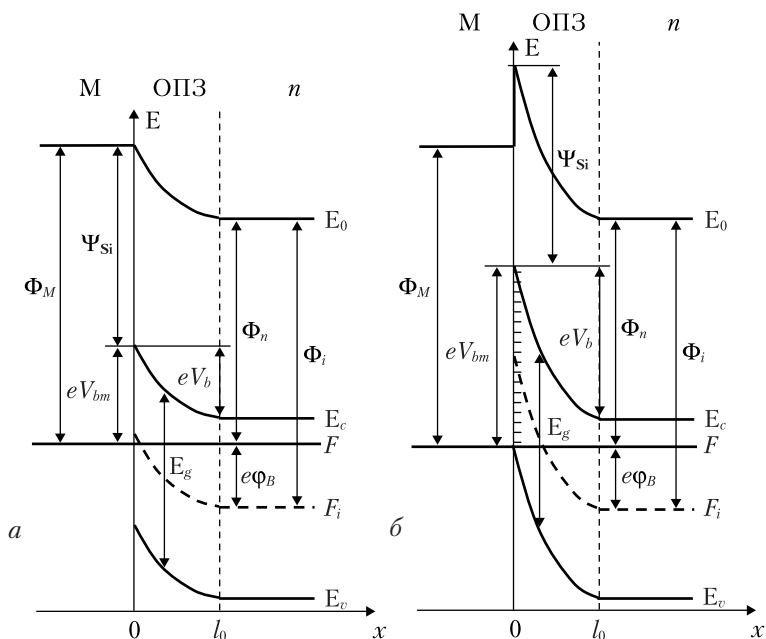


Рис. VI.1.1. Равновесные энергетические диаграммы контактов Шоттки металл–полупроводник *n*-типа:

a — идеальный контакт $N_{ss} = 0$; *б* — предельный случай $N_{ss} \rightarrow \infty$

Для электронов металла, имеющих энергию $E = F$, высота барьера составляет

$$V_{bm} = [E_c(0) - F] / e \quad (\text{VI.1.2a})$$

и для невырожденной базы $V_{bm} < E_g / e$.

Из энергетической диаграммы на рис. VI.1.1, *a* следует, что

$$V_{bm} = (\Phi_M - \Psi_{Si}) / e. \quad (\text{VI.1.2б})$$

Электронное сродство является константой полупроводника, поэтому для идеального контакта Шоттки *высота барьера в металле не зависит от степени легирования базы*.

Естественно, что *высота барьера в базе* V_b (контактная разность потенциалов металл—полупроводник) зависит от концентрации примеси в базе N_B . Она определяется разностью термодинамических работ выхода:

$$V_b(N_B) = [\Phi_M - \Phi_n(N_B)] / e, \quad (\text{VI.1.3})$$

где $\Phi_n = \Phi_i - e\phi_B$; Φ_i — термодинамическая работа выхода из собственного полупроводника (см. рис. VI.1.1 а), а величина

$$\phi_B = (F - F_i) / e = \phi_T \ln(N_B / n_i) \quad (\text{VI.1.4})$$

соответствует той же величине в структуре МДП (соотношения (III.1.5) и (III.1.21)). Таким образом,

$$V_b(N_B) = \phi_B(N_B) + (\Phi_M - \Phi_i) / e. \quad (\text{VI.1.5})$$

Из рис. VI.1.1, а следует, что *высота барьера в металле* и контактная разность потенциалов связаны соотношением

$$V_{bm} - V_b = (E_c - F_i) / e - \phi_B. \quad (\text{VI.1.6})$$

Соотношение (VI.1.6) показывает, что *разность* $V_{bm} - V_b$ *не зависит от свойств поверхности и определяется только свойствами базы.*

Равновесная поверхностная плотность заряда доноров в ОПЗ составляет

$$Q_{sB0} = eN_B l_0. \quad (\text{VI.1.7})$$

Заряд в ОПЗ компенсируется зарядом на поверхности металла с поверхностной плотностью $Q_{sM0} = -Q_{sB0}$.

При *конечной* плотности поверхностных состояний заряд в ОПЗ частично компенсируется зарядом поверхностных состояний Q_{ss} , заполненных электронами до уровня Ферми:

$$Q_{ss} = -e \int_{E_v(0)}^F g_s(E) dE, \quad (\text{VI.1.8})$$

где g_s — плотность поверхностных состояний по энергии. Это приводит к повышению барьера в металле V_{bm} .

В предельном случае *бесконечной* плотности поверхностных состояний ($N_{ss} \rightarrow \infty$, $g_s \rightarrow \infty$) заряд в ОПЗ полностью компенсируется зарядом поверхностных состояний

$$Q_{sB0} + Q_{ss} = 0. \quad (\text{VI.1.9})$$

Из (VI.1.8) следует, что при $g_s \rightarrow \infty$ величина Q_{ss} конечна только при $F = E_v(0)$ (рис. VI.1.1, б), т.е.



$$V_{bm} = E_g / e. \quad (\text{VI.1.10})$$

В *реальных* контактах Шоттки обычно

$$V_{bm} = (0,6-0,75)E_g / e. \quad (\text{VI.1.11})$$

Контактная разность потенциалов V_b соответствует контактной разности потенциалов Φ_{BE} в сильно несимметричном ступенчатом p - n переходе с $N_E \gg N_B$. Поэтому и равновесная ширина ОПЗ l_0 определяется так же, как равновесная ширина сильно несимметричного ступенчатого p - n перехода (соотношение (II.2.25)):

$$l_0 = \sqrt{2\epsilon\epsilon_0 V_b / eN_B}. \quad (\text{VI.1.12})$$

Внешнее напряжение V , приложенное к барьерному контакту, считается положительным, если потенциал металла выше потенциала n -базы (в случае p -базы — наоборот). Как и в p - n переходе, внешнее напряжение изменяет высоту потенциального барьера в полупроводнике V_b до величины $V_b - V$. Таким образом, как и в ступенчатом p - n переходе, ширина ОПЗ зависит от напряжения

$$l(V) = l_0 \sqrt{1 - V / V_b}. \quad (\text{VI.1.13})$$

По аналогии с p - n переходом контакт Шоттки имеет барьерную емкость, зависящую от напряжения,

$$C(V) = C_0 / \sqrt{1 - V / V_b}, \quad (\text{VI.1.14})$$

где $C_0 = C(0) = \epsilon\epsilon_0 S / l_0. \quad (\text{VI.1.15})$

Практически величина контактной разности потенциалов V_b находится из экспериментальной ВФХ $C(V)$, после чего высота барьера в металле может быть рассчитана из соотношения (VI.1.6). Как правило, при равной концентрации примеси в базе N_B *контактная разность потенциалов V_b в контакте Шоттки ниже, чем в p - n переходе.*

VI.1.3. Эффект Шоттки

Точечный заряд, помещенный вблизи металлической плоскости, индуцирует на поверхности металла заряд противоположного знака. Распределение этого индуцированного заряда по поверхности металла таково, что его действие на точечный заряд эквивалентно действию точечного заряда противоположного знака, расположенного за поверхностью



металла симметрично исходному (сила зеркального отображения).

Действие сил зеркального отображения в барьерном контакте металл—полупроводник иллюстрируется рис. VI.1.2. На электрон, находящийся на расстоянии x от поверхности металла, действует кулоновская сила

$$F(x) = -e^2 / 4\pi\epsilon_0(2x)^2,$$

которая снижает потенциальную энергию электрона на величину

$$\Delta E_{c0}(x) = \int_{-\infty}^x F(x)dx = e^2 / 16\pi\epsilon_0 x. \quad (\text{VI.1.16})$$

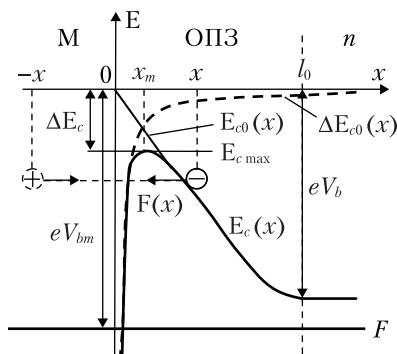


Рис. VI.1.2. Снижение равновесного потенциального барьера под действием силы зеркального отображения (эффект Шоттки)

Действие силы зеркального отображения проявляется только на малом расстоянии от поверхности металла, где в равновесном состоянии ($V=0$) зависимость исходной (без учета силы отображения) энергии $E_{c0}(x)$ близка к линейной:

$$E_{c0}(x) \approx E_{c0}(0) - eE_m x. \quad (\text{VI.1.17})$$

В (VI.1.17) E_m — максимальная напряженность электрического поля в ОПЗ (при $x=0$), которая по закону Гаусса составляет

$$E_{m0} = Q_{sB0} / \epsilon\epsilon_0, \quad (\text{VI.1.18})$$

Подстановка (VI.1.7) в (VI.1.18) с учетом (VI.1.12) дает

$$E_{m0} = \frac{eN_B I_0}{\varepsilon \varepsilon_0} = \sqrt{\frac{2eN_B V_b}{\varepsilon \varepsilon_0}}. \quad (\text{VI.1.19})$$

Из соотношений (VI.1.17) и (VI.1.16) получаем результирующую зависимость $E_c(x)$

$$E_c(x) = E_{c0}(x) - \Delta E_{c0}(x) \approx E_{c0}(0) - eE_m x - e^2 / 16\pi \varepsilon \varepsilon_0 x. \quad (\text{VI.1.20})$$

Функция $E_c(x)$ имеет максимум (рис. VI.1.2) в точке

$$x_m = \sqrt{e / 16\pi \varepsilon \varepsilon_0},$$

где значение $E_{c\max} = E_c(x_m)$ составляет

$$E_{c\max} = E_{c0}(0) - \frac{e^{3/2} (eN_B V_b)^{1/4}}{\pi^{1/2} (2\varepsilon \varepsilon_0)^{3/4}}.$$

Таким образом, под действием сил зеркального отображения высота барьеров в металле V_{bm} и в полупроводнике V_b снижается на величину

$$\Delta V_{bm} = \Delta V_b = \frac{e^{7/4} (N_B V_b)^{1/4}}{\pi^{1/2} (2\varepsilon \varepsilon_0)^{3/4}}. \quad (\text{VI.1.21})$$

Снижения высоты барьеров под действием сил зеркального отображения называется *эффектом Шоттки*.

Эффект Шоттки проявляется довольно слабо. Для типичных значений $N_B = 10^{17} \text{ см}^{-3}$, $V_b = 0,7 \text{ В}$ получаем $\Delta V_{bm} = \Delta V_b \approx 40 \text{ мВ}$.

VI.2. Теория выпрямления на контакте металл—полупроводник

VI.2.1. Общий подход

В разд. II.4 ВАХ полупроводникового диода вычислялась путем определения токов *неосновных* носителей по обе стороны p - n перехода. Для барьерных контактов металл—полупроводник такой подход неприменим, так как в полупроводниковой базе и в ОПЗ ток переносится *основными* носителями заряда в базе. В случае n -базы электронный ток в плоскости контакта является разностью токов, обусловленных инжекцией электронов из полупроводника в металл и из металла в полупроводник:

$$I = I_n - I_{nM}.$$



Поток электронов из полупроводника в металл зависит от высоты барьера в полупроводнике $V_b - V$, т.е. от приложенного напряжения: $I_n = I_n(V)$ (см. рис. VI.2.1). Поток электронов из металла в полупроводник зависит от высоты барьера в металле V_{bm} и не зависит от напряжения V : $I_{nM} = I_0$. В состоянии равновесия ($V = 0$) токи I_n и I_{nM} взаимно компенсируются. Таким образом,

$$I(V) = I_n(V) - I_0. \quad (\text{VI.2.1})$$

Метод расчета ВАХ барьерного контакта зависит от соотношения между шириной ОПЗ l и длиной свободного пробега электронов λ в ОПЗ. Существуют две теории выпрямления на барьерном контакте металл—полупроводник, соответствующие предельным случаям $l \ll \lambda$ и $l \gg \lambda$.

1. *Теория термоэлектронной эмиссии (ТТЭ) или диодная теория выпрямления* [4] применима в предельном случае $l \ll \lambda$. При этом для электронного тока в ОПЗ нельзя использовать уравнения диффузии и дрейфа, так как понятия кинетических коэффициентов (подвижности μ и коэффициента диффузии D) имеют статистический смысл. В ТТЭ поток электронов из полупроводника в металл, т.е. ток $I_n(V)$, вычисляется в плоскости $x = l$ с помощью уравнений, описывающих равновесное распределение электронов по скоростям.

2. *Диффузионная теория выпрямления* применима в предельном случае $l \gg \lambda$. При этом электронный ток в ОПЗ вычисляется как сумма диффузионной и дрейфовой составляющих с использованием обычных уравнений для тока.

VI.2.2. Теория термоэлектронной эмиссии (диодная теория выпрямления)

На рис. VI.2.1. представлена энергетическая диаграмма контакта Шоттки при положительном внешнем напряжении $V > 0$, определяющем расстояние между уровнем Ферми в металле F_M и квазиуровнем Ферми для электронов F_n в базе. При $l \ll \lambda$ в металл инжектируются все электроны, находящиеся в плоскости $x = l$, которые имеют нужное направление скорости (в сторону металла) и достаточную энергию для преодоления потенциального барьера $V_b - V$. Плотность тока этих электронов составляет

$$j_n = \frac{2e}{h^3} \int v_x f_n d\vec{p}, \quad (\text{VI.2.2})$$



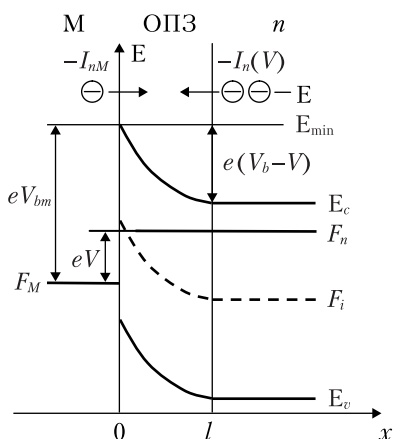


Рис. VI.2.1. Энергетическая диаграмма контакта Шоттки при $V > 0$

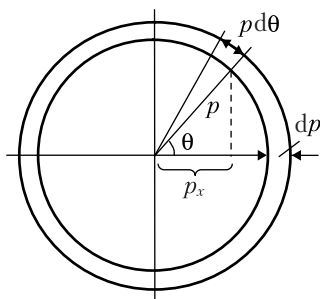


Рис. VI.2.2. Представление импульса в полярных координатах

где $2/h^3$ — плотность разрешенных состояний в пространстве импульса $\vec{p}$;

$$f_n = e^{(F_n - E)/kT} \quad (\text{VI.2.3})$$

— функция распределения. В полярных координатах (рис. VI.2.2) дифференциал импульса $d\vec{p}$ выражается через абсолютную величину p и полярный угол θ соотношением

$$d\vec{p} = 2\pi p^2 \sin \theta dp d\theta = 2\pi \sqrt{2} m_n^{3/2} (E - E_c)^{1/2} \sin \theta d\theta dE, \quad (\text{VI.2.4})$$

где
$$p = \sqrt{2m_n(E - E_c)}, \quad (\text{VI.2.5})$$

$$dp = \sqrt{m_n/2(E - E_c)} dE. \quad (\text{VI.2.6})$$

Нормальная к плоскости контакта составляющая скорости электронов составляет

$$v_x = p \cos \theta / m_n = \sqrt{2(E - E_c) / m_n} \cos \theta. \quad (\text{VI.2.7})$$

Подставляя (VI.2.3), (VI.2.4) и (VI.2.7) в (VI.2.2), получаем

$$\begin{aligned} j_n &= \frac{4e\pi m_n}{h^3} \int_{E_{\min}}^{\infty} e^{(F_n - E)/kT} (E - E_c) \int_0^{\theta_{\max}} \sin 2\theta_{\max} d\theta dE = \\ &= \frac{4e\pi m_n}{h^3} \int_{E_{\min}}^{\infty} e^{(F_n - E)/kT} (E - E_c) (1 - \cos^2 \theta_{\max}) dE. \end{aligned} \quad (\text{VI.2.8})$$

Как следует из рис. VI.2.1,

$$E_{\min} = E_c + e(V_b - V), \quad (\text{VI.2.9})$$

$$F_n - E = (F_n - E_{\min}) - (E - E_{\min}) = e(V - V_{bm}) - (E - E_{\min}). \quad (\text{VI.2.10})$$

Из рис. VI.2.2 и (VI.2.5) следует:

$$\begin{aligned} \cos \theta_{\max} &= (p_{\min} / p)^{1/2} = (E_{\min} - E_c)^{1/2} / (E - E_c)^{1/2}; \\ (E - E_c)(1 - \cos^2 \theta_{\max}) &= (E - E_c)[1 - (E_{\min} - E_c) / (E - E_c)] = \\ &= E - E_{\min}. \end{aligned} \quad (\text{VI.2.11})$$

Подстановка (VI.2.9)–(VI.2.11) в (VI.2.8) и интегрирование дают

$$j_n(V) = j_0 e^{V/\Phi_T}, \quad (\text{VI.2.12})$$

где $j_0 = A^* T^2 e^{-V_{bm}/\Phi_T}; \quad (\text{VI.2.13})$

$$A^* = (m_n / m_0) A \quad (\text{VI.2.14})$$

— *эффективная постоянная Ричардсона,*

$$A = 4\pi e m_0 k^2 / h^3 = 120 \text{ A}/(\text{K} \cdot \text{см})^2 \quad (\text{VI.2.15})$$

— *истинная постоянная Ричардсона.*

Подставляя в (VI.2.1) значение тока $j_n(V)$ из (VI.2.13), получаем ВАХ контакта Шоттки в соответствии с ТТЭ

$$I(V) = I_0 (e^{V/\Phi_T} - 1), \quad (\text{VI.2.16})$$

где $I_0 = j_0 S$; S — площадь контакта. Соотношение (VI.2.16) показывает, что в рамках ТТЭ ВАХ контакта Шоттки имеет такой же вид, как ВАХ идеализированного диода (II.4.20), хотя масштабный ток I_0 обычно значительно больше теплового тока I_S и слабее зависит от температуры.

VI.2.3. Диффузионная теория выпрямления

На рис. VI.2.3 представлена энергетическая диаграмма контакта Шоттки при $V > 0$ (а) и эпюры плотности объемного заряда $\rho(x)$, напряженности электрического поля $E(x)$ и потенциала $\phi(x)$ (б).

Согласно диффузионной теории выпрямления [1] электронный ток в ОПЗ, не зависящий от координаты x , вычисляется как сумма диффузионной и дрейфовой составляющих:

$$j_n = eD_n \frac{dn(x)}{dx} + en(x)\mu_n E(x) = eD_n \left(\frac{dn(x)}{dx} + \frac{n(x)E(x)}{\Phi_T} \right).$$

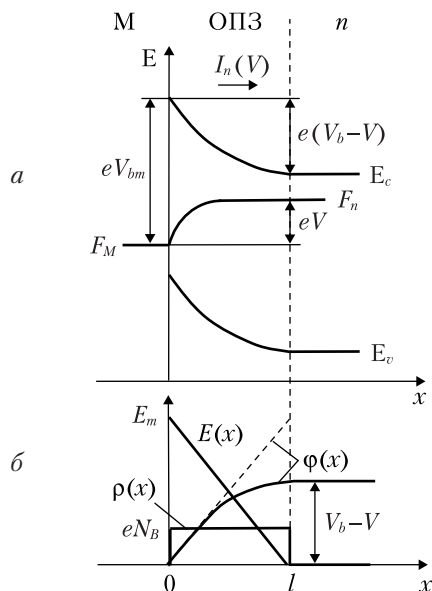


Рис. VI.2.3. Энергетическая диаграмма контакта Шоттки при $V > 0$ (а) и эшоры плотности объемного заряда $\rho(x)$, напряженности электрического поля $E(x)$ и потенциала $\phi(x)$ (б)

Умножая это уравнение на интегрирующий множитель $e^{-\phi(x)/\phi_T}$, где $\phi(x)$ — потенциал, и учитывая, что $E(x) = -d\phi(x)/dx$, получаем

$$\begin{aligned} j_n e^{-\phi(x)/\phi_T} &= eD_n \left[\exp \frac{-\phi(x)}{\phi_T} \cdot \frac{dn(x)}{dx} - \exp \frac{-\phi(x)}{\phi_T} \cdot \frac{n(x)}{\phi_T} \cdot \frac{d\phi(x)}{dx} \right] = \\ &= eD_n \cdot \frac{d}{dx} \left[n(x) \cdot e^{-\phi(x)/\phi_T} \right]. \end{aligned}$$

Интегрирование по всей ОПЗ (от $x = 0$ до $x = l$ на рис. VI.2.3) дает

$$\begin{aligned} j_n &= eD_n \frac{n(l)e^{-\phi(l)/\phi_T} - n(0)e^{-\phi(0)/\phi_T}}{\int_0^l e^{-\phi(x)/\phi_T} dx} = eD_n n(l) \frac{e^{[\phi(0)-\phi(l)]/\phi_T} - n(0)/n(l)}{\int_0^l e^{[\phi(0)-\phi(x)]/\phi_T} dx} = \\ &= eD_n n(l) \frac{e^{(V-V_b)/\phi_T} - n(0)/n(l)}{\int_0^l e^{[\phi(0)-\phi(x)]/\phi_T} dx}, \end{aligned} \quad (\text{VI.2.17})$$

где учтено, что $\varphi(0) - \varphi(l) = V - V_b$ (см. рис. VI.2.3). Для невырожденной базы согласно (I.3.27а)

$$n(x) = N_c \exp \frac{F_n(x) - E_c(x)}{kT}; \quad n(0) / n(l) = \exp \frac{F_n(0) - F_n(l) + E_c(l) - E_c(0)}{kT}.$$

Из рис. VI.2.3 следует, что $F_n(0) - F_n(l) = -eV$ (на границе с металлом $F_n(0) = F_M$) и $E_c(l) - E_c(0) = e(V - V_b)$. Таким образом, $F_n(0) - F_n(l) + E_c(l) - E_c(0) = -eV_b$ и

$$n(0) / n(l) = e^{-V_b / \Phi_T}. \quad (\text{VI.2.18})$$

В контакте Шоттки (в отличие от полупроводникового диода) инжекция неосновных носителей из металла в базу отсутствует, поэтому на границе ОПЗ с электронейтральной базой

$$n(l) = n_{B0} = N_B. \quad (\text{VI.2.19})$$

Подставляя (VI.2.18) и (VI.2.19) в (VI.2.17), получаем

$$j_n = eD_n n_{B0} e^{-V_b / \Phi_T} \frac{e^{V / \Phi_T} - 1}{\int_0^l e^{[\varphi(0) - \varphi(x)] / \Phi_T} dx}. \quad (\text{VI.2.20})$$

При $V_b - V \gg \Phi_T$ подынтегральная функция в знаменателе (VI.2.20) быстро уменьшается с ростом координаты x и существенно только ее поведение при $x \ll l$, где

$$\varphi(x) \approx \varphi(0) + E_m x,$$

$$E_m(V) = E_{m0} \sqrt{1 - V / V_b} \quad (\text{VI.2.21})$$

— максимальная напряженность электрического поля в ОПЗ, равновесное значение которой E_{m0} определено в (VI.1.19). Таким образом,

$$\int_0^l e^{[\varphi(0) - \varphi(x)] / \Phi_T} dx \approx \int_0^\infty e^{-E_m x / \Phi_T} dx = \Phi_T / E_m. \quad (\text{VI.2.22})$$

Подстановка (VI.2.22) в (VI.2.20) с учетом соотношения Эйнштейна $D_n = \Phi_T \mu_n$ дает ВАХ контакта Шоттки в виде

$$I(V) = I_0 (1 - V / V_b)^{1/2} (e^{V / \Phi_T} - 1), \quad (\text{VI.2.23})$$

где $I_0 = j_0 S$;

$$j_0 = \sigma_{B0} E_{m0} e^{-V_b / \Phi_T}; \quad (\text{VI.2.24})$$

$\sigma_{B0} = en_{B0} \mu_n$ — удельная электропроводность базы.



VI.2.4. Сравнение результатов ТТЭ и диффузионной теории выпрямления

Сравнение ВАХ (VI.2.16) и (VI.2.23) показывает, что при $V > 0$ диффузионная теория дает несколько менее резкое нарастание тока с напряжением (за счет множителя $(1 - V/V_b)^{1/2}$). При обратном напряжении согласно ТТЭ ток не зависит от напряжения, в то время как диффузионная теория предсказывает увеличение обратного тока с ростом обратного напряжения.

Ширина ОПЗ уменьшается с увеличением степени легирования базы (VI.1.12). Практически условие $l \ll \lambda$ применимости ТТЭ обычно можно считать выполненным при $N_B > 10^{17} \text{ см}^{-3}$, а условие применимости диффузионной теории — при $N_B < 10^{16} \text{ см}^{-3}$. Согласно (VI.1.13) ширина ОПЗ возрастает с ростом обратного напряжения. Поэтому при концентрации примеси в базе $10^{16} \text{ см}^{-3} < N_B < 10^{17} \text{ см}^{-3}$ может оказаться, что прямую ветвь ВАХ лучше описывает ТТЭ, а обратную — диффузионная теория.

Обобщение ТТЭ и диффузионной теории выпрямления выполнено в работе [5].

VI.3. Омические контакты

Как отмечалось в параграфе VI.1.1, большинство металлов с наиболее распространенными полупроводниковыми материалами n -типа образуют барьерные контакты, а с материалами p -типа — омические. Это означает, что вблизи границы с металлом энергетические зоны практически всегда изогнуты вверх независимо от типа проводимости полупроводника. Таким образом, в контакте металл—полупроводник p -типа ОПЗ представляет собой область, обогащенную основными носителями заряда (дырками). Ширина ОПЗ (как и в структуре МДП в режиме обогащения, см. раздел III.1) имеет порядок дебаевской длины экранирования в базе L_{DB} .

Энергетическая диаграмма омического контакта металл—полупроводник p -типа показана на рис. VI.3.1, а. Из рисунка очевидно, что ввиду отсутствия потенциального барьера дырки могут свободно перемещаться через границу металл—полупроводник при любом направлении тока.

В контакте металл—полупроводник n -типа потенциальный барьер существует всегда, однако за счет повышения

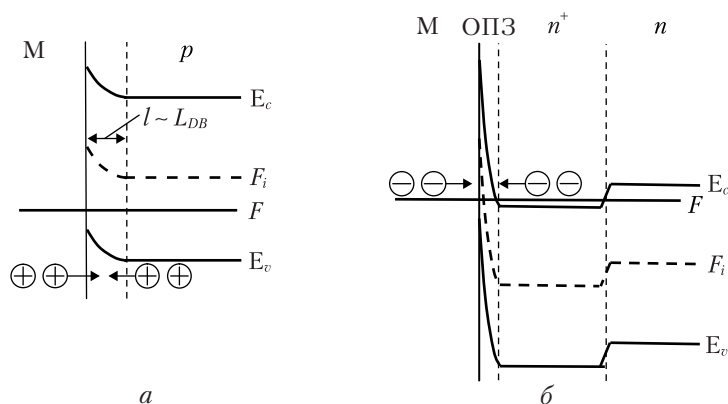


Рис. VI.3.1. Энергетические диаграммы омических контактов металл—полупроводник:

а — к полупроводнику *p*-типа; *б* — к полупроводнику *n*-типа

степени легирования базы его ширина может быть сделана чрезвычайно малой. При $N > 10^{19} \text{ см}^{-3}$ ширина равновесного барьера по основанию (ширина ОПЗ) составляет менее 10 нм, и барьер является туннельно прозрачным (напряжение туннельного пробоя равно нулю, параграф II.6.3). Это позволяет формировать омический контакт к полупроводнику *n*-типа путем сильного легирования поверхности полупроводника (рис. VI.3.1, б).

Таким образом, в МДПТ омические контакты к вырожденным областям стока и истока (см. рис. III.8.2—III.8.4) формируются автоматически. В биполярных *n-p-n* транзисторах (см. рис. IV.8.1) омические контакты к *p*-базе и *n*⁺-эмиттеру также формируются автоматически, а под коллекторным контактом создается специальная контактная *n*⁺-область (см. рис. IV.8.1). В латеральных *p-n-p* транзисторах омические контакты автоматически формируются к *p*-областям эмиттера и коллектора, а в *n*-базе также создается специальная *n*⁺-область (см. рис. IV.8.2).

Следует отметить, что между *n*-областью с концентрацией примеси *N* и контактной *n*⁺-областью существует тонкий переходный слой с повышенным сопротивлением. Сопротивление переходного слоя характеризуется *удельным контактным сопротивлением* [6]

$$\rho_c (\text{Ом} \cdot \text{см}^2) = k_c (T/300)^{1/2} \cdot 1,55 \cdot 10^{10} / N (\text{см}^{-3}), \quad (\text{VI.3.1})$$

где $k_c > 1$ — коэффициент качества омического контакта.

VI.4. Диоды Шоттки

Выпрямляющие свойства барьерных контактов металл—полупроводник используются для создания *диодов Шоттки* (ДШ) — полупроводниковых приборов, сходных по свойствам с диодами на основе p - n переходов.

Структуры ДШ на основе планарной технологии БТ и в дискретном исполнении показаны на рис. VI.4.1, *а*, *б*. Барьерные контакты формируются переходами металл—полупроводник n -типа, а омические — переходами металл—полупроводник n^+ -типа.

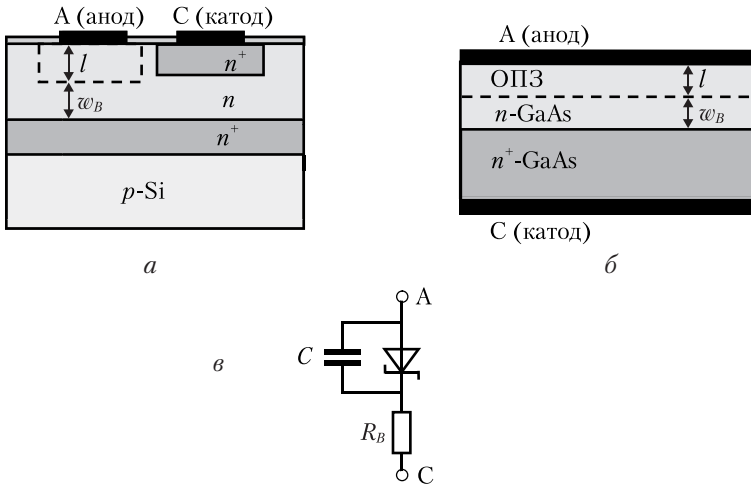


Рис. VI.4.1. Диоды Шоттки:

а — планарная структура на основе технологии БТ;
б — дискретный диод на основе GaAs; *в* — эквивалентная схема

Главное преимущество ДШ состоит в практическом отсутствии диффузионной емкости, связанной с инжекцией неосновных носителей из металла в n -базу. Поэтому ДШ широко применяются в смесительных устройствах СВЧ-диапазона. Эквивалентная схема диода Шоттки представлена на рис. VI.4.1, *в*.

Инерционные свойства ДШ характеризуются постоянной времени

$$\tau_0 = R_{B0}C_0, \quad (\text{VI.4.1})$$

где R_{B0} и C_0 — равновесные значения сопротивления базы и барьерной емкости контакта (при $V = 0$), сопротивление базы определяется соотношением (VI.1.15). Для дискретной конструкции (см. рис. VI.4.1, *б*) сопротивление базы составляет

$$R_{B0} = w_{B0} / eN_B\mu_n S, \quad (\text{VI.4.2})$$

и с учетом (VI.1.15) постоянная времени

$$\tau_0 = \frac{w_{B0}(\epsilon\epsilon_0)^{1/2}}{\mu_n(2eN_BV_b)^{1/2}} \quad (\text{VI.4.3})$$

не зависит от площади контакта S . Повышение быстродействия при сохранении напряжения пробоя достигается уменьшением длины базы до величины, соответствующей полному перекрытию базы областью пространственного заряда при максимальном обратном напряжении. С учетом (VI.1.13) минимальная длина базы составляет

$$w_{B\min} = l_0 \left(\sqrt{1 - V_{\max} / V_b} - 1 \right) \quad (\text{VI.4.4})$$

Соотношение (VI.4.3) показывает, что быстродействие ДШ повышается при повышении концентрации примеси в базе ($\tau_0 \sim N_B^{-1/2}$), которая ограничена требуемым напряжением пробоя. Обычно степень легирования базы не превышает $N_B \approx 5 \cdot 10^{17} \text{ см}^{-3}$, так как при больших значениях существенно снижается подвижность электронов μ_n .

Максимальное быстродействие имеют ДШ на основе GaAs (см. рис. VI.4.1, б), который характеризуется примерно в семь раз более высокой подвижностью электронов, чем кремний. В таких ДШ постоянная времени τ_0 может быть менее 1 пс, что позволяет применять их до частот более 100 ГГц. При прочих равных условиях планарные ДШ имеют в два-три раза меньшее быстродействие из-за большего сопротивления базы.

В кремниевых цифровых ИС на элементной базе ТТЛ (транзисторно-транзисторная логика) ДШ используются для того, чтобы предотвратить накопление заряда избыточных дырок в n -коллекторе БТ в режиме насыщения, которое значительно затягивает переходный процесс выключения БТ (см. разд. IV.4, V.5). За счет незначительного удлинения металл контакта к p -базе частично накрывает область n -коллектора (рис. VI.4.2, а), образуя с ней барьерный контакт. Как отмечено в параграфе VI.2.2, масштабный ток контакта Шоттки I_0 значительно превышает тепловой ток коллекторного перехода I_{SC} . В результате ВАХ контакта Шоттки сдвинута относительно ВАХ коллекторного перехода на величину $\approx 0,15 \text{ В}$, и при заданной величине коллекторного тока ДШ, включенный параллельно коллекторному переходу, «перехватывает» большую часть тока (см. рис. VI.4.2, а).



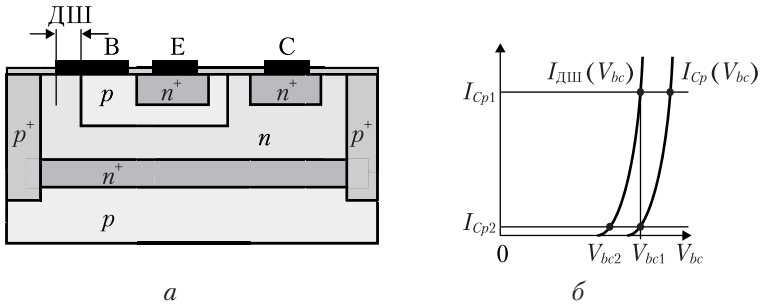


Рис. VI.4.2. Шунтирование диодом Шоттки коллекторного перехода планарного n - p - n БТ для предотвращения инжекции дырок в n -коллектор:

a — структура БТ с ДШ; $б$ — снижение дырочного тока через коллекторный переход за счет ДШ ($I_{cp1} \rightarrow I_{cp2}$)

На ВАХ реальных диодов Шоттки некоторое воздействие оказывают факторы, не учтенные в ТТЭ и диффузионной теории: а) туннелирование электронов сквозь барьер; б) рекомбинация электронно-дырочных пар в ОПЗ; в) инжекция дырок в полупроводник из металла.

Туннелирование электронов сквозь барьер имеет место вблизи его вершины, где ширина барьера мала (процесс 1 на рис. VI.4.3). ВАХ туннельного тока исследована в работах [7, 8]. Показано, что с ростом прямого напряжения туннельный ток растет медленнее, чем по закону e^{V/Φ_T} . Поэтому его роль существенна при малом прямом и особенно при обратном напряжении.

Рекомбинация в ОПЗ существенна, если на границе металл—полупроводник уровень Ферми расположен вблизи уровня E_v (процесс 2 на рис. VI.4.3). В этом случае ток рекомбинации—генерации в ОПЗ имеет тот же механизм, что и в p - n переходе (см. параграфы II.5.2—II.5.4). Прямая ветвь его ВАХ описывается соотношением (II.5.11)

$$I_r = I_{r0} e^{V/m\Phi_T},$$

где $m \approx 2$, а обратная — соотношением (II.5.4) при $k = 2$

$$I_g(V) = I_{g0} \sqrt{1 - V/V_b}.$$

При очень больших токах (и большом прямом напряжении) высота барьера для дырок в металле может снизиться настолько, что возможна инжекция дырок из металла в полупроводник (процесс 3 на рис. VI.4.3). При этом диод Шоттки имеет некоторую диффузионную емкость, пропорциональ-

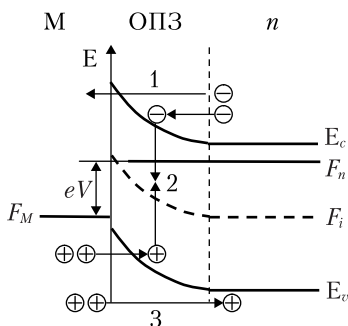


Рис. VI.4.3. Туннелирование электронов сквозь барьер (1), рекомбинация в ОПЗ (2) и инжекция дырок из металла в полупроводник (3)

ную току. Величина ее, впрочем, значительно меньше, чем в полупроводниковом диоде.

Наличие туннельного тока, тока рекомбинации в ОПЗ, а также множителя $1 - (V/V_b)^{1/2}$ в ВАХ (VI.2.23) по диффузионной теории практически учитывается аппроксимацией прямой ВАХ ДШ универсальным соотношением

$$I = I_0 (e^{V/m\Phi_T} - 1),$$

где фактор неидеальности $m > 1$. В высококачественных диодах Шоттки $m = 1,05 - 1,1$.

Основные выводы

1. Большинство металлов с наиболее распространенными полупроводниковыми материалами n -типа образуют барьерные контакты, а с материалами p -типа — омические.

2. Барьерные контакты металл—полупроводник (контакты Шоттки) характеризуются высотой барьера в металле V_{bm} и в полупроводниковой базе V_b . При нулевой плотности поверхностных состояний высота барьера в металле определяется разностью термодинамических работ выхода из металла и из полупроводника и не зависит от степени легирования базы. В реальных контактах Шоттки высота барьера в металле составляет 0,6—0,75 ширины запрещенной зоны полупроводника. Разность $V_{bm} - V_b$ не зависит от свойств поверхности и определяется только свойствами базы.

3. Ширина ОПЗ и барьерная емкость контакта Шоттки определяются так же, как в сильно асимметричном ступенчатом переходе.

4. Высота барьеров V_{bm} и V_b снижается под действием сил зеркального отображения (эффект Шоттки).

5. ВАХ барьерных контактов с малой шириной ОПЗ (меньшей длины свободного пробега электронов) определяется теорией термоэлектронной эмиссии (ТТЭ) и по форме совпадает с ВАХ идеализированного полупроводникового диода. При большой ширине ОПЗ действует диффузионная теория выпрямления, которая дает несколько более слабую зависимость тока от напряжения. ТТЭ справедлива при концентрации примеси в базе $N_B > 10^{17} \text{ см}^{-3}$, диффузионная теория — при $N_B < 10^{16} \text{ см}^{-3}$.

6. Омические контакты к полупроводнику p -типа формируются автоматически, а к полупроводнику n -типа — путем сильного легирования донорами поверхностного слоя.

7. Выпрямляющие свойства барьерных контактов металл—полупроводник используются для создания диодов Шоттки. Главное преимущество диодов Шоттки состоит в практическом отсутствии диффузионной емкости, связанной с инжекцией неосновных носителей из металла в n -базу.

8. На характеристики реальных диодов Шоттки некоторое воздействие оказывают туннелирование электронов сквозь барьер и рекомбинация-генерация носителей заряда в ОПЗ. Прямая ветвь ВАХ описывается аппроксимацией $I \approx I_0 (e^{V/m\varphi_T} - 1)$, где фактор неидеальности $m > 1$.

Литература

1. *Schottky, W.* Halbleitertheorie der Sperrschicht / W. Schottky // Naturwissenschaften. — 1938. — Vol. 26. — P. 843.

2. *Mott, N. F.* Note on the contact between a metal and an insulator or semiconductor / N. F. Mott // Proc. Cambr. Philos. Soc. — 1938. — Vol. 34. — P. 568.

3. *Henisch, H. K.* Rectifying semiconductor contacts / H. K. Henisch. — Clarendon, Oxford, 1957.

4. *Bethe, H. A.* Theory of the boundary layer of crystal rectifiers / H. A. Bethe // MIT Radiat. Lab. Rep. — 1942. — P. 43–12.

5. *Crowell, C. R.* Current transport in metal-semiconductor barriers / C. R. Crowell, S. M. Sze // Solid State Electron. — 1966. — Vol. 9. — P. 1035.

6. *Шур, М.* Современные приборы на основе арсенида галлия / М. Шур; пер. с англ. — М.: Мир, 1991. — 632 с.

7. *Стреттон, Р.* Туннелирование в выпрямителях с барьером Шоттки / Р. Стреттон Туннельные явления в полупроводниках; пер. с англ.; под ред. Э. Бурштейна и С. Лундквиста. — М.: Мир, 1973. — С. 106–124.

8. *Jeong, M. K.* Comparison of raised and Shottky source/drain MOS-FETs using novel tunneling contact model / M. K. Jeong [and other] // IEDM-98. — P. 733–736.



Глава VII

ПОЛЕВЫЕ ТРАНЗИСТОРЫ С УПРАВЛЯЮЩИМ БАРЬЕРНЫМ ПЕРЕХОДОМ

Полевой транзистор с управляющим барьерным переходом (ПТБП) в виде p - n перехода впервые предложен У. Шокли в 1952 г. [1]. В этой же работе Шокли изложена классическая теория ПТБП, которая легла и в основу теории МДПТ (см. главу III). В соответствии с идеей Шокли ПТБП с управляющим p - n переходом экспериментально реализован Г. Дейки и И. Россом [2]. Ими же развита теория ПТБП [3].

Принцип действия ПТБП основан на том, что толщина резистивного активного слоя полупроводника (канала) между двумя омическими контактами (стока и истока) модулируется областью пространственного заряда (ОПЗ). Толщина ОПЗ управляется потенциалом барьерного электрода (затвора).

Как и транзисторы МДП, ПТБП являются *униполярными* приборами. Их принципиальное отличие от МДП-транзисторов состоит в том, что модуляция проводимости канала осуществляется за счет изменения *толщины канала*, а не поверхностной плотности носителей заряда в канале.

По типу барьерного перехода можно выделить две разновидности ПТБП: полевые транзисторы с управляющим p - n переходом и полевые транзисторы с барьером Шоттки.

VII.1. Полевые транзисторы с управляющим p - n переходом

VII.1.1. Устройство и принцип действия

Полевые транзисторы с управляющим p - n переходом (ПТУП или JFET — junction field-effect transistor) являются первыми униполярными транзисторами, производство которых было освоено промышленностью. Важным факто-



ром послужила независимость порогового напряжения от состояния поверхности.

Как и МДПТ, полевые транзисторы с управляющим p - n переходом могут иметь канал n - или p -типа. Структура n -канального транзистора схематически представлена на рис. VII.1.1. Основой прибора является тонкий полупроводниковый кристалл n -типа, имеющий форму прямоугольного параллелепипеда. На торцах кристалла сформированы омические контакты стока D и истока S . Затвор G выполнен в виде p^+ -областей на боковых гранях кристалла. Омические контакты к p^+ -областям затвора объединены. Каналом является электронейтральная n -область исходного кристалла, расположенная между ОПЗ p - n переходов. «Внутренний» транзистор ограничен плоскостями, проходящими через границы p^+ -областей затвора («внутренний сток» d и «внутренний исток» s). Расстояние между этими плоскостями соответствует длине канала L (см. рис. VII.1.1).

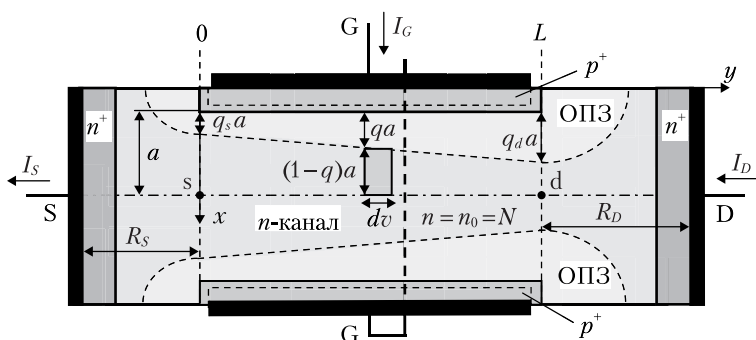


Рис. VII.1.1. Структура n -канального полевого транзистора с управляющим p - n переходом

Принцип действия прибора аналогичен принципу действия МДПТ. Напряжение затвор—исток V_{gs} определяет ширину ОПЗ p - n переходов под электродами затвора и, следовательно, ширину и сопротивление n -канала. Напряжение сток—исток V_{ds} между внутренними электродами стока и истока создает ток в цепи сток—исток, величина которого зависит от сопротивления канала и, следовательно, от напряжения затвор—исток V_{gs} .

Усилительные свойства транзистора лучше всего проявляются в режиме, когда напряжение V_{gs} не вызывает заметного тока I_G в цепи затвора:

$$V_{gs} < V_{gs\max} = k_b V_b, \quad (\text{VII.1.1})$$

где V_b — *барьерный потенциал* (контактная разность потенциалов затвор—канал); $k_b \approx 0,5$. При больших напряжениях затвор—исток часть $I_G R_S$ внешнего напряжения V_{GS} выделяется на сопротивлении R_S пассивной области истока (см. рис. VII.1.1). При этом изменение напряжения V_{GS} незначительно изменяет внутреннее напряжение V_{gs} на открытом p - n переходе затвор—канал с малым дифференциальным сопротивлением, что ведет к резкому снижению крутизны ВАХ.

Еще одно нежелательное последствие затворного тока состоит в инжекции неосновных носителей (дырок) из затвора в канал, что приводит к накоплению избыточного заряда в канале (диффузионная емкость канала) и снижению быстродействия.

VII.1.2. ВАХ идеализированного транзистора

Модель идеализированного ПТУП использует большинство допущений, принятых в модели идеализированного МДПТ (см. параграф III.3.1).

1. Малость тока неосновных носителей в канале.
2. Дрейфовый механизм тока основных носителей в канале.
3. Постоянство (независимость от электрического поля) подвижности основных носителей в канале μ .
4. Отсутствие тока в канале при напряжении затвор—исток ниже порогового.
5. Приближение плавного канала (продольное электрическое поле в канале много меньше поперечного).
6. Независимость потенциала канала $V(y)$ от координаты x (см. рис. VII.1.1).
7. Однородность канала (концентрация примеси в канале $N(x, y) = \text{const}$).
8. Электронейтральность канала ($n = n_0 = N$).
9. Малость тока затвора I_G — условие (VII.1.1).
10. Полное обеднение ОПЗ подвижными носителями заряда.
11. Нулевые сопротивления пассивных областей стока R_D и истока R_S .

Методика анализа ВАХ аналогична использованной для анализа ВАХ МДПТ. Электронный (допущение 1) ток в канале не зависит от координаты y , равен току стока I_D (допу-



щение 9) и определяется законом Ома в дифференциальной форме (допущение 2)

$$dV(y) = I_D dR(y), \quad (\text{VII.1.2})$$

где
$$dR(y) = \frac{dy}{eN\mu_n Za[1-q(y)]} = \frac{R_0 dy}{L[1-q(y)]} \quad (\text{VII.1.3})$$

— сопротивление элементарного участка канала длиной dy ; Z — ширина канала; a — толщина канала при отсутствии ОПЗ,

$$R_0 = L / 2eN\mu_n Za \quad (\text{VII.1.4})$$

— сопротивление канала при отсутствии ОПЗ;

$$q(y) = \frac{1}{a} \sqrt{\frac{2\epsilon\epsilon_0[V_b - V_g + V(y)]}{eN}} = \quad (\text{VII.1.5a})$$

$$= \sqrt{[V_b - V_g + V(y)] / V_p} \quad (\text{VII.1.5б})$$

— часть толщины канала, перекрытая областью пространственного заряда (см. рис. VII.1.1),

$$V_p = eNa^2 / 2\epsilon\epsilon_0 \quad (\text{VII.1.6})$$

— *напряжение перекрытия канала*. Соотношения (VII.1.2) — (VII.1.4) учитывают допущения 3, 6–8 и 11, а соотношение (VII.1.5) — допущения 5 и 10. Соотношения (VII.1.5а, б) справедливы при $q(y) \leq 1$. Если из (VII.1.5а, б) следует, что $q(y) > 1$, следует полагать $q(y) = 1$.

Подстановка (VII.1.3) в (VII.1.2) дает дифференциальное уравнение для функции $q(y)$:

$$\frac{I_D dy}{L} = \frac{1-q(y)}{R_0} dV. \quad (\text{VII.1.7})$$

Как и в МДПТ, проводимость канала, пропорциональная $1-q(y)$, уменьшается по направлению к стоку (см. рис. VII.1.1). На границе с внутренним истоком ($y=0$) $V(0) = V_s$ (см. рис. VII.1.1) и $V_g - V(0) = V_{gs}$. Из соотношения (VII.1.5б) следует, что канал исчезает ($q(0) = 1$) при $V_{gs} \leq V_b - V_p$. Таким образом, величина

$$V_t = V_b - V_p \quad (\text{VII.1.8})$$

имеет смысл *порогового напряжения*.

Интегрируя уравнение (VII.1.7) по всей длине канала (от $y=0$ до $y=L$), получаем

$$I_D = I_{D0} \left[(q_d^2 - q_s^2) - \frac{2}{3}(q_d^3 - q_s^3) \right], \quad (\text{VII.1.9})$$

$$\text{где } I_{D0} = V_p / R_0 = \mu_n a^3 (eN)^2 (Z / L) / \epsilon \epsilon_0, \quad (\text{VII.1.10})$$

и согласно (VII.1.56) и (VII.1.8)

$$q_s = q(0) = \sqrt{(V_b - V_{gs}) / V_p} = \sqrt{(V_b - V_{gs}) / (V_b - V_t)}; \quad (\text{VII.1.11a})$$

$$q_d = q(L) = \sqrt{(V_b - V_{gd}) / V_p} = \sqrt{(V_b - V_{gs} + V_{ds}) / (V_b - V_t)}. \quad (\text{VII.1.11б})$$

Уравнения (VII.1.9) и (VII.1.11а, б) в параметрической форме определяют выходные ВАХ транзистора. По определению потенциал стока выше потенциала истока ($V_{ds} \geq 0$), поэтому $q_s \leq q(y) \leq q_d$. Как отмечено выше, соотношение (VII.1.5а, б), использованное при выводе ВАХ, справедливо при $q(y) \leq 1$, следовательно, уравнения (VII.1.9) и (VII.1.11а, б) описывают ВАХ при $q_d \leq 1$ и согласно (VII.1.11б) соответствуют *крутой области* ВАХ

$$V_{gd} \geq V_t, \quad (\text{VII.1.12a})$$

$$\text{или } V_{ds} \leq V_{dsS} = V_{gs} - V_t, \quad (\text{VII.1.12б})$$

где V_{dsS} — *напряжение насыщения*.

При $V_{ds} = V_{dsS}$ (как и в МДПТ) канал перекрывается на границе со стоком ($q_d = 1$). Согласно (VII.1.9) ток стока достигает значения

$$I_{DS} = I_{D0} \left[(1 - q_s^2) - 2(1 - q_s^3) / 3 \right] \quad (\text{VII.1.13})$$

(ток стока насыщения). При $V_{ds} > V_{dsS}$ транзистор находится в *пологой области* ВАХ, где ток стока $I_D = I_{DS}$ не зависит от напряжения сток—исток.

При $V_{gs} < V_t$ канал перекрыт по всей длине $q(y) = 1$ и транзистор работает в *режиме отсечки*, где согласно допущению 4 $I_D = 0$.

Соотношения (VII.1.8) и (VII.1.6) показывают, что пороговое напряжение ПТУП определяется исходной толщиной канала a и степенью его легирования N . При $V_p = eNa^2 / 2\epsilon\epsilon_0 > V_b$ пороговое напряжение $V_t < 0$ (транзистор со встроенным каналом, или *нормально открытый*); при $V_p < V_b$ пороговое напряжение $V_t > 0$ (транзистор с индуцированным каналом, или *нормально закрытый*). Для типичного значения $N \approx 10^{17} \text{ см}^{-3}$ ПТУП на основе Si ($V_b \approx 0,7 \text{ В}$) имеет нулевое пороговое напряжение при исходной толщине канала $a \approx 0,1 \text{ мкм}$.



VII.1.3. Эквивалентная схема и особенности ВАХ

Эквивалентная схема идеализированного ПТУП изображена на рис. VII.1.2. ВАХ канала представлена генератором тока $I_D(V_{gs}, V_{ds})$. Диоды D_{gs} и D_{gd} моделируют *p-n* переход затвор—канал. Емкости затвор—исток и затвор—сток C_{gs}, C_{gd} включают барьерные емкости *p-n* перехода затвор—канал, а также диффузионные емкости канала, пропорциональные токам I_{gs} и I_{gd} через диоды D_{gs} и D_{gd} . Сопротивления R_D и R_S соответствуют сопротивлениям пассивных областей стока и истока (см. рис. VII.1.1), а сопротивление R_G — сопротивлению затвора.

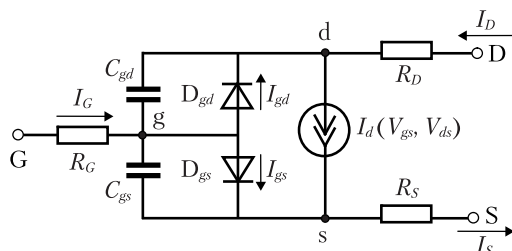


Рис. VII.1.2. Эквивалентная схема полевого транзистора с управляющим *p-n* переходом

На рис. VII.1.3, *a* представлены выходные ВАХ транзистора с пороговым напряжением $V_t = -1$ В, а на рис. VII.1.3, *б* — входная ВАХ $I_G(V_{GS})$ при напряжении сток—исток $V_{DS} = 2$ В. По сравнению с ВАХ МДПТ они имеют следующие особенности.

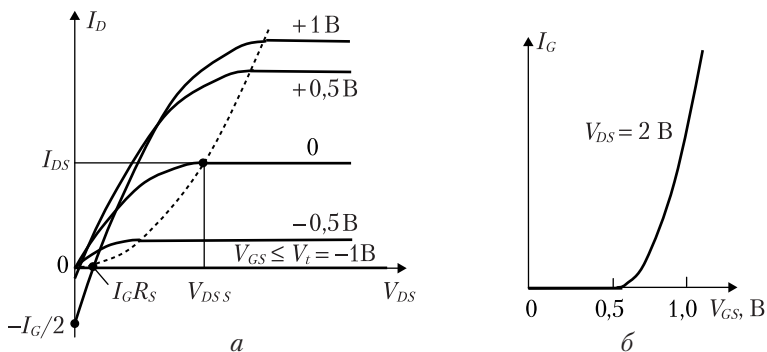


Рис. VII.1.3. ВАХ полевого транзистора с управляющим *p-n* переходом ($V_t = -1$ В):
a — выходные; *б* — входная.

Пунктир на рис. *a* разделяет крутую и пологую области ВАХ

1. Затвор ПТУП не изолирован от канала диэлектриком, поэтому при $V_{gs} > V_{gs\max} = k_b V_b \approx 0,6 \text{ В}$ (VII.1.1) в цепи затвора может протекать ток I_G (см. рис. VII.1.3, б), сравнимый с током стока. Отметим, что при достаточно высоком напряжении V_{ds} диод D_{gd} оказывается закрытым и весь ток затвора протекает в цепи затвор—исток, в результате чего ток истока больше тока стока на величину I_G .

2. При $V_{gs} > V_{gs\max}$ выходные ВАХ на рис. VII.1.3, а не проходят через начало координат. Действительно, при $V_{gs} > V_{gs\max}$ и $V_{ds} = 0$ затворный ток I_G поровну распределяется между электродами стока и истока. Таким образом, в соответствии с принятыми за положительные направления токов (см. рис. VII.1.2) $I_S = I_G / 2$ и $I_D = -I_G / 2$. Поэтому выходная ВАХ, соответствующая внешнему напряжению $V_{GS} = +1 \text{ В}$, пересекает ось тока в точке $I_D = -I_G / 2$ (см. рис. VII.1.3, а). При «оборванном» стоке ($I_D = 0$) большая часть затворного тока протекает через диод D_{gs} и сопротивление истока R_S . Поэтому выходная ВАХ пересекает ось напряжения в точке $V_{DS} \approx I_G R_S$ (см. рис. VII.1.3, а).

3. При $V_{gs} > V_{gs\max}$, когда диод D_{gs} открыт, большая часть приращения напряжения затвор—исток dV_{GS} выделяется не на малом дифференциальном сопротивлении диода D_{gs} , а на сопротивлении R_S (см. рис. VII.1.1): $dV_{gs} \ll dV_{GS}$. Поэтому в режиме затворного тока крутизна ВАХ $g = dI_D / dV_{GS}$ резко снижается. Это показывает рис. VII.1.3, а: при изменении напряжения V_{GS} от 0,5 до 1 В приращение тока стока значительно меньше, чем при изменении V_{GS} от 0 до 0,5 В.

VII.1.4. Сравнение полевого транзистора с управляющим *p-n* переходом и МДПТ

Принципиальное отличие ПТУП от МДПТ, как уже упоминалось выше, состоит в том, что модуляция проводимости канала в нем осуществляется за счет изменения *толщины электронейтрального канала*, а не поверхностной плотности носителей заряда в канале. При изменении напряжения затвор—исток приращение заряда на металлическом контакте затвора компенсируется не изменением заряда в канале (как в МДПТ), а *изменением заряда в ОПЗ *p-n* перехода* затвор-канал (т.е. в барьерной емкости *p-n* перехода).

Конструктивные различия между дискретным ПТУП (см. рис. VII.1.1) и планарным МДПТ (см. рис. III.2.1) довольно значительны. ПТУП представляет собой трехэлектродный



прибор, в то время как в МДПТ четвертый электрод подложки играет существенную роль (потенциал подложки влияет на пороговое напряжение и деформирует ВАХ). Однако эти различия нивелируются в планарной конструкции ПТУП (рис. VII.1.4), где наличие подложки обязательно.

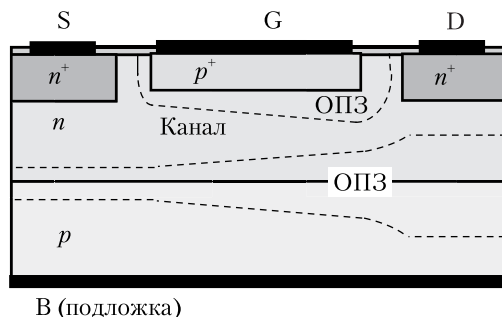


Рис. VII.1.4. Планарная конструкция полевого транзистора с управляющим p - n переходом

В ПТУП роль подзатворного диэлектрика выполняет ОПЗ p - n перехода затвор—канал. В первую очередь это различие проявляется в существовании затворного тока и диффузионных составляющих емкостей затвор—исток и затвор—сток, о чем упоминалось выше. Хотя математическая форма ВАХ ПТУП и МДПТ различна, их статические характеристики при отсутствии затворного тока весьма схожи (см. рис. VII.1.3 и рис. III.2.5), хотя ВАХ идеализированного МДПТ описывается двумя параметрами (V_t и β), а ВАХ ПТУП — тремя (I_{D0} , V_t и V_p).

Математическая запись ВАХ ПТУП (VII.1.9), (VII.1.11а, б) сложнее, чем ВАХ МДПТ — уравнения (III.3.17). Причина состоит в том, что в МДПТ эффект поля проявляется через линейную емкость диэлектрика, а в ПТУП — через нелинейную барьерную емкость p - n перехода затвор—канал.

МДП-транзисторы практически вытеснили кремниевые ПТУП из интегральной микроэлектроники. Решающими факторами явились наличие в ПТУП затворного тока, существенно ограничивающее их схемотехнические возможности, а также меньшая компактность конструкции (затвор в виде p - n перехода занимает большую площадь на кристалле, а его длина не может быть сделана столь же малой, как в МДПТ). Преимущество ПТУП состоит в том, что канал находится

не на поверхности кристалла, а в его объеме, и свойства поверхности практически не влияют на характеристики прибора. В частности, ПТУП имеют более высокую радиационную стойкость, что для некоторых областей применения имеет решающее значение.

Основные выводы

1. Принцип действия полевых транзисторов с управляющим *p-n* переходом аналогичен принципу действия МДПТ. Основное различие состоит в том, что модуляция проводимости канала осуществляется за счет изменения толщины электронейтрального канала, а не поверхностной плотности носителей заряда в канале.

2. ВАХ ПТУП схожи с ВАХ МДПТ, однако возможности повышения управляющего напряжения затвор—исток ограничены затворным током.

3. В режиме затворных токов крутизна ВАХ резко снижается, а в емкостях затвор—исток и затвор—сток появляется диффузионная составляющая.

4. По схемотехническим возможностям и возможностям достижения предельно малых размеров ПТУП уступают МДПТ. Их преимуществом является изоляция канала от поверхности полупроводникового кристалла.

VII.2. Полевые транзисторы с затвором Шоттки на основе GaAs

Преимущества кремния как базового материала микроэлектроники столь велики, что на его основе производится более 95% микроэлектронных изделий. Тем не менее возможности кремния ограничены, и имеются полупроводниковые материалы, превосходящие кремний по некоторым свойствам. К таким материалам относятся полупроводники группы A^3B^5 , из которых наиболее хорошо изучен арсенид галлия (GaAs). Хотя в целом GaAs уступает кремнию как материал микроэлектроники, его свойства позволяют создавать для некоторых областей применения полупроводниковые приборы, параметры которых не достижимы на кремнии. К таким приборам, в частности, относится полевой транзистор с затвором Шоттки (ПТШ).



VII.2.1. Особенности GaAs как материала микроэлектроники

Основные физические свойства GaAs в сравнении с другими полупроводниковыми материалами микроэлектроники приведены в приложениях 3, 4. Более подробные данные содержатся в работе [4].

Важнейшим свойством GaAs является *высокая подвижность электронов* в слабых полях (в семь раз выше, чем в кремнии), обусловленная их малой эффективной массой ($0,067 m_0$). В чистом материале основным механизмом рассеяния является рассеяние на полярных оптических фононах. При охлаждении до 77 К подвижность ограничивается рассеянием на фоновых примесях и достигает $(80-250) \cdot 10^3 \text{ см}^2/\text{В}\cdot\text{с}$. Благодаря *малой энергии активации доноров* ($< 0,005 \text{ эВ}$) они (в отличие от Si) остаются практически полностью ионизованными при температуре ниже 77 К. На рис. VII.2.1, а представлены зависимости подвижности электронов от концентрации примесных ионов для температур 300 К и 77 К.

Главные энергетические минимумы электронов и дырок в GaAs находятся в центре зоны Бриллюэна, поэтому электронные переходы между зоной проводимости и валентной зоной обычно являются прямыми. Вследствие этого основным механизмом рекомбинации является межзонная излучательная рекомбинация, что позволяет создавать на основе GaAs высококачественные оптоэлектронные приборы (в том числе лазеры). Наличие в зоне проводимости побочных энергетических минимумов с высокой плотностью состояний приводит к аномальной зависимости дрейфовой скорости электронов v_n от напряженности электрического поля (рис. VII.2.1, б). При разогреве электрическим полем наиболее энергичные электроны переходят в побочные минимумы с большей эффективной массой, в результате чего зависимость $v_n(E)$ имеет участок с *отрицательной дифференциальной подвижностью* dv_n/dE .

Важным преимуществом GaAs является *значительная ширина запрещенной зоны* (1,43 эВ при 300 К). Собственная концентрация носителей заряда весьма мала ($\approx 2 \cdot 10^6 \text{ см}^{-3}$ при $T = 300 \text{ К}$), что позволяет получить материал с высоким удельным сопротивлением (*полуизолирующий i^{*}-GaAs*). Подложки GaAs с удельным сопротивлением $\rho = 10^7-10^9 \text{ Ом}\cdot\text{см}$ обладают прекрасными диэлектрическими свойствами вплоть до частот в несколько сот гигагерц.

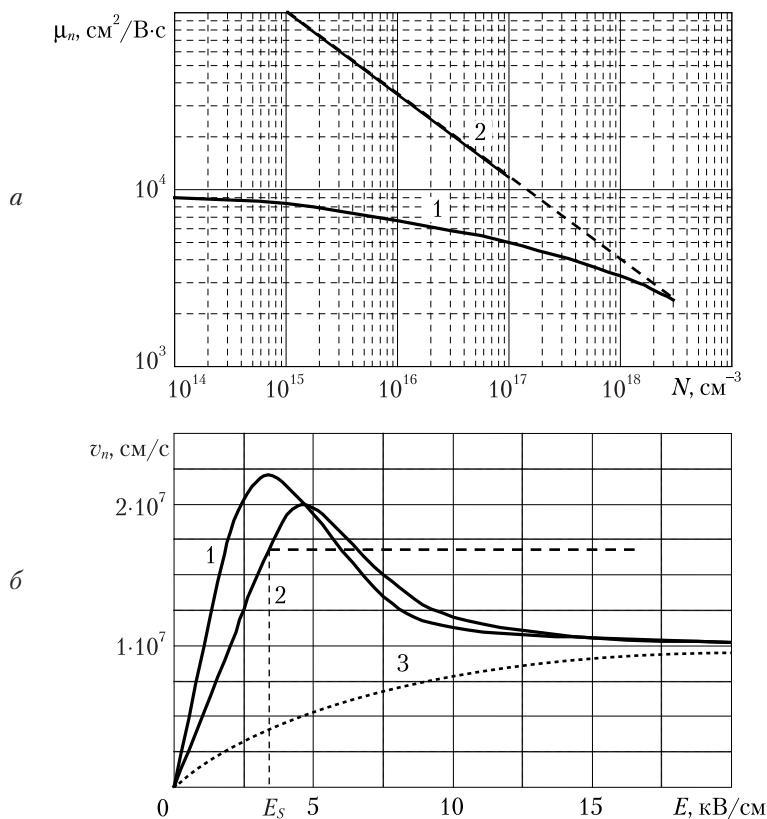


Рис. VII.2.1. Зависимости подвижности электронов в GaAs от концентрации примесных ионов (а) и дрейфовой скорости электронов от напряженности электрического поля при $T = 300 \text{ K}$ (б)
 а: 1 — $T = 300 \text{ K}$, 2 — $T = 77 \text{ K}$; б: 1 — GaAs, $N \rightarrow 0$;
 2 — GaAs, $N = 10^{17} \text{ cm}^{-3}$; 3 — Si, $N \rightarrow 0$

Высокая подвижность электронов позволяет создавать на основе GaAs электронные приборы чрезвычайно высокого быстродействия, причем диэлектрическая подложка обеспечивает малые паразитные емкости в ИС. Значительная ширина запрещенной зоны позволяет расширить *температурный диапазон* работы ИС. Прямой характер межзонных переходов приводит к резкому снижению времени жизни избыточных носителей заряда. В сочетании с большой шириной запрещенной зоны это свойство GaAs повышает *радиационную стойкость* ИС на его основе.

Весьма полезным свойством GaAs является возможность превращения легированного донорами или акцепторами материала в полупроводник путем бомбардировки его легкими ионами (водорода, бора или кислорода). Возникающие при этом радиационные дефекты создают в запрещенной зоне глубокие разрешенные уровни, поглощающие свободные носители заряда. В результате создается возможность получения высококачественной взаимной изоляции элементов ИМС без нарушения планарности структуры.

Недостатками GaAs по сравнению с кремнием являются:

- сравнительно высокая плотность дефектов;
- малая механическая прочность (хрупкость);
- меньшая теплопроводность (в 2,7 раза меньше, чем у Si);
- высокая стоимость (подложки GaAs примерно в 30 раз дороже кремниевых);
- низкая подвижность дырок (менее $400 \text{ см}^2/\text{В} \cdot \text{с}$ при 300 К);
- отсутствие высококачественного собственного оксида, подобного по свойствам SiO_2 .

Поэтому применение GaAs как исходного материала ИС целесообразно в тех случаях, когда необходимые свойства изделий не могут быть получены при использовании кремния.

VII.2.2. Устройство и особенности ПТШ на GaAs

Несмотря на многочисленные исследования, попытки создания на основе GaAs высококачественных МДП-транзисторов не достигли результата. Основные успехи связаны с созданием и промышленным освоением (в том числе в интегральной микроэлектронике) полевых транзисторов с затвором Шоттки (ПТШ). В отличие от ПТУП в этих приборах затвор выполнен в виде барьерного контакта типа Шоттки; в остальном принципы их действия идентичны. Использование в качестве затвора барьерного контакта Шоттки позволяет исключить инжекцию дырок в n -канал и связанную с ней диффузионную емкость. Таким образом, ПТШ сохраняют высокое быстродействие даже в режиме затворных токов.

Типовая структура интегрального ПТШ представлена на рис. VII.2.2. На полупроводящей подложке 1 (нелегированный i^* -GaAs) методами эпитаксии наращивается барьерный n -слой GaAs 2 ($N \approx 2 \cdot 10^{17} \text{ см}^{-3}$, $a_0 \approx 100 \text{ нм}$), а поверх него — контактный n^+ -слой 4 ($N \approx 10^{18} \text{ см}^{-3}$, $a^+ \approx 100 \text{ нм}$). Для формирования омических контактов стока и истока 5 на контактный n^+ -слой 4 наносится трехслойная композиция GeAuNi. При

специальной термообработке сплавления омических контактов образуется низкоомная эвтектика с большим содержанием Ge, который является в GaAs донорной примесью. Взаимная изоляция транзисторов осуществляется селективной бомбардировкой поверхности ионами водорода или бора (см. параграф VII.2.1), в результате чего формируются изолирующие области 6. Для создания барьерного контакта затвора используются композиции, включающие тугоплавкие металлы (W, Ti, V и др.) или их силициды. Перед нанесением барьерного контакта затвора осуществляется селективное травление поверхности барьерного n -слоя 2 до необходимой толщины a , обеспечивающей заданное пороговое напряжение. При необходимости могут быть сформированы два типа ПТШ с различными пороговыми напряжениями, например нормально открытые ($V_t < 0$) и нормально закрытые ($V_t > 0$).

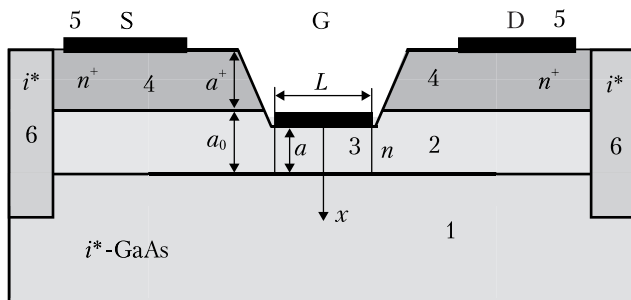


Рис. VII.2.2. Типовая структура интегрального ПТШ на основе GaAs:
 1 — полуизолирующая подложка ($\rho \approx 10^8 \text{ Ом} \cdot \text{см}$); 2 — барьерный n -слой ($N \approx 2 \cdot 10^{17} \text{ см}^{-3}$, $a_0 \approx 100 \text{ нм}$); 3 — n -канал ($a \approx 70 \text{ нм}$);
 4 — контактный n^+ -слой ($N \approx 10^{18} \text{ см}^{-3}$, $a^+ \approx 100 \text{ нм}$);
 5 — омические контакты стока и истока (GeAuNi);
 6 — изолирующие области

Принципиальная технологическая простота и использование методов самосовмещения позволяют реализовать ПТШ сверхвысокого быстродействия с длиной канала $L = 0,1\text{--}0,2 \text{ мкм}$.

Особенностью ПТШ является наличие полуизолирующей подложки, которая при высоком качестве исходного материала практически не влияет на величину порогового напряжения и вид ВАХ.

Описанная конструкция ПТШ имеет однородно легированный канал, примесный профиль которого в сечении канала изображен на рис. VII.2.3, а. Разновидностью ПТШ

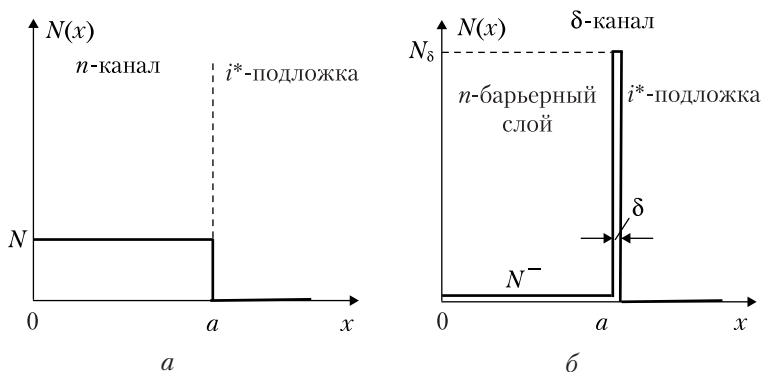


Рис. VII.2.3. Примесный профиль ПТШ с однородным (а) и δ -легированным (б) каналом

являются ПТШ с δ -легированным каналом (рис. VII.2.3, б). В этих транзисторах барьерный n -слой легируется слабее ($N \approx 5 \cdot 10^{16} \text{ см}^{-3}$), а необходимая для обеспечения заданного порогового напряжения примесная доза сосредоточена в тонком ($\approx 15 \text{ нм}$) n^+ -слое (δ -слое). При прочих равных условиях ПТШ с δ -легированным каналом обеспечивают несколько большее значение максимальной крутизны ВАХ и предельной частоты [5].

VII.2.3. Особенности характеристик и применения ПТШ

Для ПТШ с длинным однородно легированным каналом справедлива теория полевых транзисторов с управляющим p - n переходом, изложенная в параграфе VII.1.2. В частности, пороговое напряжение определяется соотношениями (VII.1.8) и (VII.1.6)

$$V_t = V_b - V_p; \quad V_p = eNa^2 / 2\epsilon\epsilon_0,$$

где V_b — высота потенциального барьера в n -канале. Благодаря широкой запрещенной зоне GaAs барьерный потенциал V_b составляет от 0,8 до 0,9 В, что обеспечивает достаточно широкий рабочий диапазон входного сигнала для нормально закрытых ПТШ. Как и для ПТУП, ВАХ в крутой области определяются соотношениями (VII.1.9)–(VII.1.11), а эквивалентная схема соответствует рисунку VII.1.2 с той разницей, что емкости C_{gs} и C_{gd} не содержат диффузионных составляющих.

Для ПТШ с δ -легированным каналом напряжение перекрытия канала вместо (VII.1.6) определяется соотношением

$$V_p = eD_\delta a / \epsilon \epsilon_0, \quad (\text{VII.2.1})$$

где $D_\delta = N_\delta \delta$ — поверхностная доза примеси в δ -канале. В δ -легированных ПТШ ОПЗ под затвором полностью перекрывает барьерный n^- -слой, выполняющий роль подзатворного диэлектрика. Поэтому удельная барьерная емкость затвор—канал

$$C_G = \epsilon \epsilon_0 / a \quad (\text{VII.2.2})$$

не зависит от напряжения, и ВАХ ПТШ с длинным δ -каналом соответствуют теории идеализированного МДПТ (см. параграф III.3.2) с той разницей, что при $V_{gs} > 0$ возможно протекание затворного тока через диоды D_{gs} и D_{gd} (см. рис. VII.1.1).

Ввиду высокой подвижности электронов в GaAs отклонение зависимости $v_n(E)$ от линейной происходит при значительно меньших электрических полях, чем в кремнии (см. рис. VII.2.1). Следует отметить, что сложный немонотонный характер зависимости $v_n(E)$ (см. рис. VII.2.1) проявляется только в образцах, длина которых значительно превышает длину свободного пробега электронов λ , которая в GaAs составляет $\approx 0,5$ мкм и возрастает с ростом поля. В ПТШ с длиной канала менее 1 мкм перенос электронов через канал в большой степени происходит баллистически (с малым числом столкновений). При этом зависимость $v_n(E)$ теряет локальный характер и приближается к монотонной [4]. Для оценочных расчетов, как и для Si, может использоваться кусочно-линейная аппроксимация (I.4.9)

$$v(E) = \begin{cases} \mu E, & E \leq E_s, \\ v_s = \mu E_s, & E > E_s, \end{cases} \quad (\text{VII.2.3})$$

где $E_s \approx 4$ кВ/см (см. рис. VII.2.1, б).

При длине канала $L = 0,5$ мкм пороговое значение поля E_s достигается при $V_{ds} = E_s L = 0,2$ В даже при равномерном распределении продольного поля в канале. Таким образом, *GaAs ПТШ практически всегда работают в режиме ограничения дрейфовой скорости электронов в канале.* Как и в МДПТ, следствиями этого являются эффекты ограничения тока стока насыщения и модуляция эффективной длины канала, которая приводит к конечному значению выходной проводимости в пологой области ВАХ (см. параграфы III.5.5—III.5.7).



Для ПТШ с δ -легированным каналом напряжение насыщения V_{ds} , ток стока насыщения I_{DS} , максимальные значения крутизны $g_{\max}$ и предельной частоты $\omega_{T\max}$, а также характерные значения выходной проводимости и значение эффективной длины канала определяются соотношениями (III.5.8a), (III.5.12), (III.5.14), (III.5.19), (III.5.23), полученными для МДПТ.

Для ПТШ с однородным каналом ток стока насыщения определяется приближенным соотношением [6]

$$I_{DS} \approx I_{D0}(1-q_s) \left[\frac{3}{(1-q_s)(1+2q_s)} + \frac{V_p/V_L}{1+2V_L/3V_p(1-q_s)} \right]^{-1}, \quad (\text{VII.2.4})$$

где $V_L = E_S L$, а значение $q_s(V_{gs})$ определено в (VII.1.11a).

Следует отметить, что в GaAs ПТШ эффект возрастания выходной проводимости и снижения коэффициента усиления по напряжению K_V при уменьшении длины канала проявляется слабее, чем МДПТ. Помимо модуляции длины канала этот эффект связан с инжекцией электронов из канала в подложку. В значительной степени его можно нейтрализовать введением между каналом и подложкой слоя широкозонного полупроводникового соединения AlGaAs [5], который является для электронов канала потенциальным барьером.

В заключение отметим, что при равной длине канала GaAs ПТШ существенно превосходят кремниевые МДПТ по удельной крутизне ВАХ (200—300 мСм/мм), предельной частоте (выше 100 ГГц при $L = 0,1$ мкм), шумовым свойствам и температурному диапазону (от 77 до ≈ 450 К). Основными областями их применения являются малошумящие СВЧ-усилители, генераторы СВЧ-мощности, смесители и другие устройства аналоговой СВЧ-техники. На основе ПТШ производятся и цифровые СБИС гигагерцового диапазона, по степени интеграции мало уступающие кремниевым.

Основные выводы

1. Основными преимуществами GaAs как материала микроэлектроники являются высокая подвижность электронов, малая энергия активации доноров, большая ширина запрещенной зоны и возможность превращения легированного полупроводника в полупроводник. Эти свойства способствуют повышению быстродействия, расширению рабочего температурного диапазона и повышению



стойкости к внешним воздействиям полупроводниковых приборов на основе GaAs.

2. К недостаткам GaAs по сравнению с традиционным кремнием относятся повышенная плотность дефектов, меньшие механическая прочность и теплопроводность, меньшая подвижность дырок, отсутствие высококачественного собственного окисла, высокая стоимость исходного материала.

3. Базовым активным прибором GaAs ИС является полевой транзистор с затвором Шоттки (ПТШ). ПТШ с однородно легированным каналом близки по своим свойствам к полевым транзисторам с управляющим p - n переходом, однако существенно превосходят их по быстродействию и компактности структуры в ИС. ПТШ с δ -легированным каналом близки по свойствам к МДПТ, но существенно превосходят кремниевые МДПТ по удельной крутизне ВАХ, предельной частоте, шумовым свойствам и температурному диапазону.

4. Применение GaAs как исходного материала ИС целесообразно в тех случаях, когда необходимые свойства изделий не могут быть получены при использовании кремния.

Литература

1. *Shockley, W.* A unipolar field-effect transistor / W. Shockley // Proc. IRE. — 1952. — Vol. 40. — P. 1365.
2. *Dacey, G. C.* Unipolar field-effect transistor / G. C. Dacey, I. M. Ross // Proc. IRE. — 1953. — Vol. 41. — P. 970.
3. *Dacey, G. C.* The field-effect transistor / G. C. Dacey, I. M. Ross // Proc Bell Syst. Techn. J. — 1955. — Vol. 34. — P. 1149.
4. *Шур, М.* Современные приборы на основе арсенида галлия / М. Шур ; пер. с англ. — М. : Мир, 1991. — 632 с.
5. *Горбацевич, А. А.* Полупроводниковые гетероструктуры и приборы на их основе / А. А. Горбацевич [и др.] // Нанотехнологии в электронике / под ред. Ю. А. Чаплыгина. — М. : Техносфера, 2005. — С. 172—242.
6. *Старосельский, В. И.* Статические характеристики полевого транзистора с затвором Шоттки на основе арсенида галлия / В. И. Старосельский // Микроэлектроника. — 1982. — Т.11. — № 3. — С. 208—212.



Глава VIII

ГЕТЕРОПЕРЕХОДНЫЕ ТРАНЗИСТОРЫ

Новые возможности совершенствования полупроводниковых приборов открывает использование гетеропереходов — переходов между полупроводниками с различной зонной структурой. Эти возможности связаны с различием в ширине запрещенной зоны элементов гетероперехода, изменением формы потенциального барьера, а также с механически напряженным состоянием какого-либо элемента.

Полупроводниковый прибор на гетероструктурах впервые был предложен У. Шокли в 1951 г. [1], а теория гетероперехода в том же году была развита в России А. И. Губановым [2]. Первая работа по биполярным транзисторам с широкозонным эмиттером была опубликована Х. Кремером в 1957 г. [3]. В 1960 г. Р. Андерсон предсказал возможность существования инверсионного слоя в гетеропереходе [4] и предложил простую модель гетероперехода [5]. Альтернативные модели развиты М. Адамсом и А. Нуссбаумом [6], а также фон В. Русом [7] и Х. Кремером [8]. В 1969 г. Л. Есаки и Р. Тсу указали на возможность повышения подвижности электронов у границы гетероперехода [9]; этот эффект подтвержден в 1978 г. экспериментально [10]. В 1980 г. Т. Мимура и другие специалисты создали первый гетероструктурный полевой транзистор с повышенной подвижностью электронов в канале [11].

VIII.1. Гетеропереходы

Возможность создания гетероперехода определяется существованием двух полупроводников с одинаковым типом кристаллической решетки, близкими значениями периода решетки и технологической возможностью их совместного изготовления. Параметры некоторых полупроводников, пригодных для создания гетеропереходов, приведены в таблице VIII.1.1.



Таблица VIII.1.1

**Параметры некоторых полупроводников,
пригодных для создания гетеропереходов**

Полупроводник	A ³ B ⁵						A ⁴	
	GaAs	AlAs	InP	InAs	InSb	GaSb	Ge	Si
Тип решетки	Цинковая обманка						Алмаз	
Период решетки a_0 , нм	5,65	5,66	5,87	6,06	6,48	6,1	5,65	5,43
Электронное сродство Ψ , эВ	4,07	2,62	5,34	4,9	4,59	4,1	4,13	4,01
Ширина запрещенной зоны E_g , эВ	1,43	2,16	1,35	0,36	0,18	0,73	0,66	1,12
Положение экстремума E_c в зоне Бриллюэна	Г	Х	Г	Г	Г	Г	L	Δ
Эффективные массы	m_n/m_0	0,07	0,78	0,08	0,023	0,01	0,04	0,55
	m_p/m_0	0,57	0,86	0,74	0,41	0,41	0,44	0,2

VIII.1.1. Особенности гетеропереходов

Основными параметрами гетероперехода являются положения границ запрещенной зоны, а именно — разрывы зоны проводимости ΔE_c и валентной зоны ΔE_v . В зависимости от соотношения знаков ΔE_c и ΔE_v различают гетеропереходы I-го и II-го рода. На гетеропереходе I-го рода параметры ΔE_c и ΔE_v имеют разные знаки и края зон смещаются в противоположных направлениях. При этом меньшая запрещенная зона на оси энергии находится целиком внутри большей запрещенной зоны. На гетеропереходе II-го рода параметры ΔE_c и ΔE_v одного знака, и края зон смещаются в одном направлении. При этом запрещенные и разрешенные зоны материалов, образующих гетеропереход, перекрываются.

Первые модели гетеропереходов предполагали, что механизм образования скачков зон аналогичен механизму образования контактной разности потенциалов на p - n переходе: уровень энергии электрона в вакууме на гетеропереходе непрерывен, а разрыв зоны проводимости определяется разностью электронных сродств полупроводников (модель Андерсона [5]):

$$\Delta E_c = \Delta \Psi. \quad (\text{VIII.1.1})$$

Однако эксперименты показали, что этот подход справедлив далеко не всегда. Физическая причина состоит в том, что



взаимодействие находящихся вблизи гетероперехода атомов отличается от их взаимодействия в объеме. При сохранении электронейтральности гетероперехода в целом относительное перераспределение электронной плотности в микроскопической области вблизи гетероперехода характеризуется дипольным моментом, который приводит к разрыву уровня энергии электрона в вакууме. Разрыв зон на гетеропереходе зависит от деталей формирования межатомных связей, причем величина ΔE_c может быть как больше, так и меньше величины $\Delta \Psi$. Существенную роль играет также деформационный потенциал в напряженных гетероструктурах.

Данные табл. VIII.1.1 показывают, что прекрасными парами для гетеропереходов I-го рода являются полупроводниковые соединения A^3B^5 , например $\text{GaAs}/\text{Al}_x\text{Ga}_{1-x}\text{As}$, где x — мольная доля Al. Различие периодов решетки GaAs и AlAs составляет всего 0,12%, однако в GaAs энергетический минимум зоны проводимости находится в центре зоны Бриллюэна, а в AlAs — на краю. Поэтому максимальная величина разрыва зоны проводимости ΔE_c достигается при $x = 0,3$ и составляет около 0,3 эВ.

Ширина запрещенной зоны в AlAs на $\Delta E_{g0} = 0,736$ эВ больше, чем в GaAs, причем в соединении $\text{Al}_x\text{Ga}_{1-x}\text{As}$ величина разрыва в ширине запрещенной зоны примерно пропорциональна мольной доле Al:

$$\Delta E_g \sim x \Delta E_{g0}. \quad (\text{VIII.1.2})$$

Ввиду чрезвычайно малой эффективной массы электронов в индии (и, следовательно, высокой подвижности) весьма перспективными для гетеропереходных транзисторов являются тройные полупроводниковые соединения InGaAs, InGaP, InAlAs и InP. Введение индия в GaAs существенно повышает подвижность электронов. По данным работы [12] в соединении $\text{In}_x\text{Ga}_{1-x}\text{As}$ подвижность электронов возрастает с увеличением мольной доли In x примерно по закону $\mu_n \approx 5000(1+2x)$ см²/В·с. Постоянные решетки InGaAs, InGaP, InAlAs и InP-подложки хорошо согласуются, что позволяет довести мольную долю In до $x \approx 0,6$ и увеличить подвижность электронов при 300 К примерно вдвое. Введение In в GaAs сужает запрещенную зону, причем разрыв зоны проводимости в гетероструктуре $\text{In}_{0,53}\text{Ga}_{0,47}\text{As}/\text{GaAs}$ составляет около 0,5 В (вдвое больше, чем в $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}/\text{GaAs}$).



Весьма перспективны гетероструктуры на основе традиционных полупроводников Si/SiGe [12–16], в которых могут быть получены в 2–3 раза более высокие, чем в кремнии, значения подвижности как электронов, так и дырок [12].

Рассогласование постоянных решетки Ge и Si составляет 4,2%, что вызывает механические напряжения в более тонком слое гетероструктуры. На рис. VIII.1.1 показаны энергетические диаграммы гетеропереходов Si/Si_{0,7}Ge_{0,3}, в которых напряженными являются слои Si/Si_{0,7}Ge_{0,3} (а) или Si (б). В первом случае реализуется гетеропереход I-го рода, во втором — гетеропереход II-го рода. В последнем случае разрывы зоны проводимости и валентной зоны примерно одинаковы при почти одинаковой ширине запрещенной зоны ($\Delta E_g = 0,1$ эВ), что позволяет создавать на ее основе комплементарные пары полевых транзисторов с высокими подвижностями электронов и дырок в канале: $\mu_n = 1270\text{--}2830$ см²/В·с, $\mu_p = 800\text{--}1000$ см²/В·с [13,14].

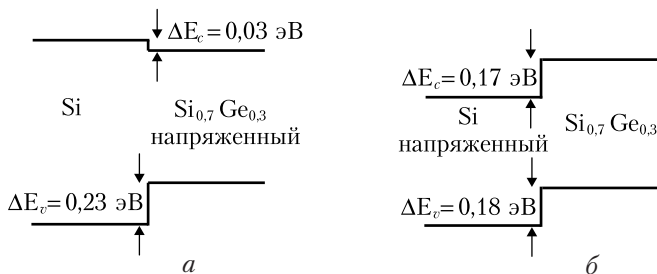


Рис. VIII.1.1. Энергетические диаграммы гетеропереходов Si/Si_{0,7}Ge_{0,3}:
а — напряженный слой Si_{0,7}Ge_{0,3}; б — напряженный слой Si

Наиболее совершенными технологическими методами изготовления гетеропереходов являются молекулярно-лучевая эпитаксия (МБЕ), позволяющая создавать сверхтонкие (вплоть до моноатомных) полупроводниковые слои, и эпитаксия металлоорганических соединений из газовой фазы (МОСVD).

VIII.1.2. Гетеропереходы на основе GaAs

Зонные диаграммы полупроводников $n\text{-Al}_x\text{Ga}_{1-x}\text{As}$, $p\text{-GaAs}$ и равновесная энергетическая диаграмма $p\text{-}n$ перехода показаны на рис. VIII.1.2.

Суммарный разрыв зоны проводимости складывается из разрыва зоны проводимости и разрыва валентной зоны:

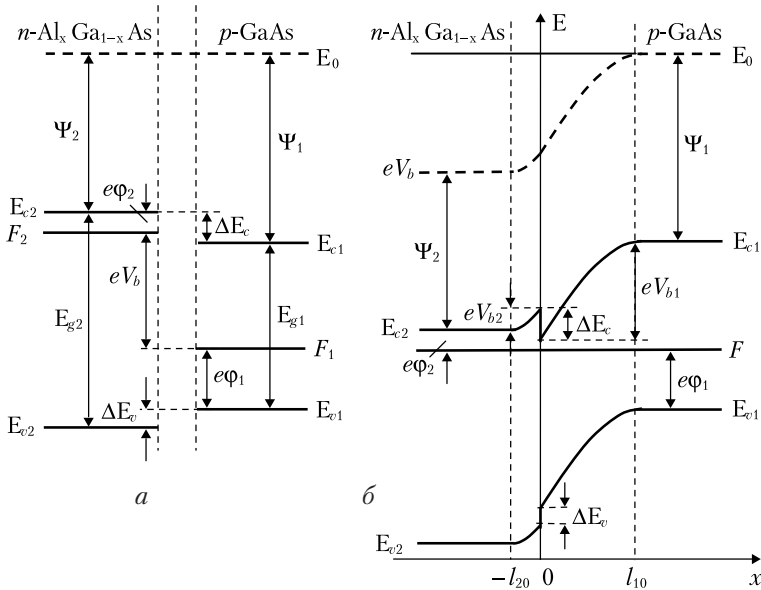


Рис. VIII.1.2. Зонные диаграммы полупроводников $n\text{-Al}_x\text{Ga}_{1-x}\text{As}$, $p\text{-GaAs}$ (а) и равновесная энергетическая диаграмма $p\text{-}n$ перехода $n\text{-Al}_x\text{Ga}_{1-x}\text{As} - p\text{-GaAs}$ (б)

$$\Delta E_g = \Delta E_c + \Delta E_v. \quad (\text{VIII.1.3})$$

В первом приближении разрыв зоны проводимости, как и ΔE_g , пропорционален мольной доле гетерокомпонента

$$\Delta E_c + x\Delta E_{c0}. \quad (\text{VIII.1.4})$$

При отсутствии вырождения расстояния $\phi_{1,2}$ между уровнем Ферми и ближайшей к нему границей запрещенной зоны в электронейтральных областях определяются степенью их легирования (концентрации примеси N_1 и N_2 предполагаются постоянными):

$$p_{10} = N_1 = N_{v1} e^{-\phi_1/\phi_T}; \quad \phi_1 = \phi_T \ln(N_{v1}/N_1); \quad (\text{VIII.1.5a})$$

$$n_{20} = N_2 = N_{c2} e^{-\phi_2/\phi_T}; \quad \phi_2 = \phi_T \ln(N_{c2}/N_2). \quad (\text{VIII.1.5б})$$

Как следует из рис. VIII.1.2, а, контактная разность потенциалов составляет

$$V_b = (F_2 - F_1)/e = (E_{g1} + \Delta E_c)/e - (\phi_1 + \phi_2). \quad (\text{VIII.1.6})$$

Обычно диэлектрические проницаемости полупроводников 1 и 2 практически одинаковы. При этом значения V_b , а также ширины ОПЗ $l_{1,2}$ в областях 1 и 2 определяются так же, как в обычном ступенчатом p - n переходе.

Контактная разность потенциалов V_b складывается из напряжений V_{b1} и V_{b2} на n - и p -участках ОПЗ: $V_b = V_{b1} + V_{b2}$ (рис. VIII.1.2, б), а значения $V_{b1,2}$ могут быть получены двукратным интегрированием уравнений Пуассона в пределах $-l_{20} \leq x \leq 0$ и $0 \leq x \leq l_{10}$:

$$V_{b1,2} = eN_{1,2}l_{10,20}^2 / 2\epsilon\epsilon_0. \quad (\text{VIII.1.7})$$

VIII.1.3. Инверсионный слой в гетеропереходе и двумерный электронный газ

Если концентрация акцепторов N_1 в p -GaAs области гетероперехода достаточно мала, а концентрация доноров N_2 в n -Al_xGa_{1-x}As области достаточно велика, уровень Ферми в GaAs вблизи границы раздела попадает в зону проводимости. Такая ситуация реализуется в гетеропереходах n^+ -Al_xGa_{1-x}As — p -GaAs (рис. VIII.1.3). В результате вблизи границы раздела формируется тонкий инверсионный слой (канал), обогащенный электронами, подобно тому, как это происходит в структуре МДП.

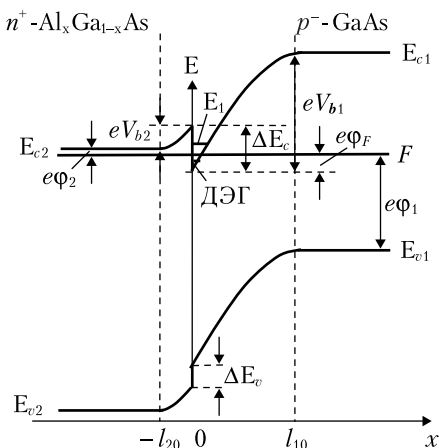


Рис. VIII.1.3. Равновесная энергетическая диаграмма гетероперехода n^+ -Al_xGa_{1-x}As — p -GaAs с двумерным электронным газом в инверсионном слое

Ввиду малой эффективной массы электронов в канале квантование их энергетического спектра в направлении оси x ограничивает движение электронов в этом направлении (электроны не могут перемещаться в направлении оси x без изменения энергии на ΔE_j). Таким образом, электроны в канале образуют систему с пониженной размерностью — *двухмерный электронный газ* (ДЭГ). Форма потенциальной ямы вблизи границы гетероперехода ($x = 0$) близка к треугольной; при этом энергетический спектр имеет вид

$$E_j \approx E_{c1}(0) + [3ehE(0)(j + 3/4)]^{2/3} / (32m_n)^{1/3}, \quad (\text{VIII.1.8})$$

где $E(0) = \frac{1}{e} \cdot \frac{\partial E_{c1}}{\partial x}(0)$ — напряженность электрического поля при $x = 0$.

Поверхностная плотность ДЭГ для j -й подзоны составляет

$$n_{sj} = \frac{m_n}{\pi \hbar^2} \sum_{j=0}^{\infty} \int_{E_j}^{\infty} \frac{dE}{1 + \exp[(E - F) / kT]},$$

где множитель в правой части соответствует плотности разрешенных состояний для двухмерного газа. Практически достаточно учесть только две первые подзоны ($j = 0$ и $j = 1$). Произведя интегрирование, получаем

$$n_s \approx \frac{m_n kT}{\pi \hbar^2} \ln \left[\left(1 + \exp \frac{F - E_0}{kT} \right) \left(1 + \exp \frac{F - E_1}{kT} \right) \right], \quad (\text{VIII.1.9})$$

где значения $E_{0,1}$ определены в (VIII.1.8).

С другой стороны, поверхностная плотность электронов ДЭГ связана с полем $E(0)$ уравнением Гаусса (см. рис. VIII.1.3)

$$E(0) = e(n_s - N_1 l_1) / \epsilon_1 \epsilon_0 \approx en_s / \epsilon_1 \epsilon_0. \quad (\text{VIII.1.10})$$

Последнее приближение оправдано малостью поверхностной плотности заряда акцепторных ионов в обедненном слое p^- -GaAs. Подставляя (VIII.1.10) в (VIII.1.8) для $j = 0$ и $j = 1$, получаем

$$E_{0,1} \approx E_{c1}(0) + e\gamma_{0,1} n_s^{2/3}, \quad (\text{VIII.1.11})$$

$$\text{где } \gamma_0 = \frac{1}{2} \left(\frac{9\pi \hbar \sqrt{e / m_n}}{4\epsilon_1 \epsilon_0} \right)^{2/3}, \quad \gamma_1 = \gamma_0 (7/4)^{2/3}.$$

Подстановка (VIII.1.11) в (VIII.1.9) дает



$$n_s \approx \frac{m_n kT}{\pi \hbar^2} \ln \left[\left(1 + \exp \frac{\Phi_F - \gamma_0 n_s^{2/3}}{\Phi_T} \right) \left(1 + \exp \frac{\Phi_F - \gamma_1 n_s^{2/3}}{\Phi_T} \right) \right], \quad (\text{VIII.1.12})$$

где $e\Phi_F = F - E_{c1}(0)$ (см. рис. VIII.1.3).

Расчеты показывают, что функция $\Phi_F(n_s)$, определяемая уравнением (VIII.1.12), близка к линейной:

$$\Phi_F(n_s) \approx en_s \Delta d / \epsilon_1 \epsilon_0 \equiv -Q_{sn} \Delta d / \epsilon_1 \epsilon_0, \quad (\text{VIII.1.13})$$

где $Q_{sn} = -en_s$ — поверхностная плотность заряда ДЭГ; Δd — параметр, смысл которого объясняется в параграфе VIII.2.2.

При $\Delta d = 8$ нм погрешность аппроксимации (VIII.1.13) для GaAs не превышает 0,1 В в температурном диапазоне 0—300 К. С ростом температуры зависимость $\Phi_F(n_s)$ приближается к экспоненциальной. Такая ситуация имеет место в кремниевых МДПТ, где эффекты квантования не сказываются при обычных рабочих температурах. Линейный характер зависимости $\Phi_F(n_s)$ является особенностью ГПТ.

Поверхностная плотность ДЭГ n_s и толщина ОПЗ l_2 в слое $n^+ - \text{Al}_x \text{Ga}_{1-x} \text{As}$ могут быть найдены из условия электронейтральности перехода в целом:

$$Q_{s2} + Q_{sn} + Q_{s1} = 0,$$

где $Q_{s2} = eN_2 l_2$ — поверхностная плотность заряда в слое $n^+ - \text{Al}_x \text{Ga}_{1-x} \text{As}$; Q_{sn} определяется соотношением (VIII.1.13), а величиной Q_{s1} , как отмечалось выше, можно пренебречь. Исключая параметры V_{b2} и ϕ_1 из соотношений (VIII.1.7), (VIII.1.13) и очевидного из рисунка VIII.1.3 соотношения $\Delta E_c / e = \phi_2 + V_{b2} + \phi_1$, получаем

$$n_s = N_2 \Delta d \left[\sqrt{1 + \frac{2\epsilon\epsilon_0 \left(\frac{\Delta E_c}{e} - \phi_2 \right)}{eN_2 \Delta d^2}} - 1 \right]; \quad (\text{VIII.1.14})$$

$$l_2 = n_s / N_2, \quad (\text{VIII.1.15})$$

где обычно $\phi_2 \ll \Delta E_c / e$. При $N_2 = 2 \cdot 10^{18} \text{ см}^{-3}$, $\frac{\Delta E_c}{e} = 0,25 \text{ В}$ и $\Delta d = 8 \text{ нм}$ параметры имеют следующие значения: $n_s = 2,5 \cdot 10^{12} \text{ см}^{-2}$; $l_2 = 12 \text{ нм}$.

Основные выводы

1. Хорошими парами для гетеропереходов являются полупроводниковые соединения A^3B^5 , из которых наиболее изученной яв-



ляется пара GaAs/Al_xGa_{1-x}As. Введение индия в GaAs существенно повышает подвижность электронов.

2. В гетероструктурах Si/SiGe могут быть получены в 2–3 раза более высокие, чем в кремнии, значения подвижности как электронов, так и дырок. Использование механических напряжений в более тонком слое гетероструктуры позволяет направленно изменять свойства энергетической диаграммы.

3. Если значения диэлектрических проницаемостей в гетерообластях различаются незначительно, ширина гетероперехода, ее зависимость от напряжения и контактная разность потенциалов определяются теми же соотношениями, что и для обычного *p-n* перехода.

4. В гетеропереходах *n*⁺-Al_xGa_{1-x}As — *p*⁺-GaAs вблизи границы раздела формируется тонкий инверсионный слой, обогащенный электронами, с пониженной размерностью — двухмерный электронный газ (ДЭГ).

5. Поверхностная концентрация ДЭГ является линейной функцией расстояния между уровнем Ферми и дном зоны проводимости на границе раздела.

VIII.2. Гетеропереходные полевые транзисторы на основе GaAs

VIII.2.1. Устройство и принцип действия ГПТ на основе GaAs

Гетеропереходные полевые транзисторы (ГПТ) на основе GaAs были предложены как модификация GaAs полевых транзисторов с затвором Шоттки (ПТШ) [11], в которых потенциально высокая подвижность электронов реализуется не в полной мере. Максимальное значение предельной частоты ПТШ достигается при концентрации доноров в канале $N = (1-3) \cdot 10^{17} \text{ см}^{-3}$. При этом подвижность электронов при 300 К составляет $(3-4) \cdot 10^3 \text{ см}^2/\text{В} \cdot \text{с}$, т.е. $\approx 40\%$ максимально возможной. Основная идея ГПТ состоит в отделении электронов в канале от ионов примеси с целью повышения подвижности электронов до предельного значения $(8-9) \cdot 10^3 \text{ см}^2/\text{В} \cdot \text{с}$. С этой идеей связано широко распространенное название ГПТ на основе GaAs — HEMT (High Electron Mobility Transistor).

Структура ГПТ представлена на рис. VIII.2.1. На полуизолирующей подложке *i*⁺-GaAs 1 методом молекулярно-лучевой эпи-



таксии выращивается вначале нелегированный слой i -GaAs 2 с высокой подвижностью электронов, а затем сильнолегированный ($\approx 10^{18} \text{ см}^{-3}$) гетерослой n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ 3 толщиной $d_2 \sim 100 \text{ нм}$. Тонкий промежуточный спейсер i -AlGaAs 4 (spacer — прокладка) толщиной $\sim 10 \text{ нм}$ нейтрализует поверхностные дефекты и влияние ионов примеси в n^+ - $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ на подвижность электронов в канале. Барьерный контакт затвора и омические контакты стока и истока формируются таким же образом, как в GaAs ПТШ (см. раздел VII.2).

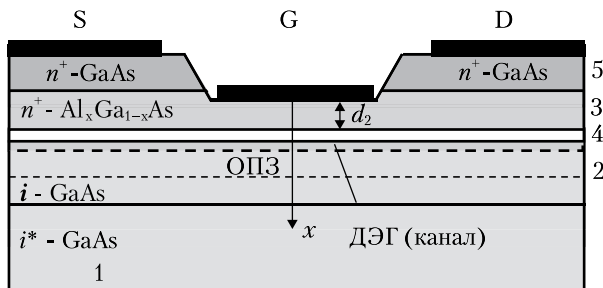


Рис. VIII.2.1. Структура ГПТ на основе GaAs:

- 1 — подложка i^* -GaAs; 2 — i -GaAs; 3 — n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ (10^{18} см^{-3} , $\approx 100 \text{ нм}$);
4 — i -AlGaAs (спейсер, $\approx 10 \text{ нм}$); 5 — контактный слой n^+ -GaAs;
ДЭГ — двухмерный электронный газ (канал)

Энергетическая диаграмма ГПТ в поперечном сечении канала (ось x на рис. VIII.2.1) при $V_{gs} > 0$ представлена на рис. VIII.2.2. Потенциальный барьер eV_{bm} на границе металл затвора— n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ изгибает вверх энергетические уровни E_{c2} и E_{v2} таким образом, что при достаточно малой толщине d_2 барьерного слоя он оказывается полностью обедненным электронами. Таким образом, несмотря на высокую степень легирования, барьерный слой n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ выполняет функцию подзатворного диэлектрика с той лишь разницей, что содержит положительный объемный заряд доноров с объемной плотностью eN_2 . Заметим, что при большой толщине этого слоя уровни E_{c2} и E_{v2} снижались бы монотонно по мере удаления от границы с i -GaAs (против оси x). На рис. VIII.2.2 этот случай представлен штриховыми линиями.

Поверхностная концентрация электронов в канале контролируется напряжением затвор—исток. Таким образом, свойства ГПТ близки к свойствам МДПТ.

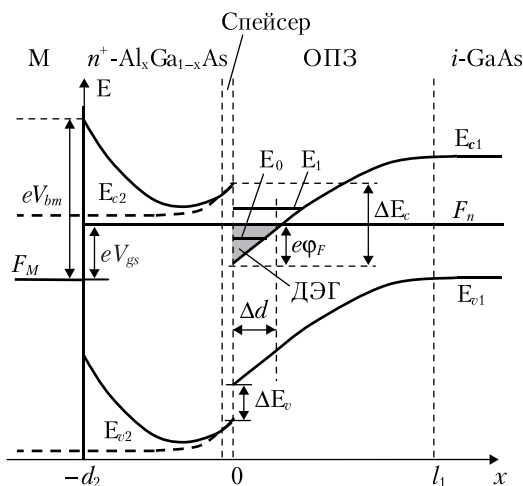


Рис. VIII.2.2. Энергетическая диаграмма ГПТ на основе GaAs при $V_{gs} > 0$.

Штриховые линии соответствуют границам запрещенной зоны E_{c2} и E_{v2} в слое $n^+-Al_xGa_{1-x}As$ большой толщины

VIII.2.2. Пороговое напряжение

Пороговое напряжение ГПТ может быть определено так же, как в МДПТ (см. параграф III.2.4). Различие состоит в том, что подзатворный диэлектрик (обедненный слой $n^+-Al_xGa_{1-x}As$) содержит положительный заряд доноров, поверхностный заряд на границе диэлектрик—канал практически отсутствует, а зависимость плотности электронов в канале от положения квазиуровня Ферми линейна (VIII.1.13).

Согласно энергетической диаграмме на рис. VIII.2.2 напряжение затвор—исток (без учета спейсера) может быть представлено в виде

$$V_{gs} = V_b + \int_{-d_2}^0 E(x) dx - \Delta E_c / e + \phi_F, \quad (\text{VIII.2.1})$$

где при $\epsilon_1 = \epsilon_2 = \epsilon$

$$E(x) = E(-d_2) + \frac{1}{\epsilon \epsilon_0} \int_{-d_2}^x e N_2 dx = E(-d_2) + \frac{e N_2 (x + d_2)}{\epsilon \epsilon_0} \quad (\text{VIII.2.2})$$

— напряженность поля в диэлектрике. По теореме Гаусса граничное значение поля $E(-d_2)$ составляет

$$E(-d_2) = Q_{sM} / \epsilon \epsilon_0. \quad (\text{VIII.2.3})$$

Плотность поверхностного заряда Q_{sM} в металлическом электроде затвора может быть найдена из уравнения общей электронейтральности структуры

$$Q_{sM} + Q_{s2} + Q_{sn} + Q_{s1} = 0, \quad (\text{VIII.2.4})$$

где $Q_{s2} = eN_2d_2$ — поверхностная плотность заряда в слое $n^+ - \text{Al}_x\text{Ga}_{1-x}\text{As}$. Как отмечалось в параграфе VIII.1.3, поверхностной плотностью заряда Q_{s1} в слое $i\text{-GaAs}$ можно пренебречь. При этом из (VIII.2.3) и (VIII.2.4) получаем

$$E(-d_2) = (-Q_{sn} + eN_2d_2) / \epsilon \epsilon_0. \quad (\text{VIII.2.5})$$

Подстановка этого соотношения в (VIII.2.3) и затем в (VIII.2.1) дает

$$V_{gs} = V_b - \frac{eN_2d_2^2}{2\epsilon\epsilon_0} - \frac{\Delta E_c}{e} + \phi_F - \frac{Q_{sn}d_2}{\epsilon\epsilon_0}. \quad (\text{VIII.2.6})$$

Используя зависимость (VIII.1.13), связывающую величину ϕ_F с поверхностной плотностью заряда электронов в канале Q_{sn} , из (VIII.2.6) получаем

$$Q_{sn} = -C_{Gs}(V_{gs} - V_t), \quad (\text{VIII.2.7})$$

$$\text{где} \quad V_t = V_b - V_{p2} - \Delta E_c / e \quad (\text{VIII.2.8})$$

— пороговое напряжение ГПТ (при $V_{gs} < V_t$ транзистор закрыт: $Q_{sn} = 0$);

$$C_{Gs} = \epsilon \epsilon_0 / (d_2 + \Delta d) \quad (\text{VIII.2.9})$$

— удельная (на единицу площади канала) емкость затвор—канал,

$$V_{p2} = eN_2d_2^2 / 2\epsilon\epsilon_0 \quad (\text{VIII.2.10})$$

— напряжение перекрытия слоя $n^+ - \text{Al}_x\text{Ga}_{1-x}\text{As}$.

Нужная величина порогового напряжения может быть обеспечена соответствующим выбором толщины и степени легирования слоя $n^+ - \text{Al}_x\text{Ga}_{1-x}\text{As}$ (т.е. величины V_{p2}). Практически подгонка порогового напряжения осуществляется сравнением с поверхности кристалла контактного слоя $n^+ - \text{GaAs}$ и части слоя $n^+ - \text{Al}_x\text{Ga}_{1-x}\text{As}$ (см. рис. VIII.2.1). Таким способом можно создавать как нормально открытые, так и нормально закрытые ГПТ.



Интересно отметить, что в соответствии с (VIII.2.9) эффективная толщина подзатворного диэлектрика составляет $d_{ef} = d_2 + \Delta d$, что является следствием линейной зависимости $\varphi_F(n_s)$ (VIII.1.13) в структуре с ДЭГ. Избыточную толщину Δd можно трактовать как эффективную толщину ДЭГ.

VIII.2.3. ВАХ ГПТ

Соотношение (VIII.2.7), определяющее зависимость поверхностной плотности заряда электронов в канале, в точности совпадает с аналогичным соотношением для идеализированного МДПТ (III.3.9). Поэтому ВАХ идеализированного ГПТ (нулевые сопротивления стока и истока, постоянная подвижность электронов в канале) в крутой области совпадают по виду с ВАХ идеализированного МДП-транзистора до тех пор, пока ток затвора мал, а электроны в подзатворном слое $n^+ - \text{Al}_x \text{Ga}_{1-x} \text{As}$ отсутствуют:

$$I_d = \beta V_{ds}(V_{gst} - V_{ds} / 2), \quad (\text{VIII.2.11})$$

где
$$\beta = \mu_n Z C_{Gs} / L. \quad (\text{VIII.2.12})$$

Количественное отличие состоит в значительно большей подвижности электронов и, соответственно, крутизне ВАХ.

Ввиду отсутствия проводящей подложки эффекты короткого канала в ГПТ проявляются весьма слабо за исключением эффекта насыщения дрейфовой скорости электронов. Этот эффект, как и в ПТШ, проявляется в большей степени, так как значение порогового поля $E_S = v_s / \mu_n$ мало. Количественно эффект насыщения дрейфовой скорости описывается так же, как в МДПТ (см. параграфы III.5.5—III.5.7).

Сопротивление истока R_S значительно снижает крутизну ГПТ ввиду высокой крутизны «внутреннего» транзистора, поэтому проходные ВАХ ГПТ ближе к линейным, чем к квадратичным (как в МДП-транзисторе).

Барьерный переход затвор—канал имеет диодную характеристику, как в ПТШ. Поэтому при протекании затворного тока ВАХ ГПТ имеет те же особенности, что и ВАХ ПТШ (см. параграф VII.2.3).

Специфической особенностью ГПТ является эффект экранирования ДЭГ в канале электронами, которые появляются в слое $n^+ - \text{Al}_x \text{Ga}_{1-x} \text{As}$ при повышении напряжения затвор—исток. Этот эффект может проявляться уже при таких напряжениях V_{gs} , при которых ток затвора остается малым. С ростом напряжения V_{gs} уменьшается зазор между энергетическими



уровнями E_{c2} и F_n (см. рис. VIII.2.2). Когда этот зазор становится достаточно малым, в слое $n^+-Al_xGa_{1-x}As$ создается второй канал, параллельный каналу с ДЭГ, который экранирует канал с ДЭГ от поля затвора. Поскольку подвижность электронов в слое $n^+-Al_xGa_{1-x}As$ значительно меньше, чем в ДЭГ, крутизна ВАХ при этом резко снижается и ток стока практически насыщается. Можно показать, что максимально достижимая поверхностная плотность электронов ДЭГ, при которой начинается экранирование ДЭГ, составляет

$$n_{smax} \approx \Delta d N_2 (\sqrt{1 + N_0 / N_2} - 1), \quad (VIII.2.11)$$

$$\text{где} \quad N_0 = 2\epsilon\epsilon_0\Delta E_c / (e\Delta d)^2 \approx 5 \cdot 10^{18} \text{ см}^{-3}. \quad (VIII.2.13)$$

VIII.2.4. Разновидности ГПТ на основе соединений A^3B^5

Разновидностью ГПТ на основе GaAs являются приборы с *обращенной* структурой [17]. В обращенном ГПТ узкозонный слой GaAs, в котором формируется канал, расположен между барьерным контактом и широкозонным гетерослоем AlGaAs. Такая структура имеет некоторые преимущества. В частности, открытая поверхность GaAs более стабильна, чем AlGaAs. Кроме того, роль «подзатворного диэлектрика» исполняет нелегированный слой GaAs, что обеспечивает более высокую производительность порогового напряжения.

Характеристики ГПТ могут быть существенно улучшены за счет использования новых полупроводниковых соединений A^3B^5 . Весьма перспективными оказались соединения InGaAs, InGaP, InAlAs и InP. Введение индия в GaAs существенно повышает подвижность электронов и сужает запрещенную зону, причем разрыв зоны проводимости в гетероструктуре $In_{0,53}Ga_{0,47}As/GaAs$ составляет около 0,5 В (вдвое больше, чем в $Al_{0,3}Ga_{0,7}As/GaAs$). Это обстоятельство способствует нейтрализации паразитного канала. На основе такой структуры реализуются *псевдоморфные* ГПТ (p-HEMT) на InP-подложке (структура InGaAs/InP) [18–21]. Простейшая структура p-HEMT $In_{0,53}Ga_{0,47}As/InP$ представлена на рис. VIII.2.3. В работах [18, 20] описаны структуры InGaP/InGaAs p-HEMT с двойным δ -легированным каналом. Отмечается снижение выходной проводимости, повышение линейности проходных ВАХ, а также существенное расширение температурного диапазона работы прибора (до 450 °С [20]).



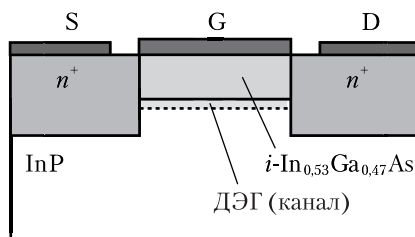


Рис. VIII.2.3. Простейшая структура псевдоморфного НЕМТ
 $\text{In}_{0,53}\text{Ga}_{0,47}\text{As}/\text{InP}$

Главным недостатком р-НЕМТ InGaAs/InP является сложность изготовления InP -подложек и, как следствие, их более высокая стоимость. Это обстоятельство стимулировало разработку *метаморфных* НЕМТ (m-НЕМТ), имеющих структуру $\text{AlInAs}/\text{InGaAs}$ на GaAs -подложке. Для согласования атомных решеток слоя InGaAs с GaAs -подложкой используется многослойная варизонная система буферных слоев [22, 23]. В настоящее время характеристики m-НЕМТ почти не уступают р-НЕМТ на InP -подложке. Более того, на СВЧ подложка GaAs обладает значительно лучшими диэлектрическими свойствами, чем InP .

Сравнительные характеристики полевых транзисторов на полупроводниках A^3B^5 представлены в табл. VIII.2.1 [24].

Таблица VIII.2.1

Сравнительные характеристики полевых транзисторов
на полупроводниках A^3B^5

Тип ПТ	Длина канала, мкм	Пороговое напряжение, В	Крутизна, мСм/мм	Предельная частота, ГГц
ПТШ	0,12		500	160
	0,3		265	40
	0,3		200	56
ГПТ $\text{AlGaAs}/\text{GaAs}$	1,8		190	15
	0,03		215	-
$\text{InGaP}/\text{InGaAs}$ р-НЕМТ	1	+0,26	150	8,3
		-0,6	150	16
	0,25		800	87
	0,13		1500	245
	0,03		1500	540
m-НЕМТ	0,07		800	> 300
	0,13			235
	0,35	+0,1	660	75
		-0,55	510	80

Для использования в мощных СВЧ-устройствах, а также в устройствах, предназначенных для работы в жестких эксплуатационных условиях, значительные преимущества имеют транзисторы на основе широкозонных полупроводниковых соединений — GaN и SiC. Структура ГПТ AlGaN/GaN схематично представлена на рис. VIII.2.4 [25]. Канал с ДЭГ формируется в узкозонном слое AlGaN. Теория ГПТ AlGaN/GaN, методы их изготовления и характеристики (в том числе на кремниевой подложке) описаны в работах [25—28 и др.].

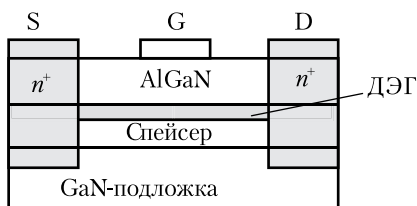


Рис. VIII.2.4. Структура ГПТ AlGaN/GaN

VIII.2.5. ГПТ на основе структур Si/SiGe

Ввиду высокой степени освоения промышленностью кремниевой технологии и низкой стоимости исходного материала важное практическое значение имеют гетероструктурные полевые транзисторы на основе традиционного материала микроэлектроники — кремния. Наличие высококачественного собственного оксида делает возможным создание на основе Si гетероструктурных МДПТ.

Типовые конструкции комплементарных МДПТ на основе гетероструктур Si/SiGe показаны на рис. VIII.2.5 [29]. Структуры используют гетеропереходы II-го рода с тонким напряженным слоем Si (см. рис. VIII.1.1, б), в котором и формируются каналы. Как и в обычной МДП-технологии, *p*-канальные транзисторы формируются в *n*-карманах. Подстройка их порогового напряжения осуществляется подбором степени легирования единственного эпитаксиального слоя под *p*-каналом. Характеристики МДПТ в значительной степени зависят от качества поверхности раздела Si/SiO₂ [30].

На рис. VIII.2.6 показаны многослойные гетероструктуры, на основе которых формируются комплементарные ГПТ с затворным контактом барьерного типа [29]. Достигнутые параметры ГПТ со структурой Si/SiGe приведены

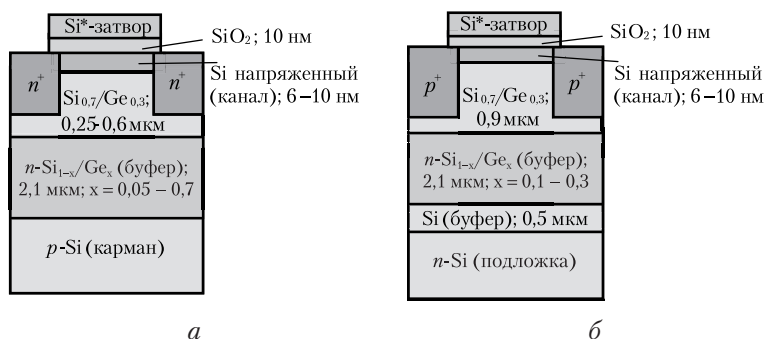


Рис. VIII.2.5. Типовые структуры гетеропереходных МДПТ Si/SiGe:
a — *n*-канал, *б* — *p*-канал

Si – контактный слой	Si – контактный слой	Si – контактный слой
Si _{1-x} Ge _x	Si _{1-x} Ge _x	Si _{1-x} Ge _x
<i>n</i>	<i>p</i> ⁺	<i>p</i> ⁺
Si _{1-x} Ge _x – спейсер	Si – спейсер	Si _{1-x} Ge _x – спейсер
Si – канал	Si _{1-x} Ge _x – канал	Ge _x – канал
Si _{1-x} Ge _x – ненапряженный буфер	Si – буфер	Si _{1-x} Ge _x – ненапряженный буфер
Si – подложка	Si – подложка	Si – подложка

Рис. VIII.2.6. Варианты структуры ГПТ Si/SiGe:
a – Si – *n*-канал на ненапряженном буфере SiGe; *б* – SiGe – *p*-канал; *в* – Ge – *p*-канал на ненапряженном буфере SiGe

в табл. VIII.2.3 [24]. Данные таблицы показывают, что по частотным и усилительным свойствам ГПТ Si/SiGe в 2–3 раза уступают р- и m-НЕМТ на основе соединений A³B⁵.

Таблица VIII.2.3

Характеристики ГПТ Si/SiGe

Тип транзистора	Тип канала	Длина канала, мкм	Крутизна, мСм/мм	Предельная частота, ГГц
ГПТ	<i>n</i>	0,08	250	70–80
ГПТ	<i>n</i>	1,4	340	
ГПТ	<i>n</i>	0,25	330	
ГПТ	<i>n</i>	0,5	390	
ГПТ	<i>p</i>	0,25	230	24
ГПТ	<i>p</i>	0,23	280	30



Основные выводы

1. В ГПТ на основе GaAs высокая подвижность электронов в канале достигается за счет отделения электронов в канале от ионов примеси. Функцию подзатворного диэлектрика выполняет сильно легированный обедненный барьерный слой $n^+-\text{Al}_x\text{Ga}_{1-x}\text{As}$.

2. По своим свойствам ГПТ на основе GaAs близок к МДП транзистору. Главным их различием является возможность протекания в ГПТ затворного тока. Пороговое напряжение определяется так же, как в МДПТ, с той разницей, что подзатворный диэлектрик содержит положительный заряд доноров, поверхностный заряд на границе диэлектрик—канал практически отсутствует, а зависимость плотности электронов в канале от положения квазиуровня Ферми линейна.

3. Специфической особенностью ГПТ является эффект экранирования ДЭГ в канале электронами, которые появляются в слое $n^+-\text{Al}_x\text{Ga}_{1-x}\text{As}$ при повышении напряжения затвор—исток. Этот эффект ограничивает максимально достижимую поверхностную плотность электронов ДЭГ.

4. Характеристики ГПТ могут быть существенно улучшены за счет введения в GaAs индия. Псевдоморфные ГПТ на InP-подложке обладают наиболее высокими значениями удельной крутизны ВАХ и предельной частоты. К ним близки по характеристикам более дешевые метаморфные ГПТ, имеющие структуру AlInAs/InGaAs на GaAs-подложке.

5. Для использования в мощных СВЧ-устройствах, а также в устройствах, предназначенных для работы в жестких эксплуатационных условиях, значительные преимущества имеют транзисторы на основе широкозонных полупроводниковых соединений — GaN и SiC.

6. ГПТ на основе Si/SiGe в 2–3 раза уступают по частотным и усилительным свойствам метаморфным и псевдоморфным ГПТ на основе соединений A^3B^5 , однако позволяют реализовать комплементарные транзисторные пары, превосходящие по характеристикам традиционные МДП-транзисторы.

VIII.3. Гетеропереходные биполярные транзисторы

VIII.3.1. Возможности улучшения характеристик биполярных транзисторов за счет использования гетеропереходов

Требования к электрофизическим и геометрическим параметрам структуры биполярных транзисторов во многом являются противоречивыми. Действительно, для обеспечения



высокого коэффициента передачи базового тока необходимо выполнение условий (IV.1.1) и (IV.1.2)

$$w_B / L_B \ll 1; \quad (\text{VIII.1.1})$$

$$N_E / N_B \gg 1, \quad (\text{VIII.1.2})$$

обеспечивающих близкие к 1 значения коэффициента переноса и эффективности эмиттера.

Снижение степени легирования базы трудно осуществимо технологически. Кроме того, уменьшение толщины базы w_B и степени ее легирования ведет к увеличению сопротивления активной базы, что ухудшает частотные свойства транзистора и снижает его усилительные свойства при больших плотностях тока (эффект оттеснения эмиттерного тока, см. параграф IV.5.4). Чрезмерное легирование эмиттера для выполнения условия (VIII.1.2) не приводит к повышению его эффективности вследствие эффектов вырождения (см. параграф IV.1.5). Поэтому концентрация примеси в активной базе не должна превышать $(1-3) \cdot 10^{17} \text{ см}^{-3}$. При этом уменьшение толщины базы до $\sim 0,1 \text{ мкм}$ снижает напряжение прокола базы (см. параграф IV.5.7) и увеличивает выходную проводимость транзистора вследствие эффекта Эрли (см. параграф IV.5.2), что в конечном счете снижает его усилительные свойства.

Повышение степени легирования эмиттера и базы приводит к увеличению барьерной емкости эмиттерного перехода и ухудшению частотных характеристик транзистора.

Таким образом, при выборе электрофизических и геометрических параметров структуры приходится руководствоваться компромиссными соображениями, так как улучшение одних характеристик прибора неизбежно сопровождается ухудшением других.

Радикальным способом преодоления указанных противоречий является выполнение биполярного транзистора в виде гетероструктуры, в которой эмиттер имеет большую ширину запрещенной зоны, чем база.

В бездрейфовом приближении при малой длине базы и эмиттера ($w_{B,E} / L_{B,E} \ll 1$) и без учета токов рекомбинации—генерации в ОПЗ эмиттерного перехода эффективность n -эмиттера полупроводникового диода определяется следующими соотношениями из раздела II.4:

$$j_{Sn} = eD_B n_{p0} / L_B; \quad (\text{VIII.1.3a})$$

$$j_{Sp} = eD_E p_{n0} / L_E; \quad (\text{VIII.1.3б})$$



$$\gamma = \frac{1}{1 + I_{Sp} / I_{Sn}} = \frac{1}{1 + (\omega_B / D_B n_{p0}) / (\omega_E / D_E p_{n0})}. \quad (\text{VIII.1.4})$$

С учетом возможного вырождения электронейтральных областей равновесные концентрации неосновных носителей в базе n_{p0} и эмиттере могут быть выражены через эффективные концентрации примеси

$$n_{p0}(N_B) = n_i^2 / N_{Bef}(N_B); \quad (\text{VIII.1.5a})$$

$$p_{n0}(N_E) = n_i^2 / N_{Eef}(N_E), \quad (\text{VIII.1.5б})$$

причем способ вычисления функций $N_{Bef, Eef}(N_{B, E})$ указан в параграфе I.3.1. Соотношения (VIII.1.3)–(VIII.1.5) использованы в разделе IV.3.2 для представления эффективности эмиттера биполярного транзистора в виде (IV.3.15)

$$\gamma_N = \frac{1}{1 + G_B / G_E}. \quad (\text{VIII.1.6})$$

Числа Гуммеля в базе и эмиттере в случае произвольной степени легирования составляют

$$G_B = N_{Bef} \omega_B / D_B; \quad (\text{VIII.1.7a})$$

$$G_E = N_{Eef} \omega_E / D_E. \quad (\text{VIII.1.7б})$$

Для гетероструктурного биполярного транзистора (ГБТ) с различной шириной запрещенной зоны базы и эмиттера соотношения (VIII.1.5a, б) должны быть скорректированы:

$$n_{p0}(N_B) = n_{iB}^2 / N_{Bef}(N_B); \quad (\text{VIII.1.8a})$$

$$p_{n0}(N_E) = n_{iE}^2 / N_{Eef}(N_E), \quad (\text{VIII.1.8б})$$

где собственные концентрации носителей в базе и эмиттере различны:

$$n_{iB}^2 = N_{cB} N_{vB} e^{-E_{gB} / kT}; \quad (\text{VIII.1.9a})$$

$$n_{iE}^2 = N_{cE} N_{vE} e^{-E_{gE} / kT}. \quad (\text{VIII.1.9б})$$

Таким образом, вместо (VIII.1.6) для ГБТ получаем

$$\gamma_N = \frac{1}{1 + (G_B / G_E) \vartheta_{BE} e^{-\Delta E_g / kT}} = \quad (\text{VIII.1.10a})$$

$$= \frac{1}{1 + G_{Bef} / G_E}, \quad (\text{VIII.1.10б})$$



где $\vartheta_{BE} = N_{cE} N_{vE} / N_{cB} N_{vB}$, $\Delta E_g = E_{gE} - E_{gB}$ — разрыв ширины запрещенной зоны в эмиттере и базе;

$$G_{Bef} = G_B \vartheta_{BE} e^{-\Delta E_g / kT}. \quad (\text{VIII.1.11})$$

Значение коэффициента ϑ_{BE} для полупроводников, образующих хорошие пары для гетеропереходов I-го рода (см. табл. VIII.1.1), обычно не слишком сильно отличается от единицы, но разрыв ширины запрещенной зоны радикально влияет на эффективность эмиттера. При $\Delta E_g \gg kT$ эффективность эмиттера может оставаться чрезвычайно высокой (близкой к единице) даже в том случае, когда эффективность эмиттера в гомогенной структуре мала.

Например, при $\Delta E_g = 0,25$ эВ эффективность эмиттера может составить более 0,99 даже в случае $N_E / N_B < 1/20$. Это обстоятельство позволяет в ГBT легировать базу значительно сильнее, чем эмиттер.

Основные преимущества ГBT состоят в следующем:

1) высокая степень легирования базы обеспечивает малое сопротивление базы;

2) малая степень легирования эмиттера обеспечивает увеличение ширины эмиттерного перехода и, следовательно, снижение его барьерной емкости;

3) использование неоднородной по составу базы позволяет создавать в базе чрезвычайно сильные встроенные электрические поля, которые ускоряют перенос неосновных носителей через базу и улучшают частотные характеристики ГBT;

4) при высокой по сравнению с коллектором степени легирования базы ($N_B \gg N_C$) лишь очень малая часть ОПЗ коллекторного перехода расположена в базе, что способствует нейтрализации эффекта Эрли и повышению напряжения прокола базы даже в случае очень малой ее длины.

К недостаткам ГBT можно отнести снижение подвижности основных и неосновных носителей в базе при сильном легировании.

VIII.3.2. ГBT на основе GaAs

Первые высококачественные ГBT были созданы на основе GaAs с использованием широкозонного эмиттера $\text{Al}_x\text{Ga}_{1-x}\text{As}$ [31–34] и принадлежат к наиболее быстродействующим приборам микроэлектроники.

Простейшая типовая структура ГBT, создаваемая методом молекулярно-лучевой эпитаксии, или вакуумным, или



осаждением металлоорганических соединений (МОСVD), представлена на рис. VIII.3.1 (пирамидальная структура). На полуизолирующей подложке i^* -GaAs последовательно выращиваются контактная область коллектора n^+ -GaAs ($N_C^+ \approx 4 \cdot 10^{18} \text{ см}^{-3}$; $w_C^+ \approx 0,5 \text{ мкм}$) и собственно коллектор n -GaAs ($N_C \approx 4 \cdot 10^{16} \text{ см}^{-3}$; $w_C \approx 0,4 \text{ мкм}$), база p^+ -GaAs ($N_B \approx 5 \cdot 10^{18} \text{ см}^{-3}$; $w_B \approx 0,15 \text{ мкм}$), широкозонный гетерослой эмиттера $n\text{-Al}_{0,3}\text{Ga}_{0,7}\text{As}$ ($N_E \approx 3 \cdot 10^{17} \text{ см}^{-3}$; $w_E \approx 0,2 \text{ мкм}$) и контактная область эмиттера n^+ -GaAs ($N_E^+ \approx 4 \cdot 10^{18} \text{ см}^{-3}$; $w_E^+ \approx 0,1 \text{ мкм}$). Топологическое формирование взаимно изолированных областей коллектора, а также базы и эмиттера осуществляется селективным обтачиванием соответствующих объемов.

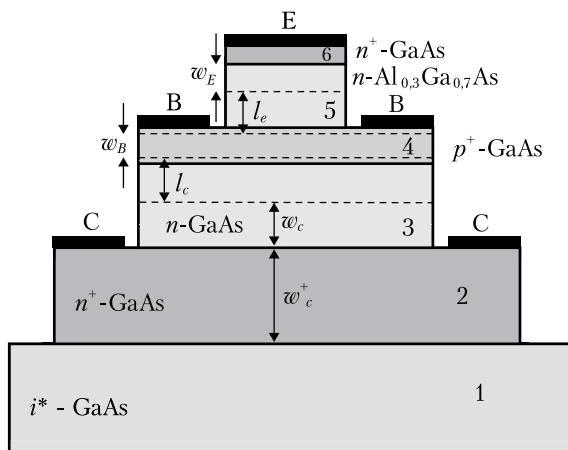


Рис. VIII.3.1. Пирамидальная структура ГБТ на основе GaAs:

- 1 — подложка i^* -GaAs; 2 — контактная область коллектора n^+ -GaAs ($\approx 4 \cdot 10^{18} \text{ см}^{-3}$; 0,5 мкм); 3 — коллектор n -GaAs ($\approx 4 \cdot 10^{16} \text{ см}^{-3}$; 0,4 мкм); 4 — база p^+ -GaAs ($\approx 5 \cdot 10^{18} \text{ см}^{-3}$; 0,15 мкм); 5 — эмиттер $n\text{-Al}_{0,3}\text{Ga}_{0,7}\text{As}$ ($\approx 3 \cdot 10^{17} \text{ см}^{-3}$; 0,2 мкм); 6 — контактная область эмиттера n^+ -GaAs ($\approx 4 \cdot 10^{18} \text{ см}^{-3}$; 0,1 мкм).

Штриховыми линиями показаны границы эмиттерного и коллекторного переходов

Низкая концентрация примеси в активном коллекторном слое обеспечивает малые барьерные емкости коллекторного перехода. Снижение барьерной емкости эмиттерного перехода достигается сравнительно невысокой степенью легирования эмиттера, в области которого расположена большая часть ОПЗ.

Энергетическая диаграмма и профиль мольной доли гетерокомпонента (Al) $x(x)$ в простейшей структуре ГБТ на основе GaAs схематично представлены на рис. VIII.3.2, а.

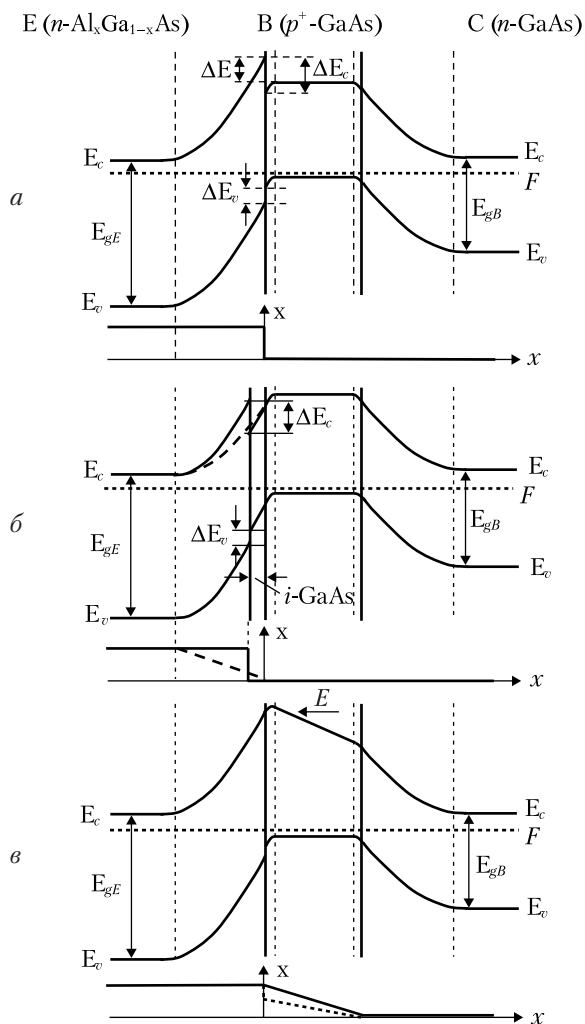


Рис. VIII.3.2. Энергетические диаграммы и профили мольной доли гетерокомпонента (Al) $x(x)$ в структурах ГБТ на основе GaAs:

а — простейшая структура; б — структура с нелегированным спейсером $i\text{-GaAs}$ и варизонным эмиттером; в — структура с варизонной базой

Особенностью энергетической диаграммы является повышение потенциального барьера для электронов эмиттера за счет энергетического «пика» высотой ΔE , связанного с разрывом зоны проводимости (кривая $E_c(x)$ на рис. VIII.3.2, а). Этот «пик» препятствует потоку электронов из эмиттера в базу и может значительно снизить эффективность эмиттера. Существует несколько способов нейтрализации этого вредного эффекта.

Первый из них состоит во введении *нелегированного спейсера* (*i-GaAs*) на границе между базой и эмиттером (рис. VIII.3.2, б). Падение напряжения на спейсере позволяет снизить или вообще исключить дополнительный потенциальный барьер для электронов, как показывает сплошная кривая $E_c(x)$ на рис. VIII.3.2, б.

Второй способ состоит в создании *варизонного эмиттера* (эмиттера с изменяющейся шириной запрещенной зоны) путем плавного уменьшения мольной доли гетерокомпонента до нуля в эмиттерной области перехода (штриховая линия $x(x)$ на рис. VIII.3.2 б). При этом разрывы запрещенной зоны полностью устраняются (штриховая линия $E_c(x)$ на рис. VIII.3.2, б).

Рисунок VIII.3.2, в соответствует ГБТ с *варизонной базой*, сформированной за счет плавного уменьшения мольной доли гетерокомпонента до нуля в области электронейтральной базы. Поскольку база легирована акцепторами равномерно, равновесная концентрация дырок и, следовательно, расстояния между уровнем Ферми и потолком валентной зоны $F - E_v$ не зависят от координаты x в базе. Уменьшение ширины запрещенной зоны в базе по направлению к коллектору осуществляется за счет снижения дна зоны проводимости $E_c(x)$ (см. рис. VIII.3.2, в). Это означает, что в электронейтральной базе существует встроенное *квазиполе для электронов* напряженностью

$$E(x) = \frac{1}{e} \cdot \frac{dE_c(x)}{dx}. \quad (\text{VIII.1.12})$$

Таким образом, ГБТ с варизонной базой является *дрейфовым транзистором*. Встроенное квазиполе направлено от коллектора к эмиттеру (см. рис. VIII.3.2, в) и является *ускоряющим* для электронов, инжектированных в базу из эмиттера. Падение напряжения от встроенного квазиполя на базе составляет

$$\Delta V_B = \Delta E_g / e, \quad (\text{VIII.1.13})$$

где $\Delta E_g = E_{gE} - E_{gB}$. Фактор поля для дрейфовых транзисторов определяется соотношением (IV.6.46), приведенным в параграфе IV.6.1,

$$\eta = \Delta V_B / \Phi_T. \quad (\text{VIII.1.14})$$

Для типичного значения $\Delta E_g = 0,25$ эВ фактор поля $\eta \sim 10$, в то время как в обычных дрейфовых транзисторах его значение редко превышает 2. В параграфе IV.6.3 показано, что дрейфовый механизм переноса неосновных носителей через базу преобладает при $\eta > 2$. Таким образом, ГБТ с варизонной базой может считаться *чисто дрейфовым* транзистором. Более того, при типичной длине базы $w_B \approx 0,15$ мкм средняя напряженность квазиполя в базе составляет $\bar{E} = \Delta V_B / w_B = \Delta E_g / ew_B \approx 17$ кВ/см. При этом электроны пролетают базу с максимальной дрейфовой скоростью $v_s \approx 10^7$ см/с, и снижение их подвижности, связанное с повышенной степенью легирования базы, практически нейтрализуется.

Недостатком ГБТ с варизонной базой является снижение эффективности эмиттера по сравнению со значением, которое дает соотношение (VIII.1.10). По аналогии с соотношением (IV.6.27), обобщающим определение числа Гуммеля в базе на случай произвольного примесного профиля, эффективное число Гуммеля в (VIII.1.11) можно представить в виде

$$G_{B\text{eff}} = \frac{G_B}{w_B} \int_0^{w_B} \vartheta_{BE} e^{-\Delta E_g(x)/kT} dx.$$

Полагая $\vartheta_{BE} \approx 1$ и

$$\Delta E_g(x) = \Delta E_g(w_B - x) / w_B \quad (\text{VIII.1.15})$$

(см. рис. VIII.3.2, в), при $\Delta E_g \gg kT$ получаем

$$G_{B\text{eff}} = G_B \frac{kT}{\Delta E_g} (1 - e^{-\Delta E_g/kT}) \approx G_B \frac{kT}{\Delta E_g}. \quad (\text{VIII.1.16})$$

Сравнение соотношений (VIII.1.16) и (VIII.1.11) показывает, что в ГБТ с варизонной базой эффективное число Гуммеля в базе снижается значительно меньше, чем в ГБТ с *i*-GaAs спейсером или с варизонным эмиттером. Соответственно, меньшим получается выигрыш в эффективности эмиттера, и главные преимущества ГБТ не могут быть использованы в полной мере. Компромиссным решением является снижение мольной доли гетерокомпонента в базе $x(0)$ на границе с эмиттерным переходом (пунктир на рис. VIII.3.2, в).



VIII.3.3. ГBT на основе гетероструктур Si-Ge/Si

В последнее время интенсивно развивается промышленное производство кремниевых ИС, использующих в качестве активных элементов гетеропереходные биполярные транзисторы со структурой Si-Ge/Si [35–38]. В отличие от ГПТ на основе GaAs эти транзисторы имеют узкозонную базу — тонкий напряженный слой $p^+-\text{Si}_{0.7}\text{Ge}_{0.3}$. При этом эмиттерный переход представляет собой гетеропереход I рода (см. рис. VIII.1.1, *a*). Как очевидно из рис. VIII.1.1, *a*, основная часть разрыва ширины запрещенной зоны приходится на валентную зону, что позволяет обойтись без специальных мер по устранению энергетического «пичка», связанного с разрывом зоны проводимости (см. рис. VIII.3.2, *a*).

Типовая структура гетеропереходного биполярного n - p - n транзистора с базой Si/Ge показана на рис. VIII.3.3.

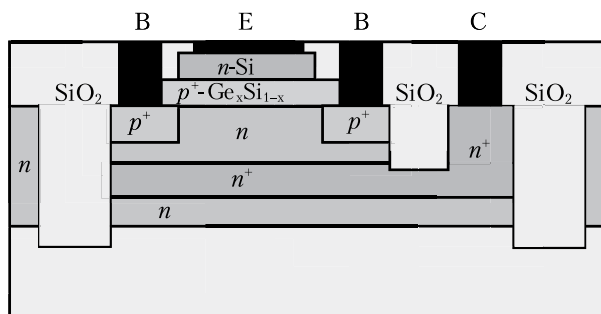


Рис. VIII.3.3. Гетеропереходный биполярный транзистор с базой Si/Ge

В настоящее время гетеропереходные биполярные транзисторы являются наиболее быстродействующими традиционными приборами микроэлектроники. Их предельная частота приближается к 400 ГГц [39], а основные области применения составляют аналоговые и аналого-цифровые ИС для устройств коммуникационной и измерительной техники. Для этих применений отработаны технологические процессы, позволяющие создавать в одном кристалле ГBT и комплементарные пары МДПТ [40–42]. Такой подход обеспечивает оптимальные параметры по критерию цена/качество.

Основные выводы

1. В гетеропереходных биполярных транзисторах с узкозонной базой высокий коэффициент усиления базового тока может быть получен при большей концентрации примеси в базе, чем в эмиттере. За счет этого значительно снижается сопротивление базы и барьерной емкости эмиттерного перехода, а также нейтрализуется действие эффекта Эрли, что позволяет существенно повысить частотные характеристики транзистора, расширить рабочий диапазон плотности эмиттерного тока и повысить напряжение прокола базы.

2. В базе ГПТ с варизонной базой создаются чрезвычайно сильные встроенные электрические поля, которые ускоряют перенос неосновных носителей через базу и улучшают частотные характеристики ГБТ.

3. Наиболее высокими скоростными характеристиками обладают ГПТ на основе гетероструктур A^3B^5 . Широкое практическое применение находят ГПТ на основе гетероструктур Si-Ge/Si.

Литература

1. *Shockley, W.* Пат. США № 2569347 / W. Shockley, 1951.
2. *Губанов, А. И.* ЖТФ / А. И. Губанов. — 1951. — Т. 21. — С. 721, 304.
3. *Kroemer, H.* Theory of wide-gap emitter for transistors / H. Kroemer // Proc. IRE. — 1957. — Vol. 45. — P. 1535.
4. *Anderson, R. A.* IBM J. Res. Dev. / R. A. Anderson. — 1960. — Vol. 4. — P. 283.
5. *Anderson, R. L.* Experiments on Ge-GaAs heterojunctions / R. L. Anderson // Solid State Electron. — 1962. — Vol. 5. — P. 341–345.
6. *Adams, M. J.* A proposal for new approach to heterojunction theory / M. J. Adams, A. Nussbaum, // Solid State Electron. — 1979. — Vol. 22. — P. 783–791.
7. *Von Roos, O.* Theory of extrinsic and intrinsic heterojunctions in thermal equilibrium / O. Von Roos // Solid State Electron. — 1990. — Vol. 23. — P. 1069–1075.
8. *Kroemer, H.* Critique of two recent theories of heterojunction lineups / H. Kroemer // IEEE EDL. — 1983. — Vol. EDL-4(2). — P. 25–27.
9. *Esaki, L.* Internal Report RC 2418 / L. Esaki, R. Tsu // IBM Research. — 1969. — March 26.
10. *Dingle, R.* / R. Dingle [and other] // Appl. Phys. Lett. — 1978. — Vol. — 37. — P. 805.
11. *Mimura, T.* A new FET with selectively doped GaAs/n- Al_xGa_{1-x} As heterojunctions / T. Mimura [and other] // J. Appl. Phys. — 1980. — Vol. 19. — L225–L227.
12. *Hoke, W. E.* / W. E. Hoke [and other] // J. Vac. Sci. Technol. 1999. — Vol. B17. — P. 1131.



13. *Pearsall, T. P.* IEEE Electron Dev. Lett. / T. P. Pearsall, J. C. Bean — 1986. — P. 308.
14. *Mueller, C. H.* High electron mobility in SiGe/Si n-MODFET structure on sapphire substrates / C. H. Mueller, E. T. Croce // Electron. Lett. — 2003. — Vol. 39. — № 18. — P. 1353–1354.
15. *Maiti, C. K.* Strained-Si heterostructure field effect transistors / C. K. Maiti, L. K. Bera, S. Chattopadhyay // Semicind. Sci. Technol. — 1998. — Vol. 13. — P. 1225–1246.
16. *Scäffler, F.* High-mobilities Si and Ge structure / F. Scäffler // Semicond. Sci Technol. — 1997. — Vol. 12. — P. 1515–1549.
17. *Thorne, R. E.* Performance of inverted structure modulation doped Schottky barrier field effect transistors / R. E. Thorne [and other] // Jpn. J. Appl. Phys. Lett. — 1982. — Vol. 21. — L223–L224.
18. *Chuang H.-M.* InGaP/InGaAs double delta-doped channel transistor / H.-M. Chuang [and other] // Electron. Lett. — 2003. — Vol. 39. — № 13. — P. 1016–1018.
19. *Arai, T.* Suppression of drain conductance in InP-based HEMTs by eliminating hole accumulation / T. Arai [and other] // IEEE Trans. — 2003. — Vol. ED-50. — № 5. — P. 1189–1193.
20. *Chuang, H.-M.* Investigation of a new InGaP-InGaAs pseudomorphic double doped channel heterostructure FET (PDDCHFET) / H.-M. Chuang, S.-Y. Cheng, C.-Y. Chen // IEEE Trans. — 2003. — Vol. ED-50. — № 8. — P. 1717–1722.
21. *Sleiman, A.* Channel thickness dependence of breakdown dynamic in InP-based lattice-matched HEMTs. / A. Sleiman [and other] // IEEE Trans. — 2003. — Vol. ED-50. — № 10. — P. 2009–2014.
22. *Hoke, W. E.* Progress in GaAs-based metamorphic technology / W. E. Hoke [and other] // — GaAs MANTECH, Inc., 2001.
23. *Bollaert, S.* Metamorphic $\text{In}_{0.4}\text{Al}_{0.6}\text{As}/\text{In}_{0.4}\text{Ga}_{0.6}\text{As}$ HEMT's on GaAs substrate / S. Bollaert [and other] // IEEE Electron. Dev. Lett. — 1999. — Vol. 20. — № 3. — P.123–125.
24. *Горбацевич, А. А.* Полупроводниковые гетероструктуры и приборы на их основе / А. А. Горбацевич [и др.] // Нанотехнологии в электронике / под ред. Ю. А. Чаплыгина. — М.: Техносфера, 2005. — С. 172–242.
25. *Yu, T. H.* Theoretical study of a GaN-AlGaN high electron mobility transistor including a nonlinear polarization model / T. H. Yu, F. Brennan // IEEE Trans. — 2003. — Vol. ED-50. — № 2. — P. 315–323.
26. *Katz, O.* Characteristics of $\text{In}_x\text{Al}_{1-x}\text{N}$ -GaN high electron mobilities FET / O. Katz [and other] // IEEE Trans. — 2005. — Vol. ED-52. — № 2. — P. 146–150.
27. *Behtash, R.* AlGaN/GaN HEMTs on Si (111) with 6,6 W/mm output power density/ R. Behtash// Electron. Lett. — 2003. — Vol.39. — № 7. — P. 626–628.
28. *Dumka, D. C.* AlGaN/GaN HEMT on Si substrate with 7 W/mm output power density at 10 GHz / D. C. Dumka, C. Lee, H. Q. Tserng // Electron. Lett. — 2004. — Vol. 40. — № 16. — P. 1023–1024.



29. *Maiti, C. K.* Strained-Si heterostructure field effect transistors / C. K. Maiti, L. K. Bera, S. Chattopadhyay // *Semicind. Sci. Technol.* — 1998. — Vol. 13. — P. 1225–1246.
30. *Olsen, S. H.* Strained Si/Ge n-channel MOSFETs: impact of cross-hatching on device performance / S. H. Olsen [and other] // *Semicind. Sci. Technol.* — 2002. — Vol. 17. — P. 655–661.
31. *Bailbe, J. P.* Design and fabrication of high-speed GaAlAs/GaAs heterojunction transistors / J. P. Bailbe [and other] // *IEEE Trans.* — 1980. — Vol. ED-27. — P. 1160–1164.
32. *Kroemer, H.* Heterostructure bipolar transistors and integrated circuits / H. Kroemer // *Proc. IEEE.* — 1982. — Vol. 70. — P. 13025.
33. *Ankri, D.* Diffused epitaxial GaAlAs/GaAs heterojunction bipolar transistor for high frequency operation / D. Ankri [and other] // *Appl. Phys. Lett.* — 1982. — Vol. 40. — P. 816.
34. *Su, S. I.* Double heterojunction GaAs/Al_xGa_{1-x}As bipolar transistors prepared by molecular beam epitaxy / S. I. Su [and other] // *J / Appl. Phys.* — 1983. — Vol. 54. — P. 6725–6731.
35. *Harame, D.* Si/SiGe Epitaxial-Base Transistors-Part1: Materials, Physics, and circuits / D. Harame [and other] // *IEEE Trans.* — 1995. — Vol. ED-42. — №3. — P. 455–468.
36. *Harame, D.* Si/SiGe Epitaxial-Base Transistors-Part2: Process Integration and Analog Applications / D. Harame [and other] // *IEEE Trans.* — 1995. — Vol. ED-42. — №3. — P. 469–482.
37. *Rieh, J. S.* Scaling of SiGe Heterojunction Bipolar Transistors. *Proc / J. S. Rieh* [and other] // *IEEE.* — 2005. — Vol. 93. — № 9. — P. 1522–1538.
38. *Khater, M.* SiGe HBT technology with $f_{\max}/f_T = 350/300$ GHz and gate delay below 3.3 ps. / M. Khater [and other] // *Dig. Int. Electron Devices Meeting, 2004.* — P. 247–250.
39. *Hafer, W.* Record f_T and $f_{\max}$ performance of InP/InGaAs single HBT / W. Hafer, J.W. Lai, M. Feng // *Electron. Lett.* — 2003. — Vol. 39. — № 10. — P. 811–813.
40. *Ahlgren, D. C.* A SiGe HBT BiSmos Technology for Mixed-Signal RF Applications / D. C. Ahlgren. — *Proc. of the IEEE Bipolar/BiCMOS Circuits and technology Meeting (BCTM), 1997.* — P. 195–197.
41. *Joseph, A.* A 0.18 μm BiCMOS Technology Featuring 120/100 GHz ($f_{\max}/f_T$) HBT and ASIC Compatible CMOS using Copper Interconnect / A. Joseph. — *Proc. of the IEEE Bipolar/BiCMOS Circuits and technology Meeting (BCTM), 2001.* — P. 143–146.
42. *Rucker H.* Integration of high-performance SiGe:C HBTs with thin-film SOI CMOS / H. Rucker [and other]. — *Dig. Int. Electron Devices Meeting 2004.* — P. 239–242.



Приложения

Приложение 1

Физические константы

№	Константа	Обозначение	Значение
1	Заряд электрона	e	$1,6 \cdot 10^{-19}$ Кл
2	Масса свободного электрона	m_0	$9,1 \cdot 10^{-31}$ кг
3	Постоянная Больцмана	k	$1,38 \cdot 10^{-23}$ Дж/К
4	Постоянная Планка	$h = 2\pi\hbar$	$6,62 \cdot 10^{-34}$ Дж с
5	Диэлектрическая проницаемость вакуума	ϵ_0	$8,86 \cdot 10^{-14}$ Ф/см
6	Магнитная проницаемость вакуума	μ_0	$1,26 \cdot 10^{-8}$ Гн/см
7	Волновое сопротивление вакуума	$\rho_0 = \sqrt{\mu_0 / \epsilon_0}$	377 Ом
8	Скорость света в вакууме	c	$3 \cdot 10^{10}$ см/с

Приложение 2

Список общих обозначений

ϵ — относительная диэлектрическая проницаемость полупроводника

ϵ_d — относительная диэлектрическая проницаемость диэлектрика

T — абсолютная температура

$N_{d,a}$ — концентрация донорной, акцепторной примеси

$N = |N_d - N_a|$ — результирующая концентрация примеси

n_i — собственная концентрация носителей заряда в полупроводнике

n, p — концентрации электронов, дырок



- n_0, p_0 — равновесные концентрации электронов, дырок
 $\Delta n = n - n_0, \Delta p = p - p_0$ — избыточные концентрации электронов, дырок
 E — энергия электрона
 E_0 — потенциальная энергия свободного электрона в вакууме
 $E_{c,v}$ — энергии дна зоны проводимости, потолка валентной зоны
 $E_{d,a}$ — энергии донорных, акцепторных уровней
 $\Delta E_{d,a} = |E_{c,v} - E_{d,a}|$ — энергии активации доноров, акцепторов
 E_t — энергия ловушечных уровней
 F — энергия Ферми
 $F_{n,p}$ — квазиуровни Ферми для электронов, дырок
 F_i — электростатическая энергия (энергия Ферми в собственном полупроводнике)
 E — напряженность электрического поля
 ϕ — электрический потенциал
 $\phi_T = kT / e$ — температурный потенциал
 V — напряжение
 I — ток
 j — плотность тока
 $\mu_{n,p}$ — подвижности электронов, дырок
 $D_{n,p}$ — коэффициенты диффузии электронов, дырок
 $\tau_{n,p}$ — время жизни электронов, дырок
 $L_{n,p} = \sqrt{D_{n,p} \tau_{n,p}}$ — диффузионные длины электронов, дырок
 L_D — дебаевская длина экранирования
 l — ширина области пространственного заряда (ОПЗ)



Приложение 3

Свойства некоторых полупроводниковых материалов ($T = 300 \text{ К}$)

№	Параметр, единица измерения	Полупроводник									
		A ⁴		A ³ B ⁵							
		Si	Ge	GaAs	AlAs	GaP	InP	InSb	InAs	GaN	
1	Атомный вес (средний)	[1,2] 28,09	[1,2] 72,59	[1–3] 72,32	[1, 2, 4] 50,95	[1, 2, 5] 50,35	[1, 2, 5] 72,9	[1, 2, 5] 128,3	[2, 5] 94,84	[2, 5] 41,9	
2	Плотность, г/см ³	2,33	5,33	5,32	3,81 [4] 3,76 [4]	4,13	4,79	5,78	5,68	6,15	
3	Тип кристаллической решетки	алмаза									
4	Постоянная решетки, Å	5,431	5,646	5,653	5,660	5,451	5,869	6,48	6,058	$a = 3,189;$ $c = 5,185$	
5	Удельная теплопроводность, Вт/см·К	1,24 [1] 1,5 [2]	0,64 [1] 0,6 [2]	0,56	0,84	0,752 [1] 1,1 [5]	0,800 [1] 0,68 [5]	0,16 [1] 0,18 [5]	0,27	Нет данных	
6	Относительная диэлектрическая постоянная	11,8 [1] 11,9 [2]	16,0	13,2 [1] 12,9 [2,3]	10,1 [2] 10,9	11,1	12,6 [2] 12,4	16,8 [2] 17,7	15,1	10,4 [2] 12,2 [5]	
7	Электронное сродство, эВ	4,05	4,0	4,07	3,5 [4]	3,8	5,34	4,59	4,9	4,1	
8	Ширина запрещенной зоны, эВ	1,12	0,66	1,43	2,36 [2] 2,17	2,26	1,35	0,17 [2] 0,18	0,36 [2] 0,354 [5]	3,44 [2] 3,36 [5]	
9	Эффективная плотность квантовых состояний, см ⁻³	2,8·10 ¹⁹	1,04·10 ¹⁹	4,7·10 ¹⁷ [2] 4,21·10 ¹⁷ [3]	1,5·10 ¹⁹	1,8·10 ¹⁹	5,7·10 ¹⁷	4,2·10 ¹⁶	8,7·10 ¹⁶	2,3·10 ¹⁸	
	Валентная зона	1,04·10 ¹⁹	6,0·10 ¹⁸	7,0·10 ¹⁷ [2] 9,51·10 ¹⁸ [3]	1,7·10 ¹⁹	1,9·10 ¹⁹	1,1·10 ¹⁹	7,3·10 ¹⁸	6,6·10 ¹⁸	4,6·10 ¹⁹	



Продолжение

№	Параметр, единица измерения	Полупроводник									
		A ⁴			A ³ B ⁵						
		Si [1,2]	Ge [1,2]	GaAs [1–3]	AlAs [1, 2, 4]	GaP [1, 2, 5]	InP [1, 2, 5]	InSb [1, 2, 5]	InAs [2, 5]	GaN [2, 5]	
10	Собственная концентрация носителей, см ^{–3}	1,5·10 ¹⁰	2,4·10 ¹³	≈2·10 ⁶	≈1·10 ^{–2}	≈2	1,2·10 ⁷	≈2·10 ¹⁶	≈8·10 ¹⁴	≈1·10 ^{–9}	
11	Эффективная масса	электрон, m _n /m ₀	1,64	0,063 [2]	1,1	0,82	0,073 [1] 0,077 [2]	0,0145	0,023	0,27 [2] 0,19 [5]	
				0,067	0,11 [2]						
		дырки, m _p /m ₀	0,082	0,5 [2]	0,22 [2]	0,57	0,58	0,42	0,41		
				0,53	0,76	0,60 [2]	0,64 [2]	0,40 [2]	0,40 [2]		
12	Подвижность, см ² /В·с	электрон	3800 [1] 3900 [2]	0,076 [2]	0,22 [2]	0,18	0,12	0,016	0,026	0,8 [2]	
				0,08	0,15						
				8800 [1] 8000 [2]	1200 [1] 180 [2] 200 [4]	300 [1] 110 [2]	4600	78 000 [1] 80 000 [2]	33 000 [2] 40 000 [5]		
13	Собственное удельное сопротивление, Ом·см	500	1820 [2] 1900 [2]	400	420 [1] 140 [4]	150 [1] 75 [2]	150	750 [1] 1250 [2]	450 [2] 500 [5]	10 [2]	
14	Время жизни носителей, с	2,3·10 ⁵	47	≈2·10 ⁸	1,6·10 ¹⁵	≈10 ¹⁶	8,6·10 ⁷	4·10 ^{–3}	0,19	≈10 ²⁵	
		2,5·10 ^{–3}	1·10 ^{–3}	1·10 ^{–8}	Нет данных						



Приложение 4

Энергии ионизации примесей в Si, Ge и GaAs

Примесь	Группа	Тип	Энергия ионизации, эВ			
			Si	Ge	GaAs	Ссылка
B	III	Акцептор	0,045	0,010	—	[1]
Al	III	Акцептор	0,067	0,011	—	[1]
Ga	III	Акцептор	0,071	0,011	—	[1]
In	III	Акцептор	0,155	0,012	—	[1]
P	V	Донор	0,045	0,013	—	[1]
As	V	Донор	0,054	0,014	—	[1]
Sb	V	Донор	0,043	0,010	—	[1]
Bi	V	Донор	0,071	0,013	—	[1]
Si	IV	Донор	—	—	0,002	[7]
Sn	IV	Донор	—	—	< 0,005	[7]
Se	VI	Донор	—	—	0,005	[6]
Te	VI	Донор	—	—	0,003	[6]
Ge	IV	Донор	—	—	< 0,005	[7]
Be	II	Акцептор	—	—	0,028	[6]
Mg	II	Акцептор	—	—	0,028 0,0125	[6] [7]
Zn	II	Акцептор	—	—	0,024	[7]
Cd	II	Акцептор	—	—	0,021	[7]

Приложение 5

Греческий алфавит

Α, α	Альфа	Η, η	Эта	Ν, ν	Ню	Τ, τ	Тау
Β, β	Бэта	Θ, θ(ϑ)	Тэта	Ξ, ξ	Кси	Φ, φ	Фи
Γ, γ	Гамма	Ι, ι	Иота	Ο, ο	Омикрон	Χ, χ	Хи
Δ, δ	Дельта	Κ, κ	Каппа	Π, π	Пи	Υ, υ	Ипсилон
Ε, ε	Эпсилон	Λ, λ	Ламбда	Ρ, ρ	Ро	Ψ, ψ(φ)	Пси
Ζ, ζ	Дзэта	Μ, μ	Мю	Σ, σ	Сигма	Ω, ω	Омега



Литература к приложениям 3, 4

1. *Пул, Ч.* Мир материалов и технологий. Нанотехнологии / Ч. Пул, Ф. Оуэнс ; пер. с англ. — М. : Техносфера, 2005. — 336 с.
2. Sze, S. M. Physics of Semiconductor Devices (3-rd edition) / S. M. Sze, K. Ng. Kwok. — Wiley-Interscience, 2007. — 815 p.
3. *Шур, М.* Современные приборы на основе арсенида галлия / М. Шур ; пер. с англ. — М. : Мир, 1991. — 632 с.
4. *Goldberg, Yu. A.* Handbook of semiconductor parameters / Yu. A. Goldberg. — Vol. 2. — Chapter 1: www.worldscibooks.com/phy-etextbook/2046/2046_chap1_1.pdf.
5. Ioffe Physico-Technical Institute. Electronic archive. New Semiconductor Materials. Characteristics and Properties: www.ioffe.rssi.ru/SVA/NSM.
6. Арсенид галлия в микроэлектронике : пер. с англ. ; под ред. Н. Айнспрука, У. Уиссмана. — М. : Мир, 1988. — 555 с.
7. Арсенид галлия. Получение, свойства и применение / под ред. Ф. П. Кесамаилы, Д. Н. Наследова. — М. : Наука, 1973. — 209 с.



Покупайте наши книги:

Оптом в офисе книготорга «Юрайт»:

1400004, Московская обл., г. Люберцы, 1-й Панковский проезд, д.1,
тел. (495) 744-00-12, e-mail: sales@urait.ru, www.urait.ru

В розницу в интернет-магазине: www.urait-book.ru,
e-mail: order@urait-book.ru, тел.: (495) 742-72-12

Для закупок у Единого поставщика в соответствии с Федеральным законом
от 21.07.2005 № 94-ФЗ обращайтесь по тел. (495) 744-00-12,
e-mail: sales@urait.ru, kea@urait.ru

Учебное издание

Виктор Игоревич Старосельский

Физика полупроводниковых приборов микроэлектроники

УЧЕБНОЕ ПОСОБИЕ

Редактор *Ярмахова Т.А.*

Корректор *Духовская Т. Н.*

Художественное оформление *А. И. Гиренко*

Компьютерная верстка *Дроздова В. В.*

Формат 84×108 ¹/₃₂.

Гарнитура «PetersburgC». Печать офсетная.

Усл. печ. л. 24,36. Доп. тираж 1000 экз. Заказ №

ООО «ИД Юрайт»

140004, Московская обл., г. Люберцы, 1-й Панковский проезд, д. 1.

Тел.: (495) 744-00-12. E-mail: izdat@urait.ru. www.urait.ru

