

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ «МИСиС»

А.Н. Ковалев

Транзисторы на основе полупроводниковых гетероструктур

Монография



Москва 2011

УДК 621.315.592
К56

Рецензент

д-р физ.-мат. наук, проф. *В.И. Трухин*
(МГУ им. М.В. Ломоносова)

Ковалев, А.Н.

К56 Транзисторы на основе полупроводниковых гетероструктур : моногр. / А.Н. Ковалев. – М. : Изд. Дом МИСиС, 2011. – 364 с. ISBN 978-5-87623-489-6

Рассмотрены физические принципы работы наиболее распространенных приборов полупроводниковой электроники – биполярных и полевых транзисторов – в их современном модернизированном исполнении на основе гетероструктурных композиций. На их примере дан анализ нового направления электроники – гетероструктурной наноэлектроники. Это часть более масштабного направления, называемого нанотехнологией, и охватывает она разработку полупроводниковых приборов и устройств субмикронных размеров. Приведены примеры реализации транзисторов на основе гетеросистем из материалов Ge/Si и $A^{III}B^V$. Установлены зависимости между размерами активных областей, составом материала и параметрами прибора. Выполнено сравнение и оценка возможностей материалов и приборов на их основе. Рассмотрены механизмы формирования и условия получения гетероэпитаксиальных структур с квантовыми точками в системах Ge/Si и InAs/GaAs с учетом элементов самоорганизации при эпитаксии. Интерес к самопорядоченным наноструктурам обусловлен созданием нанотранзисторов, а также фотоприемников и источников излучения в диапазоне длин волн 1,3...1,5 мкм.

Книга предназначена для широкого круга научных и инженерно-технических работников, специализирующихся в области разработки новых полупроводниковых микро- и наногетероструктур и приборов на их основе, а также для студентов и аспирантов.

УДК 621.315.592

ISBN 978-5-87623-489-6

© А.Н. Ковалев, 2011

ОГЛАВЛЕНИЕ

| | |
|---|-----|
| Введение | 6 |
| 1. Новое направление полупроводниковой электроники – гетероструктурная нанoeлектроника | 7 |
| 2. Энергетическая диаграмма идеального гетероперехода..... | 13 |
| 3. Биполярные гетеротранзисторы на $\text{Si}/\text{Ge}_x\text{Si}_{1-x}$ и $A^{III}B^V$ | 23 |
| 3.1. Биполярный транзистор..... | 24 |
| 3.1.1. Основные характеристики | 24 |
| 3.1.2. Эквивалентная схема транзистора. Предельные частоты | 28 |
| 3.1.3. Дрейфовый планарный биполярный транзистор | 32 |
| 3.2. Принцип действия биполярного гетеротранзистора | 35 |
| 3.3. Биполярные гетеротранзисторы на $\text{Si}/\text{Ge}_x\text{Si}_{1-x}$ | 40 |
| 3.3.1. Общая характеристика | 40 |
| 3.3.2. Углеродное ограничение профиля бора в базе SiGe-БГТ | 46 |
| 3.3.3. Механизм сегрегации бора при отжиге структур SiGeC/SiGeB/SiGeC | 52 |
| 3.4. Биполярные гетеротранзисторы на $A^{III}B^V$ | 58 |
| 3.4.1. Биполярные гетеротранзисторы на обычных материалах | 58 |
| 3.4.2. Биполярные гетеротранзисторы на основе нитридов III группы | 66 |
| 3.5. Сравнение достигнутых результатов по разработке БГТ с теоретическими оценками | 73 |
| 4. Современные полевые гетеротранзисторы на основе соединений $A^{III}B^V$ | 86 |
| 4.1. Основные представления и параметры | 86 |
| 4.1.1. Модель полевого транзистора | 87 |
| 4.1.2. Особенности переноса в субмикронных структурах | 93 |
| 4.2. Полевые транзисторы на $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ -гетероструктуре | 97 |
| 4.3. Псевдоморфные $\text{AlGaAs}/\text{InGaAs}$ -PHEMT на GaAs | 104 |
| 4.4. InP-HEMT и метаморфные GaAs-MHEMT | 117 |
| 5. Полевые гетеротранзисторы на материалах $A^{III}N$ | 133 |
| 5.1. Преимущества нитридных полупроводников | 134 |
| 5.2. Конструктивные особенности полевых транзисторов на основе GaN | 140 |

| | |
|---|-----|
| 5.3. Получение и свойства слоев гетероструктуры AlGaIn/GaN | 156 |
| 5.3.1. Получение слоев нитридных соединений МОС-гидридным методом..... | 157 |
| 5.3.2. Получение слоев гетероструктуры AlGaIn/GaN модифицированными методами МЛЭ | 164 |
| 5.3.3. Электропроводность слоев GaN | 167 |
| 5.4. Влияние поляризационных эффектов и поверхностных состояний на характеристики полевых GaN-гетеротранзисторов .. | 171 |
| 5.5. Коллапс тока и его связь с поверхностными состояниями | 181 |
| 5.6. Конструкции GaN-HEMT с полевой платой | 189 |
| 6. Основные направления и проблемы создания полевых транзисторов на AlGaIn/GaN-гетероструктурах | 199 |
| 6.1. Процессы латеральной эпитаксии GaN-HEMT | 199 |
| 6.2. Подложки для GaN-гетеротранзисторов | 205 |
| 6.3. Применение встроенных AlN(Al _x Ga _{1-x} N)-слоев | 212 |
| 6.4. Элементы постростовой технологии | 223 |
| 6.4.1. Травление нитридов III группы при изготовлении GaN-HEMT | 225 |
| 6.4.2. Контакты металл – нитриды III группы | 229 |
| 6.4.3. Ионная имплантация | 239 |
| 6.4.4. Применение диэлектрических слоев | 246 |
| 7. Современные методы усовершенствования полевых AlGaIn/GaN-гетеротранзисторов | 251 |
| 7.1. Основные направления разработки полевых гетеротранзисторов | 251 |
| 7.2. Последние разработки и достигнутые результаты | 254 |
| 7.2.1. Учет поляризационных эффектов при создании барьерного слоя | 254 |
| 7.2.2. Способы повышения предельной частоты | 260 |
| 7.2.3. Разработка GaN-HEMT для переключателей мощности ... | 271 |
| 7.2.4. Изготовление нормально закрытого E-GaN-HEMT | 274 |
| 7.2.5. Расширение возможных применений GaN-HEMT-структур | 277 |
| 7.3. Внедрение In-содержащих слоев в структуру GaN-HEMT | 279 |
| 7.3.1. Использование In в барьерном слое полевого AlGaIn/GaN-гетеротранзистора | 279 |
| 7.3.2. Создание проводящего канала в полевом AlGaIn/GaN-гетеротранзисторе с применением слоя InGaIn | 288 |
| 7.3.3. Особенности выращивания InGaIn-слоев на GaN | 294 |
| 7.4. От исследований к производству приборов | 302 |

| | |
|---|-----|
| 8. Квантоворазмерные структуры и их применение | 306 |
| 8.1. Механизмы формирования гетероэпитаксиальных структур с квантовыми точками | 307 |
| 8.1.1. Самоорганизация при эпитаксии | 309 |
| 8.1.2. Теоретические представления о достижении равновесного состояния в системе гетеронаноостровков | 315 |
| 8.2. Квантовые точки Ge/Si | 321 |
| 8.3. Фотоприемники на основе квантоворазмерных структур | 333 |
| 8.4. Квантовые точки в системе InAs/GaAs | 339 |
| 8.5. Полупроводниковые нанотрубки | 346 |
| Заключение | 351 |
| Библиографический список | 353 |

ВВЕДЕНИЕ

Многokратное увеличение уровня интеграции, быстрогодействия, расширения функциональных возможностей в электронике, оптике, информатике и других областях науки и техники возможно на основе нанотехнологии. Нанотехнология способна изменить все аспекты человеческого существования. Возможность синтезировать наномасштабные элементы, а затем собирать такие элементы в более крупные структуры, обладающие уникальными свойствами и функциями, приведет к революционным изменениям во многих отраслях. Изменение характеристик и появление новых свойств нанобъектов обусловлены не только уменьшением размеров элементов, но и проявлением квантовомеханических эффектов, волновой природой процессов переноса и доминирующей ролью поверхностей раздела.

Полупроводниковая гетероструктурная наноэлектроника представляет собой составную часть нанотехнологии и охватывает прежде всего разработку полупроводниковых приборов и устройств с размерами субмикронного диапазона, приближающегося к нанометровому диапазону, т.е. к объектам меньше 100 нм.

Рассмотрены области применения и перспективы данного направления электроники. Приведены примеры реализации биполярных и полевых транзисторов на основе гетеросистем из материалов Ge/Si, $A^{III}B^V$, размеры активных областей которых менее 100 нм.

Значительная часть работы посвящена механизмам формирования и условиям получения гетероэпитаксиальных структур с квантовыми точками в системах Ge/Si и InAs/GaAs. Рассмотрены элементы самоорганизации при эпитаксии и применение явлений самоорганизации для эпитаксиального наращивания наногетероструктурных объектов полупроводниковой электроники на основе массива квантовых точек.

1. НОВОЕ НАПРАВЛЕНИЕ ПОЛУПРОВОДНИКОВОЙ ЭЛЕКТРОНИКИ – ГЕТЕРОСТРУКТУРНАЯ НАНОЭЛЕКТРОНИКА

Вторая половина прошедшего столетия ознаменована грандиозными достижениями в развитии полупроводниковой электроники и, прежде всего, микро- и оптоэлектроники. Эти достижения обеспечили невиданный прогресс в вычислительной технике, информатике, радиоэлектронике, энергетике и в других передовых областях науки и техники. Совершенно очевидно, что все эти преобразования были бы вряд ли возможны без выдающихся достижений в развитии материаловедения и технологии полупроводниковых материалов.

В XXI век микроэлектроника вошла с производством УСБИС динамической памяти на 1 Гбит и микропроцессоров с тактовыми частотами до 1,2 ГГц. К 2012 г. предполагается довести эти показатели до 64 Гбит и 10 ГГц соответственно. Если сегодня размер единичного транзистора в УСБИС составляет 0,18...0,13 мкм, к 2012 г. эти значения уменьшатся до 0,03...0,035 мкм. Аналогичная тенденция к резкому повышению степени микроминиатюризации наблюдается и в оптоэлектронике. Уже сегодня размеры активных областей инжекционных лазеров на квантовых ямах вышли на нанометровый уровень, а впереди – активное освоение производства лазеров на основе квантовых точек. Освоение размерного диапазона 1,0...0,1 мкм представляет собой весьма сложную технологическую задачу, но происходит оно с использованием традиционной элементной базы. Совершенно другая ситуация складывается при освоении диапазона линейных размеров менее 0,1 мкм. Здесь возникает фундаментальный физический барьер, обусловленный резкими изменениями практически всех свойств твердого тела, в том числе и электропроводности. При достижении таких размеров в соответствующих объектах начинают в полной мере проявляться квантовые эффекты, что требует совершенно иного подхода к конструированию приборов, которые должны работать на новых физических принципах. Вот почему освоение нанометрового диапазона размеров в современной твердотельной электронике выделено в специальное направление, названное нанoeлектроникой.

Нанoeлектроника является областью науки и техники, сформировавшейся на основе достижений физики твердого тела, квантовой

электроники, физической химии и технологии полупроводниковой микроэлектроники. Разрабатываемые для нанoeлектроники технологии должны быть ориентированы на массовое производство приборов и интегральных схем с минимальными размерами элементов в диапазоне от 100 до 1 нм.

В нанoeлектронике используются следующие основные квантовые эффекты, лежащие в основе функционирования наноразмерных элементов: интерференция, квантовое ограничение, туннелирование через потенциальные барьеры.

Свободному электрону в твердом теле соответствует электромагнитная волна, способная распространяться в любом направлении. Однако поведение электрона изменяется, если он находится в области твердого тела, ограниченной потенциальными барьерами, примером которой может являться квантовый шнур с ограниченными размерами структуры. При этом соответствующие электроны могут иметь только определенные фиксированные значения энергии, тогда как вдоль шнура могут двигаться электроны с любой энергией. Запирание электрона хотя бы в одном из направлений сопровождается увеличением его импульса. Данное явление называется квантовым ограничением и приводит, с одной стороны, к увеличению минимальной энергии электрона, а с другой, – к дополнительному квантованию энергетических уровней, вследствие чего свойства наноразмерных структур отличаются от свойств материала, из которого они сформированы.

На туннелирование электронов в наноразмерных структурах существенное влияние оказывает квантовое ограничение. Квантование энергетических состояний в тонких периодически расположенных ямах приводит к резонансному характеру туннелирования.

Интенсивное исследование квантовых эффектов в сверхтонких полупроводниковых гетероструктурах привело к появлению полупроводниковых приборов новых классов – резонансных туннельных диодов и транзисторов, обладающих потенциально очень высоким быстродействием (предельные частоты до 10^{12} Гц) и широким спектром других возможностей, а также инжекционных лазеров на квантовых ямах и квантовых точках с уникальными рабочими характеристиками.

Объединение в самые последние годы физики и технологии наногетероструктур с нанотехнологией изготовления приборов на их основе привело к созданию наногетероструктурной электроники.

Элементная база наногетероструктурных монолитных интегральных схем включает:

– гетеробиполярные транзисторы, т.е. НВТ (Heterojunction Bipolar Transistor);

– униполярные полевые транзисторы с электронным газом высокой подвижности и высокой плотности, или НЕМТ (High Electron Mobility Transistor);

– туннельно-резонансные приборы на квантовых ямах, обладающие, наряду с рекордным быстродействием, уникальными функциональными возможностями;

– гетероструктурные транзисторы на квантовых точках (т.е. на квазиульмерных электронных системах);

– фотодетекторы и лазеры на квантовых точках.

Активно обсуждаются проблемы создания квантовых интегральных схем, основными элементами которых должны стать квантовые точки, квантовые проводники, квантовые ямы, транзисторные структуры на основе квантоворазмерных эффектов и устройств с управляемой интерференцией электронов.

Все перечисленное стало возможным лишь благодаря выдающимся достижениям в развитии технологии молекулярно-пучковой и МОС-гидридной эпитаксии, обеспечившим возможность синтеза высококачественных квантоворазмерных композиций широкого круга полупроводниковых материалов. Однако по существу развитие технологии и материаловедения наноструктур лишь только начинается. С материаловедческих позиций наноструктуры являются весьма специфическими объектами, свойства которых в значительной степени определяются свойствами их поверхности и явлениями на границах раздела фаз. Все это определяет специфику межфазных взаимодействий и особенностей поведения примесей и структурных дефектов в наноразмерных многофазных композициях. Ключ к получению недеградирующих наноструктур с контролируемыми свойствами лежит в детальном исследовании всех этих явлений.

Решающую роль в формировании наноконпозиций типа сверхрешеток, состоящих из квантовых ям, проволок или точек, играют эффекты самоорганизации. Однако для того чтобы обеспечить, например, получение композиций с однородным распределением необходимого количества квантовых точек контролируемого размера, надо иметь четкое представление о механизме явлений, лежащих в основе самопроизвольного возникновения макроскопического порядка в первоначально однородной системе, т.е. выявить основные движущие силы самоорганизации. Для наноструктур различных типов причины

неустойчивости однородного состояния системы могут существенно различаться, и в каждом конкретном случае в этом надо детально разбираться. Только такого рода подходы позволяют с наибольшим эффектом реализовать возможности процессов самоорганизации. В последние годы исследования в этом направлении развиваются весьма успешно. Развитие технологии и материаловедения полупроводниковых наноструктур вряд ли возможно без глубокого проникновения в природу явлений, происходящих в традиционных полупроводниковых средах на атомном (молекулярном) уровне. Это, в свою очередь, требует разработки новых нестандартных методов исследования с использованием сканирующей атомно-силовой и туннельной микроскопии, электронной микроскопии высокого разрешения, рентгеновской спектроскопии и ряда других современных подходов.

Развитие нанотехнологии требует применения достижений современной электронной нанолитографии для формирования активных элементов малых размеров.

Развитие методов литографии шло по пути уменьшения длины используемого излучения (ультрафиолетовое или синхротронное) и применения частиц с меньшей длиной волны (высокоэнергетичная электронная и ионная литография). Существенным ограничением возможности применения методов оптической литографии является необходимость фокусировки света. От этого недостатка свободны методы ионной и электронной литографии.

Развитие нанотехнологий направлено на создание электронных устройств будущего, способных стать базой для дальнейшего развития информационных технологий. Увеличение мощностей компьютерных систем достигается уменьшением размеров применяющихся в настоящее время кремниевых интегральных схем. Однако, по оценке ведущих международных исследовательских организаций, при сохранении современных темпов развития интегральных технологий через 10–15 лет будет достигнут предел микроминиатюризации. Дальнейшее повышение вычислительных мощностей компьютерных систем будет неразрывно связано с применением нанотехнологий.

Созданный в мире за предыдущие десятилетия научный и технологический задел дал мощный импульс быстрому промышленному освоению этой новейшей технологии. Объем мировых ежегодных продаж только самих гетероструктурных транзисторов и монокристаллических интегральных схем уже приблизился к 6 млрд долл. США, ежегодно увеличиваясь примерно на 3%. По всем прогнозам, рынок

этих изделий в ближайшие 12–15 лет будет оставаться наиболее привлекательным для инвестиций.

К настоящему времени налажено массовое производство и реализация широкого спектра гетероструктур, гетероструктурных транзисторов и монолитных интегральных схем (МИС), характеризующихся следующими особенностями:

- наиболее массовой областью применения этой технологии остается сотовая связь, сейчас это цифровая связь 3-го поколения; она занимает около 57 % «гетероструктурного» рынка; гетеротранзисторы содержатся почти в каждом выпускаемом в мире сотовом телефоне;

- около 25 % рынка занимает быстро прогрессирующая (наиболее высокоскоростная) волоконно-оптическая связь, используются гетероструктурные МИС, рассчитанные на частоты от 60 ГГц;

- около 12 % рынка принадлежат так называемой потребительской электронике, связанной с цифровым ТВ (частоты от 12 до 40 ГГц);

- наблюдается быстрый рост рынка гетероструктурных МИС (77 ГГц) для автомобильных антистолкновительных радаров;

- быстро развивается и рынок СВЧ МИС для спутниковой связи на частоты до 60 ГГц;

- перспективной областью применения гетероструктурной СВЧ электроники является беспроводная связь широкополосного доступа для работы на частотах 40...60 ГГц и выше с предоставлением услуг мультимедиа (передача сложных изображений, видеоконференций, дистанционное обучение, высокоскоростной интернет и т.д.);

- более 4 % рынка принадлежит военной электронике; это в основном бортовые и мобильные наземные радиолокаторы на активных фазированных антенных решетках (АФАР) X-диапазона (частота около 10 ГГц) для ведения разведки, обнаружения целей, наведения и сопровождения ракет, а также для радиоэлектронной борьбы.

Там, где требуются рабочие частоты выше 5 ГГц, наногетероструктурная технология быстро вытесняет не только кремниевую технологию, но и более высокоскоростную арсенид-галлиевую MESFET-технологию, захватывая все большую долю СВЧ телекоммуникационного и радиолокационного мирового рынка.

Главными тенденциями развития гетероструктурной электроники в соответствии с растущими требованиями рынка являются:

- стремительное продвижение в более высокие частоты, т.е. в миллиметровый и субмиллиметровый диапазоны; успехи, достигнутые в освоении частот до 300 ГГц, базируются в значительной мере на технологии изоморфных, псевдоморфных и метаморфных гетероструктур

InAlAs/InGaAs/InAlAs в сочетании с нанометровой (20...50 нм) технологией изготовления приборов;

– тенденция к созданию мощных высокочастотных приборов (для передающих устройств) со все более высокой выходной мощностью $P_{\text{вых}}$ и КПД; наилучшие перспективы связываются с «широкозонной» гетеросистемой Al(In)GaN/Ga(In)N, где за счет большой ширины запрещенной зоны достигается удельная мощность, на порядок большая, чем в GaAs;

– особое и весьма существенное место в наногетероструктурной электронике и оптоэлектронике занимают гетероструктуры SiGe/Si, твердые растворы SiGe, напряженные сверхрешетки SiGe/Si, на основе которых реализуют полевые и биполярные наногетеротранзисторы; на квантовых точках Ge/Si прорабатывают создание фотоприемников и излучателей на длину волны 1,5 мкм, которая соответствует окну максимальной прозрачности оптических волноводов из кварцевого стекла;

– требования рынка по расширению функциональных возможностей СВЧ приборов, уменьшению их массогабаритов, увеличению надежности и снижению стоимости стимулируют непрерывное увеличение степени интеграции наногетероструктурных МИС, т.е. создание многофункциональных МИС, объединяющих на одном чипе несколько различных по своим функциям приборов.

Достижения в разработке и изготовлении наноструктур различного назначения определяются уровнем развития нанотехнологий, которые позволяют получать их с атомной точностью и необходимого химического состава.

Нанотехнологии призваны решить следующие основные задачи современной полупроводниковой электроники:

– повышение производительности вычислительных систем и, в перспективе, создание квантового компьютера;

– увеличение пропускной способности каналов связи;

– увеличение информационной емкости и качества систем отображения информации с одновременным снижением энергозатрат;

– расширение возможностей сенсорных и энергосберегающих устройств;

– увеличение доли использования электронных и оптоэлектронных компонентов в биологических, медицинских, химических, машиностроительных и других технологиях.

2. ЭНЕРГЕТИЧЕСКАЯ ДИАГРАММА ИДЕАЛЬНОГО ГЕТЕРОПЕРЕХОДА

В отличие от p - n перехода, образованного в результате изменения типа и концентрации примесей в одном полупроводниковом материале (гомопереход), гетеропереходом называют переход, образованный полупроводниками различной физико-химической природы или состава. Примерами гетеропереходов являются Si/Ge, Ge/GaAs, GaAs/GaP, GaAs/Al_xGa_{1-x}As и др. Материалы с двух сторон гетероперехода обладают различными электрофизическими характеристиками.

Для получения гетеропереходов с минимальным количеством дефектов на границе раздела кристаллическая решетка одного полупроводника должна с минимальными нарушениями переходить в кристаллическую решетку другого. В связи с этим контактирующие полупроводники должны иметь близкие значения постоянной решетки, идентичные кристаллические структуры, близкие температурные коэффициенты расширения. Для построения зонных диаграмм, детального анализа распределения электрического поля и потенциала в области пространственного заряда гетероперехода, а также величины и компонент электрического тока для гетеропереходов, необходимо учитывать, что у различных полупроводников будут отличаться значения электронного сродства χ , ширины запрещенной зоны E_g и диэлектрической проницаемости ϵ_s . В табл. 2.1 приведены основные характеристики различных полупроводников.

Таблица 2.1

Ширина запрещенной зоны E_g , электронное сродство χ и параметр кристаллической решетки a для различных полупроводников

| Полупроводник | E_g , эВ | χ , эВ | a , нм |
|---------------|------------|-------------|----------|
| Ge | 0,66 | 4,0 | 0,5646 |
| Si | 1,12 | 4,05 | 0,5431 |
| GaAs | 1,42 | 4,07 | 0,5658 |
| GaP | 2,26 | 4,3 | 0,5451 |
| InAs | 0,36 | 4,9 | 0,6058 |
| InP | 1,35 | 5,34 | 0,5869 |
| GaN | 3,44 | 4,2 | 0,5185 |
| AlN | 6,28 | – | 0,4982 |
| InN | 1,89 | – | 0,5705 |

На рис. 2.1 показана энергетическая диаграмма идеального гетероперехода.

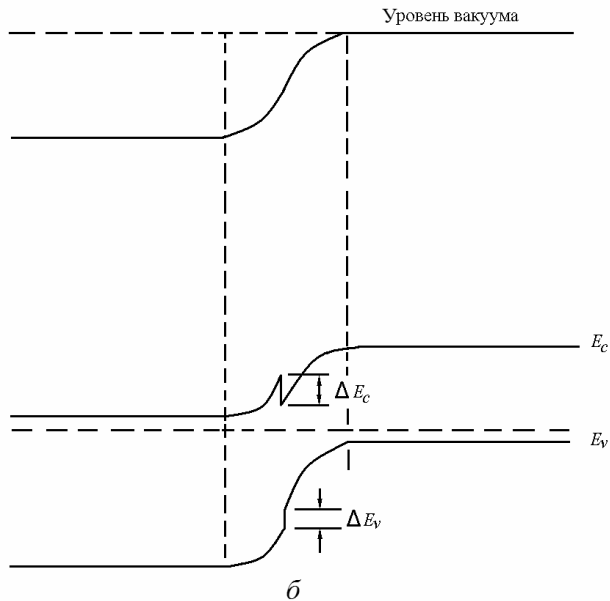
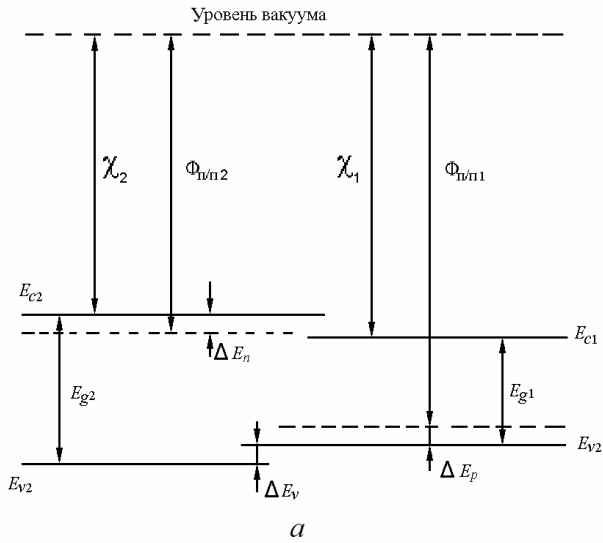


Рис. 2.1. Зонная диаграмма двух различных полупроводников и гетероперехода:
a – зонная диаграмма на примере AlGaAs *n*-типа и GaAs *p*-типа;
б – зонная диаграмма *p-n* гетероперехода при нулевом смещении

Здесь представлены зонные диаграммы двух разделенных в пространстве полупроводников разного состава: широкозонного полупроводника n -типа и узкозонного полупроводника p -типа (рис. 2.1, *a*).

Уровень вакуума выбран в качестве точки отсчета энергии. Величина $\chi_{1,2}$ – энергия, необходимая для переноса электрона со дна зоны проводимости в вакуум, называемая электронным сродством; работа выхода $\Phi_{n/n1,2}$ – разность энергий уровня Ферми и вакуума (индексы 1 и 2 относятся к узко- и широкозонным материалам соответственно). Когда два полупроводника соединяются, узкозонный полупроводник обогащается электронами, а широкозонный обедняется. Уровни Ферми выравниваются. Изгиб уровня электрона в вакууме равен разности работ выхода из полупроводников и характеризуется контактной разностью потенциалов $q\Phi_k = \Phi_k = \Phi_{n/n2} - \Phi_{n/n1}$. Энергетические зоны узкозонного полупроводника у границы раздела изгибаются вниз на величину Φ_1 , а широкозонного – вверх на величину Φ_2 (рис. 2.1, *b*).

Отличия данной зонной диаграммы от энергетической диаграммы p - n перехода заключаются в возникновении разрывов краев зоны проводимости ΔE_c и в валентной зоне ΔE_v , а также в образовании пика и ямы на границе раздела полупроводников.

Величина разрыва края зоны проводимости обусловлена разностью значений электронного сродства:

$$\Delta E_c = \chi_1 - \chi_2.$$

Разрыв в энергии края валентной зоны составляет

$$\Delta E_v = \Delta E_g - \Delta E_c,$$

где ΔE_g – разность значений ширины запрещенных зон.

Потенциальные барьеры для электронов и дырок оказываются различными:

– высота барьера для электронов и дырок выражается соответственно:

$$\Phi_n = \Phi_1 + \Phi_2 - \Delta E_c;$$

$$\Phi_p = \Phi_1 + \Phi_2 + \Delta E_v;$$

– разница высоты барьеров для дырок и электронов составляет

$$\Delta\Phi = \Phi_p - \Phi_n = \Delta E_v + \Delta E_c = E_{g2} - E_{g1},$$

т.е. равна разности значений ширины запрещенных зон контактирующих полупроводников.

Рассмотрим зонную диаграмму гетероперехода из материалов германия и арсенида галлия $n\text{Ge}/p\text{GaAs}$ (рис. 2.2). Используем те же принципы при построении этой зонной диаграммы. В этом случае разрыв наблюдается в энергетическом положении дна зоны проводимости и величина этого разрыва $\Delta E_c = \chi_{\text{Ge}} - \chi_{\text{GaAs}}$.

В области металлургической границы для энергии края валентной зоны E_v наблюдается пик. Высота пика ΔE_v составляет

$$\Delta E_v = \chi_{\text{GaAs}} + E_{g \text{ GaAs}} - \chi_{\text{Ge}} - E_{g \text{ Ge}} = (E_{g \text{ GaAs}} - E_{g \text{ Ge}}) - \Delta E_c.$$

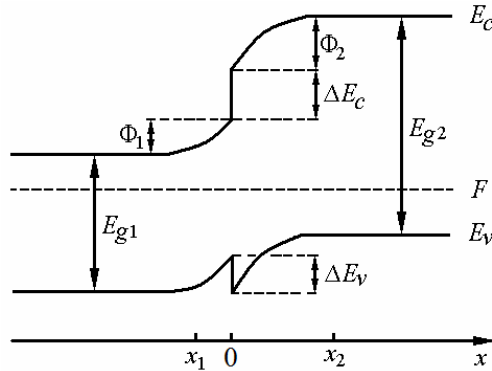


Рис. 2.2. Зонная диаграмма гетероперехода $n\text{Ge}/p\text{GaAs}$ в равновесных условиях

На рис. 2.3 приведена зонная диаграмма гетероперехода $p\text{Ge}/n\text{GaAs}$.

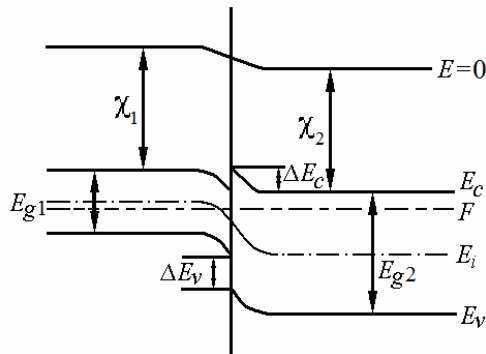


Рис. 2.3. Зонная диаграмма гетероперехода $p\text{Ge}/n\text{GaAs}$ в равновесных условиях

При построении зонной диаграммы такого гетероперехода на металлургической границе образуются пик $\Delta E_c = \chi_{\text{Ge}} - \chi_{\text{GaAs}}$ и разрыв валентной зоны $\Delta E_v = -\chi_{\text{Ge}} - E_{g \text{ Ge}} + \chi_{\text{GaAs}} + E_{g \text{ GaAs}} = -\Delta E_c + (E_{g \text{ GaAs}} - E_{g \text{ Ge}})$.

Суммарная высота пика ΔE_c и разрыва ΔE_v составляет $\Delta E_c + \Delta E_v = (E_{g \text{ GaAs}} - E_{g \text{ Ge}})$.

Значения параметров полупроводниковых материалов, выбранных для расчета этой зонной диаграммы, приведены в табл. 2.2.

Таблица 2.2

Значения параметров полупроводниковых материалов

| Параметр | Обозначение параметра | Германий (pGe) | Арсенид галлия (nGaAs) |
|---|-----------------------|-------------------|------------------------|
| Концентрация легирующей примеси, см ⁻³ | $N_{A,D}$ | $3 \cdot 10^{16}$ | 10^{16} |
| Расстояние от уровня Ферми до зоны разрешенных энергий, эВ | W_0 | 0,12 | 0,17 |
| Расстояние от уровня Ферми до середины запрещенной зоны, эВ | ϕ_0 | 0,21 | 0,54 |
| Электронное сродство, эВ | χ | 4,00 | 4,07 |

Аналогичным образом можно построить зонные диаграммы для гетеропереходов при любых соотношениях ширины запрещенной зоны и электронного сродства, типа и степени легирования. На рис. 2.4 приведены соответствующие зонные диаграммы для гетеропереходов различных типов. При этом пик и разрыв энергетических зон E_v , E_c в области металлургической границы могут наблюдаться в различных комбинациях.

Распределение потенциалов, поля и ширины обедненного слоя в каждом полупроводнике находят с помощью решения уравнения Пуассона для резкого p - n перехода с обеих сторон от границы раздела. Одним из граничных условий при этом является непрерывность электрической индукции на границе раздела, т.е. $\epsilon_1 E_1 = \epsilon_2 E_2$, где ϵ_1 и ϵ_2 – диэлектрические проницаемости контактирующих полупроводников; E_1 и E_2 – максимальные значения электрических полей на границе раздела.

На рис. 2.5 показаны распределения электрического поля и потенциала в области пространственного заряда гетероперехода $n\text{Ge}/p\text{GaAs}$.

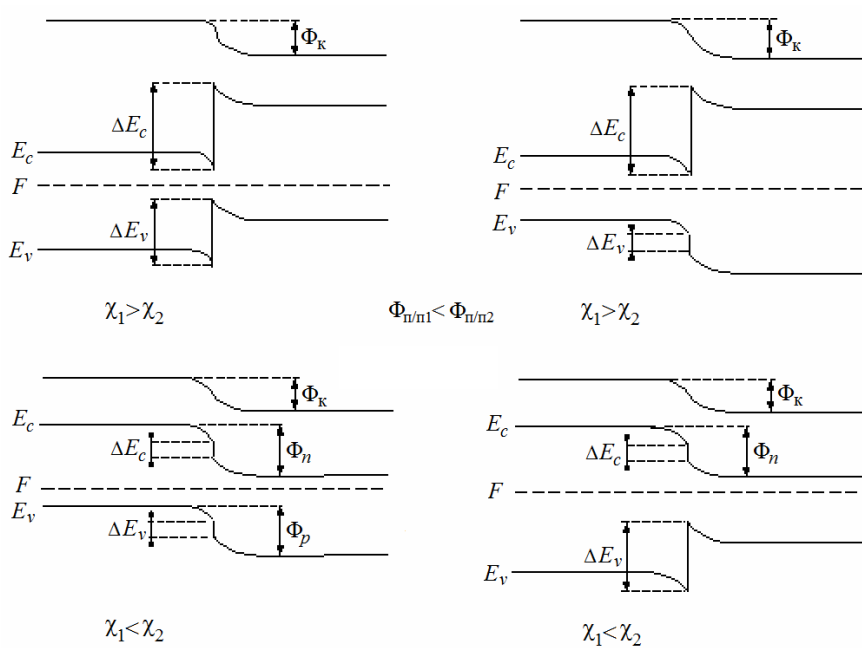


Рис. 2.4. Зонные диаграммы для гетеропереходов различных типов при условии, что термодинамическая работа выхода первого слоя меньше, чем второго ($\Phi_{n/n1} < \Phi_{n/n2}$), и при различных соотношениях электронного свойства

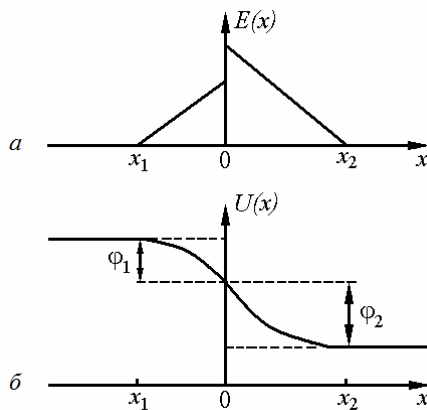


Рис. 2.5. Распределение электрического поля (а) и потенциала (б) в области пространственного заряда

Функциональная зависимость электрического поля и потенциала в области пространственного заряда гетероперехода от координаты будет соответственно линейной и квадратичной, как и в случае резкого p - n перехода. Скачок электрического поля в гетеропереходе на металлургической границе обусловлен различными значениями диэлектрических постоянных ϵ_1 и ϵ_2 .

Контактную разность потенциалов ϕ_k определяют по разности уровней Ферми в полупроводниках до их соприкосновения.

Полагают, что внешнее смещение формирует области объемного заряда аналогично контактному потенциалу в равновесии. Следовательно, внешнее смещение $U = U_1 + U_2$ распределяется между частями гетероперехода в той же пропорции, что и контактные потенциалы:

$$\frac{U_1}{U_2} = \frac{\phi_1}{\phi_2}.$$

В результате получаем следующие выражения для размеров переходных областей:

$$x_1 = \left[\frac{2\epsilon_1\epsilon_2\epsilon_0 N_A (\phi_k - U)}{qN_D(\epsilon_1 N_D + \epsilon_2 N_A)} \right]^{1/2};$$

$$x_2 = \left[\frac{2\epsilon_1\epsilon_2\epsilon_0 N_D (\phi_k - U)}{qN_A(\epsilon_1 N_D + \epsilon_2 N_A)} \right]^{1/2}.$$

Здесь N_A и N_D – соответственно концентрация акцепторов и доноров в p - и n -полупроводниках.

Барьерная емкость гетероперехода вычисляется по формуле для плоского конденсатора с двухслойным диэлектриком:

$$C^{-1} = \frac{1}{\epsilon_0} \left(\frac{x_1}{\epsilon_1} + \frac{x_2}{\epsilon_2} \right)$$

или

$$C = \left[\frac{\epsilon_1\epsilon_2\epsilon_0 N_A N_D}{2(\epsilon_1 N_D + \epsilon_2 N_A)(\phi_k - U)} \right]^{1/2}.$$

На вольт-амперные характеристики гетеропереходов влияют различные факторы, определяемые разрывами зон на границе раздела и плотностью граничных состояний. Механизм токопереноса определяется прежде всего конкретной формой потенциального рельефа гетероструктуры. Так, например, если пик ΔE_c на рис. 2.1, б расположен ниже E_c , то в прямом направлении инжекция электронов происходит с помощью диффузионного механизма, характерного для гомоперехода. Если же пик расположен выше, то действует эмиссионный механизм, как в выпрямляющем контакте металл – полупроводник (диод Шоттки). Следует учитывать такие явления, как туннельный эффект в области пика и через запрещенную зону, рекомбинация в промежуточном слое границы раздела.

В общем виде при смещении в прямом направлении зависимость $I = f(U)$ допускает аппроксимацию $I \sim \exp(qU/nkT)$, обратный ток не имеет насыщения.

В отличие от гомопереходов изотипные гетеропереходы ($n-n$, $p-p$) являются выпрямляющими, использующими основные носители.

В приведенном на рис. 2.1, б анизотипном $p-n$ гетеропереходе потенциальный барьер для электронов в зоне проводимости меньше, чем для дырок в валентной зоне. При подаче напряжения в прямом направлении потенциальный барьер для электронов уменьшается и электроны из n -полупроводника инжектируются в p -полупроводник. Потенциальный барьер для дырок в p -области также уменьшается, но все же остается достаточно большим для того, чтобы инжекции дырок из p -области в n -область практически не было. Эта особенность гетеропереходов – способность создавать высокий уровень инжекции носителей одного знака – определяет одно из их главных преимуществ перед гомопереходами. Эту способность гетеропереходов называют суперинжекцией.

Для оценки инжекционных свойств $p-n$ перехода обычно пользуются коэффициентом инжекции γ , определяемым отношением тока из сильнолегированной области к общему току. Для n^+-p перехода

$$\gamma = I_n / (I_n + I_p) \approx 1 - I_p / I_n \approx 1 - \sigma_p / \sigma_n,$$

где σ_p , σ_n – проводимость p - и n -областей соответственно.

Минимум I_p / I_n достигается созданием в гомопереходе сильнолегированной n^+ -области, т.е. большой разницей концентраций легирующих примесей, доноров и акцепторов: $N_D \gg N_A$. Этот способ имеет ограничения, поскольку существует предел растворимости

примеси, кроме того, при сильном легировании возникает много дефектов, ухудшающих параметры p - n перехода.

Для гетеропереходов

$$I_p / I_n \approx (n_{i2}/n_{i1})^2 N_A / N_D \exp(-\Delta\Phi/kT),$$

где n_{i1} , n_{i2} – концентрация носителей заряда в собственных полупроводниках.

Разность высот барьеров $\Delta\Phi \cong 0,2 \dots 0,5$ эВ.

Так как при 300 К $kT = 0,026$ эВ, то $\exp(-\Delta\Phi/kT) \sim 10^{-3} \dots 10^{-8}$, т.е. ток через гетеропереход состоит только из носителей одного знака и не зависит от соотношений концентраций носителей заряда в собственном полупроводнике, $\gamma = 1$.

Для гетеропереходов типа $p\text{Ge}/n\text{GaAs}$ легко реализовать одностороннюю инжекцию, даже в случае одинакового уровня легирования в эмиттере $p\text{Ge}$ и базе $n\text{GaAs}$ гетероперехода. Действительно, при прямом смещении отношение дырочной J_p и электронной J_n компонент инжекционного тока будет определяться отношением концентраций неосновных носителей:

$$\frac{J_p}{J_n} = \frac{qL_p p_n / \tau_p}{qL_n n_p / \tau_n} \approx \frac{p_n}{n_p} = \frac{n_{i2}^2 / N_D}{n_{i1}^2 / N_A} = \frac{n_{i2}^2}{n_{i1}^2},$$

где L_p , L_n , τ_p , τ_n – соответственно диффузная длина и время жизни неосновных носителей.

Поскольку арсенид галлия – более широкозонный полупроводник, чем германий, то собственная концентрация в арсениде галлия (n_{i2}) будет много меньше, чем в германии (n_{i1}), следовательно, дырочная компонента J_p инжекционного тока будет много меньше, чем электронная компонента J_n . Весь инжекционный ток в гетеропереходе $p\text{Ge}/n\text{GaAs}$ будет определяться электронной компонентой.

Односторонняя инжекция носителей заряда в гетеропереходе широко используется при создании высокоэффективных эмиттеров в биполярных гетеротранзисторах и инжекционных лазерах.

Важным является применение гетеропереходов для изготовления полевых гетеропереходных транзисторов на материалах $A^{\text{III}}B^{\text{V}}$. В этих приборах в качестве проводящего канала используют продольную энергетическую яму, образующуюся на гетерогранице со стороны узкозонного материала в n - n гетеропереходе (рис. 2.6).

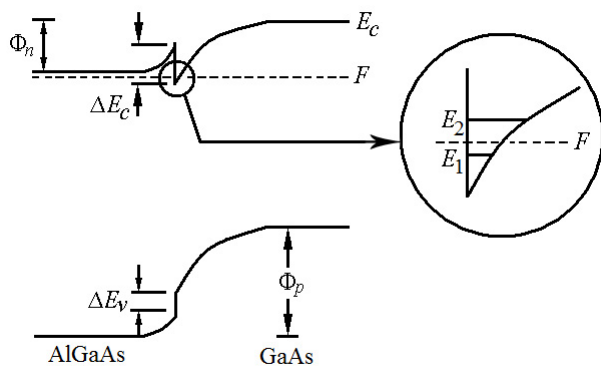


Рис. 2.6. Зонная диаграмма гетероперехода, иллюстрирующая присутствие квантовой ямы

Ширина этой ямы соизмерима с длиной волны де Бройля (10^{-6} см), в такой яме энергетический спектр электронов при движении в поперечном направлении квантуется, т.е. электроны могут занимать только определенные энергетические уровни, соответствующие целому числу полувольт. Перемещение электронов в канале оказывается возможным только в плоскости гетерограницы (по осям Y и Z). Такой электронный газ получил название двумерного – $2D$ (Dimension). Рассеяние $2D$ -электронов на границе раздела в результате рассогласования решеток будет незначительным. При обеспечении слабого уровня легирования узкозонного полупроводника рассеяние на ионизованных примесях сводят к минимуму, в результате подвижность электронов вдоль $2D$ -канала может достигать рекордных значений (10^5 см²/(В · с) при 300 К).

3. БИПОЛЯРНЫЕ ГЕТЕРОТРАНЗИСТОРЫ НА $\text{Si/Ge}_x\text{Si}_{1-x}$ И $A^{\text{III}}B^{\text{V}}$

В данном разделе приведены основные характеристики обычных гомобиполярных транзисторов, а также рассмотрен принцип действия и дан анализ состояния разработок биполярных гетеропереходных транзисторов (БГТ) на основе твердых растворов AlGaAs/GaAs , SiGe , нитридов III группы. Показано, что в результате односторонней инжекции через эмиттерный гетеропереход повышается усиление по току и предельная частота прибора. Выполнено сравнение параметров БГТ с теоретическими оценками. Рассмотрены элементы конструирования БГТ: градиент состава в переходе эмиттер – база (ЭБ), варизонная база, БГТ с двойным гетеропереходом, углеродное ограничение профиля бора в базе SiGe -БГТ. Продемонстрированы достижения современных БГТ, изготовленных на различных гетеропереходных парах. БГТ на SiGe и обычных материалах $A^{\text{III}}B^{\text{V}}$ характеризуются низким уровнем шумов, высокими предельной частотой и усилением. Особые перспективы имеют InP -БГТ. Нитриды III группы (AlGaIn/GaN) являются также перспективными широкозонными материалами для БГТ, однако на пути их создания имеются определенные технологические трудности.

В 1948 г. американские ученые Дж. Бардин и У. Браттейн создали полупроводниковый триод, или транзистор. Это событие имело громадное значение для развития полупроводниковой электроники. Транзисторы могут работать при значительно меньших напряжениях, чем ламповые триоды, и предназначены для усиления, генерации и переключения электрических сигналов. Создание и совершенствование транзисторов в свою очередь стимулировали исследования в области физики полупроводников, способствовали развитию других полупроводниковых приборов.

Транзистором называется полупроводниковый прибор с двумя электронно-дырочными переходами. В транзисторе используются оба типа носителей – основные и неосновные (электроны и дырки), поэтому его называют биполярным.

Основные характеристики транзистора определяются в первую очередь процессами, происходящими в базе. В зависимости от распределения примесей, в базе может присутствовать или отсутствовать электрическое поле. Если при отсутствии токов в базе существ-

вует электрическое поле, которое способствует движению неосновных носителей заряда от эмиттера к коллектору, то транзистор называют дрейфовым, если же поле в базе отсутствует – бездрейфовым (диффузионным) [1].

3.1. Биполярный транзистор

3.1.1. Основные характеристики

Биполярный транзистор состоит из двух $p-n$ переходов, расположенных на близком расстоянии друг от друга: толщина базы – расстояние между ними – $W_b < L_{p,n}$ – меньше диффузионной длины неосновных носителей. При работе в активном режиме первый $p-n$ переход – ЭБ – смещен в прямом направлении, второй – БК – находится при обратном смещении. Неосновные носители, инжектированные в базу через ЭБ-переход, собираются благодаря диффузии или дрейфу в коллекторной области за коллекторным переходом. Базовый ток обуславливает рекомбинацию в базе, инжекцию неосновных носителей в эмиттер и ток утечки через обратносмещенный переход БК.

Для токов транзистора справедливо соотношение $I_3 = I_k + I_6$, при этом $I_3 \approx I_k \ll I_6$.

Биполярный транзистор является активным прибором, позволяющим осуществлять усиление мощности электрических сигналов. Усиление в транзисторе происходит в результате модуляции сопротивления коллекторного $p-n$ перехода.

Коэффициент передачи тока эмиттера

$$\alpha = \frac{dI_k}{dI_3} = \gamma\beta M ,$$

где $\gamma = \frac{dI_{p3}}{dI_3} = 1 - \frac{\sigma_6 W}{\sigma_3 L_n} = 1 - \frac{\rho_3 W}{\rho_6 L_n}$ – эффективность, или коэффициент инжекции, эмиттера (для $p-n-p$ транзистора);

$$\beta = \frac{dI_{pk}}{dI_{p3}} = 1 - \frac{1}{2} \left(\frac{W}{L_p} \right)^2 - \text{коэффициент переноса, характеризующий}$$

рекомбинационные потери неосновных носителей в базе;

здесь $dI_{pэ}$, dI_{pk} – приращения тока дырок, инжектируемых эмиттером в базу и достигающих коллекторного перехода соответственно; W – толщина базы;

M – эффективность коллектора, полагаем ее равной единице.

Ток эмиттера

$$I_э = A_э q D_p \frac{dp}{dx} = A_э q D_p \frac{p_{nэ}}{W - d_k},$$

где $A_э$ – площадь эмиттерного перехода;

$$p_{nэ} = p_{n0} \exp\left(\frac{qU_э}{kT}\right) - \text{концентрация дырок в базе у эмиттерного}$$

перехода;

d_k – толщина области пространственного заряда (ОПЗ) коллекторного перехода.

Для значительного по величине сигнала применяют интегральный коэффициент передачи тока эмиттера, или коэффициент передачи по постоянному току:

$$\bar{\alpha} = \frac{I_k - I_{к60}}{I_э}, \text{ откуда } I_k = \bar{\alpha} I_э + I_{к60},$$

где $I_{к60}$ – обратный ток насыщения коллекторного перехода при $I_э = 0$.

Между α и $\bar{\alpha}$ существует связь:

$$\alpha = \bar{\alpha} + I_э \frac{d\bar{\alpha}}{dI_э},$$

откуда следует, что при малых $I_э$ и для идеализированной модели, когда $\bar{\alpha}$ не зависит от тока, $\bar{\alpha} = \alpha$.

Транзистор может работать в схемах с общей базой (ОБ), с общим эмиттером (ОЭ) и общим коллектором (ОК).

Для схемы с ОБ коэффициент передачи по току

$$K_{I_{ОБ}} = \alpha = \frac{dI_k}{dI_э} < 1,$$

т.е. усиление тока не происходит. Однако схема с ОБ позволяет получить усиление по напряжению. Выходное дифференциальное со-

противление транзистора в этой схеме $r_{\text{ВЫХ}}$ велико и предполагает включение большого сопротивления нагрузки $R_{\text{Н}}$. Входное дифференциальное сопротивление транзистора в данной схеме значительно меньше выходного: $r_{\text{ВХ}} \ll r_{\text{ВЫХ}}$. Изменение входного напряжения при изменении входного тока на $dI_{\text{э}}$ составляет

$$dU_{\text{э}} = r_{\text{ВХ}} dI_{\text{э}}.$$

Соответственно выходное напряжение изменяется на величину

$$dU_{\text{к}} = (R_{\text{Н}} \parallel r_{\text{ВЫХ}}) dI_{\text{к}},$$

где $R_{\text{Н}} \parallel r_{\text{ВЫХ}} = \frac{R_{\text{Н}} r_{\text{ВЫХ}}}{R_{\text{Н}} + r_{\text{ВЫХ}}}$ – сопротивление, определяемое параллельным включением $R_{\text{Н}}$ и $r_{\text{ВЫХ}}$.

Коэффициент усиления по напряжению в схеме с ОБ

$$K_{U \text{ ОБ}} = \frac{dU_{\text{к}}}{dU_{\text{э}}} = \frac{dI_{\text{к}} (R_{\text{Н}} \parallel r_{\text{ВЫХ}})}{dI_{\text{э}} r_{\text{ВХ}}} = \alpha \frac{R_{\text{Н}} \parallel r_{\text{ВЫХ}}}{r_{\text{ВХ}}} > 1.$$

Коэффициент усиления по мощности также заметно больше единицы:

$$K_{P \text{ ОБ}} = \frac{dP_{\text{ВЫХ}}}{dP_{\text{ВХ}}} = \frac{dU_{\text{к}} dI_{\text{к}}}{dU_{\text{э}} dI_{\text{э}}} = K_{U \text{ ОБ}} \cdot K_{I \text{ ОБ}} = \alpha^2 \frac{R_{\text{Н}} \parallel r_{\text{ВЫХ}}}{r_{\text{ВХ}}} > 1.$$

В транзисторе, включенном по схеме с ОЭ, входным током является ток базы, а выходным – ток коллектора.

Коэффициент усиления по току в схеме с ОЭ

$$K_{I \text{ ОЭ}} = B = \frac{dI_{\text{к}}}{dI_{\text{б}}} = \frac{dI_{\text{к}}}{dI_{\text{э}} - dI_{\text{к}}} = \frac{\alpha}{1 - \alpha}.$$

Интегральный коэффициент усиления по току

$$\bar{B} = \frac{I_{\text{к}} - I_{\text{кб0}}}{I_{\text{б}} + I_{\text{кб0}}}.$$

Тогда в схеме с ОЭ

$$I_{\text{к}} = \bar{B} I_{\text{б}} + \frac{I_{\text{кб0}}}{1 - \alpha} = \bar{B} I_{\text{б}} + I_{\text{кэ0}},$$

где $I_{\text{кэ0}} = \frac{I_{\text{кб0}}}{1 - \alpha}$ – обратный ток коллектора при $I_{\text{б}} = 0$.

Коэффициент усиления по напряжению в схеме с ОЭ

$$K_{U \text{ ОЭ}} = \frac{dI_{\kappa}}{dI_{\delta}} \frac{(R_{\text{H}} \parallel r_{\text{ВЫХ}}^x)}{r_{\text{ВХ}}^x} = B \frac{(R_{\text{H}} \parallel r_{\text{ВЫХ}}^x)}{r_{\text{ВХ}}^x},$$

где $r_{\text{ВХ}}^x = r_{\text{ВХ}}(B+1)$ и $r_{\text{ВЫХ}}^x = \frac{r_{\text{ВЫХ}}}{(B+1)}$ – соответственно входное и выход-

ное сопротивление транзистора в схеме с ОЭ.

Коэффициент усиления по мощности

$$K_{P \text{ ОЭ}} = B^2 \frac{(R_{\text{H}} \parallel r_{\text{ВЫХ}}^x)}{r_{\text{ВХ}}^x}$$

оказывается наибольшим по сравнению с другими схемами включения, так как в схеме с ОЭ имеется усиление и по току, и по напряжению.

Коэффициент усиления по току B в схеме с ОЭ имеет верхний предел

$$B_{\text{max}} = I_{n3} / I_{p3},$$

где I_{n3} – ток электронов, инжектированных из эмиттера;

I_{p3} – ток дырок, инжектируемых в эмиттер.

Таким образом, верхний предел усиления связан с эффективностью инжекции эмиттера:

$$\gamma = I_{n3} / I_3.$$

Выражение для B_{max} имеет вид

$$B_{\text{max}} = N_{D3} v_{n\delta} / N_{A\delta} v_{p3},$$

где $v_{n\delta}$ и v_{p3} – эффективные скорости электронов в базе и дырок в эмиттере соответственно, включающие диффузионную и дрейфовую составляющие.

Транзистор в схеме с ОК имеет усиление по току:

$$K_{I \text{ ОК}} = \alpha_{\text{ОК}} = \frac{dI_3}{dI_6} = \frac{1}{1-\alpha} = (B+1) > 1.$$

Однако коэффициент усиления по напряжению в этой схеме незначительно меньше единицы, так как разность потенциалов между

базой и эмиттером практически не зависит от тока базы. Потенциал эмиттера повторяет потенциал базы, поэтому схему с ОК называют эмиттерным повторителем. Коэффициент усиления по мощности больше единицы. Большое входное сопротивление и малое выходное сопротивление позволяют использовать схему с ОК для согласования сопротивлений каскадов в радиотехнических схемах.

3.1.2. Эквивалентная схема транзистора. Предельные частоты

Транзистор на малом переменном сигнале, $U_3 < kT/q$, может быть представлен в виде эквивалентной схемы (рис. 3.1). Перечислим составляющие ее элементы.

Дифференциальное сопротивление эмиттера

$$r_3 = \left. \frac{dU_3}{dI_3} \right|_{U_3} = \frac{kT}{qI_3}.$$

Для значений тока $I_3 = 1 \dots 20$ мА $r_3 \leq 25$ Ом. Обычно внутреннее сопротивление генератора на входе $R_r \gg r_3$. Поэтому в схеме с ОБ осуществляется режим управления током ($I_3 = \text{const}$).

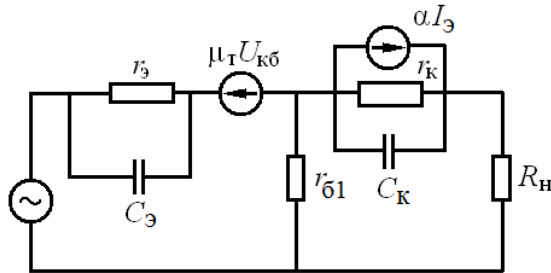


Рис. 3.1. Эквивалентная схема биполярного транзистора

Дифференциальное сопротивление коллектора при заданном токе эмиттера $I_3 = \text{const}$ (схема с ОБ) и при $n_б \ll p_к$

$$r_к = \frac{dU_к}{dI_к} = \frac{U_к W}{I_3 d_к (1 - \alpha)};$$

для значений $U_к = 10$ В, $I_3 = 5$ мА, $W/d_к \approx 10$, $\alpha = 0,98$ ($B = 0$) получаем $r_к = 10^6$ Ом.

Для схемы с ОЭ ($I_6 = \text{const}$) $r'_k = r_k (1 - \alpha)$, т.е. $r'_k \ll r_k$; для значений $B = 10 \dots 100$ r'_k оказывается в $10 \dots 100$ раз меньше r_k , т.е. составляет $10^3 \dots 10^5$ Ом и может изменяться в широких пределах в зависимости от способа включения и значения сопротивления источника сигнала на входе.

Сопротивление базы

$$r_6 = r_{61} + r_{62},$$

где $r_{61} = k\rho_6 / W$ – объемное сопротивление базы.

Наличие r_{61} приводит к уменьшению действующего напряжения на эмиттере и появлению отрицательной обратной связи. Тот же эффект возникает за счет модуляции ширины базы при $I_3 = \text{const}$. С ростом U_k уменьшается W , что эквивалентно появлению дополнительного сопротивления

базы $r_{62} = \frac{dU_3}{dI_6} = \frac{r_3}{2(1-\alpha)}$, или, поскольку $I_6 = I_3 - I_k$,

$$r_{62} = -\frac{dU_3}{dI_k} = -\frac{dU_k}{dI_k} \frac{dU_3}{dU_k} = r_k (-\mu_T),$$

где $\mu_T = \frac{dU_3}{dU_k}$ – коэффициент обратной связи по напряжению; знак

минус указывает на то, что с ростом напряжения U_k напряжение на эмиттере U_3 убывает.

Влияние I_3 на I_k в эквивалентной схеме учитывают введением генератора тока αI_3 , влияние эффекта модуляции – с помощью r_{62} или введением генератора напряжения $\mu_T U_{к6}$.

Расчеты показывают, что

$$r_{\text{вх.об}} = r_3 + r_6(1 - \alpha) = r_3 + \frac{r_6}{B + 1};$$

$$r_{\text{вх.оэ}} = r_6 + r_3(B + 1); \quad r_{\text{вых.оэ}} = \frac{r_k}{(B + 1)}.$$

При работе на высоких частотах коэффициент усиления транзистора уменьшается из-за инерционности переноса носителей через структуру транзистора, а также шунтирующего действия и перезарядки емкостей эмиттерного и коллекторного переходов. С увеличением частоты возрастает роль диффузионной емкости эмиттерного перехода

$$C_3 = \frac{qI_3 W^2}{2kT D_p},$$

где D_p – коэффициент диффузии дырок, и зарядной емкости коллекторного перехода

$$C_k = S \left[qn_0 \left(\frac{\varepsilon \varepsilon_0}{2U_k} \right) \right]^{\frac{1}{2}}.$$

Явления, связанные с модуляцией ширины базы, проявляются лишь при сравнительно невысоких частотах $f < f_n = (2\pi r_k C_k)^{-1}$. При $f > f_n$ изменением W можно пренебречь и обратная связь становится пренебрежимо малой. В эквивалентной схеме следует заменить r_k на $Z_k = \frac{1}{i\omega C_k}$. При высоких частотах сопротивление Z_k шунтирует сопротивление нагрузки R_n , полезная выходная мощность при этом уменьшается.

Влияние C_3 не столь велико, поскольку ее шунтирующее действие проявляется лишь на очень высокой частоте, когда сопротивление $\frac{1}{i\omega C_3}$ становится соизмеримым с малым значением r_3 .

Еще одной причиной уменьшения коэффициента усиления транзистора является инерционность переноса носителей через базу.

В итоге при работе на высоких частотах коэффициент усиления транзистора уменьшается из-за инерционности переноса носителей через структуру транзистора, а также шунтирующего действия и перезарядки емкостей эмиттерного и коллекторного переходов. Существует несколько предельных частот транзистора.

Частота, при которой в схеме с ОБ α уменьшается в $\sqrt{2}$ раза по сравнению с ее значением при низкой частоте α_0 , называется предельной частотой и определяется так:

$$f_\alpha = \frac{1,2D_p}{\pi W^2} = \frac{1}{\pi t_{\text{диф}}},$$

где $t_{\text{диф}}$ – время диффузии носителей через базу.

В схеме с ОЭ предельной частотой является частота, при которой B уменьшается в $\sqrt{2}$ раз:

$$f_B = \frac{1}{2\pi\tau_p}.$$

Между предельными частотами существует соотношение:

$$f_B = f_\alpha(1 - \alpha_0), \text{ т.е. } f_B \ll f_\alpha, \text{ так как } t_{\text{диф}} \ll \tau_p.$$

Причина существенной разницы предельных частот f_B и f_α заключается в том, что в отличие от схемы с ОБ, когда избыточные электроны через время $t_{\text{диф}}$ свободно выходят через базовый контакт после прохождения дырок, несущих сигнал, из эмиттера через базу в коллектор, в случае схемы с ОЭ при заданном токе $I_0 = \text{const}$ заряд электронов, вошедший через базовый контакт в качестве полезного сигнала, остается в базе и попеременно нейтрализуется дырками, выходящими из эмиттера и уходящими в коллектор, до тех пор, пока через время жизни τ_p все вошедшие электроны не исчезнут благодаря рекомбинации.

Еще одна критическая частота, при которой в схеме с ОЭ (при $R_H = 0$) коэффициент B становится равным единице, определяется выражением $f_k = 2\pi(t_0)^{-1}$, где $t_0 \approx \tau_3 + t_{\text{диф}} + t_k + \tau_k$ – суммарное время задержки. Здесь $\tau_3 = r_3 C_3$ – постоянная времени эмиттера; $t_k = \frac{d_k}{2v}$ – время пролета через обедненный слой коллектора d_k , $v \approx 10^7$ см/с.

На высоких частотах C_k шунтирует r_k (и R_H), но заряжается через r_6 , поэтому постоянная времени коллектора $\tau_k = r_6 C_k$.

Наличие усиления по току не является обобщающей характеристикой биполярного транзистора. Транзистор может быть использован в качестве усилителя, генератора только при наличии усиления по мощности. Поэтому вводится понятие максимальной частоты, при которой коэффициент усиления по мощности снижается до единицы:

$$f_{\text{max}} = \sqrt{\frac{\alpha_0 f_\alpha}{30r_{61} C_k}}.$$

Из приведенного выше соотношения при коэффициенте передачи по току $\alpha_0 = 1$ получаем для n - p - n транзистора $f_{\max} = \sqrt{Q}$, где выражение под корнем Q называют фактором качества транзистора:

$$Q = (p_6 U_k)^{1/2} \mu_p \mu_n / W S_k.$$

Улучшения высокочастотных свойств транзистора достигают уменьшением толщины базы и площади S_k коллекторного p - n перехода. В современных высокочастотных транзисторах W и S_k уменьшены до технически возможных пределов.

Увеличение p_6 не приводит к повышению Q , так как при этом снижается напряжение пробоя коллектора. Следовательно, снижается и максимально допустимое рабочее напряжение коллектора. Кроме того, с ростом p_6 уменьшается эффективность эмиттера и подвижность носителей заряда в базе. Так обстоит дело в гомобиполярном транзисторе.

Биполярный гетеротранзистор (БГТ) дает дополнительные возможности увеличения f_{\max} , жестко не связывая величины, входящие в формулу для фактора качества.

3.1.3. Дрейфовый планарный биполярный транзистор

В предыдущих разделах был представлен перенос инжектированных носителей через базу биполярного транзистора. Процесс переноса рассматривался как диффузионный, поскольку электрическое поле в базе отсутствует. При диффузионном переносе скорость направленного движения носителей невысока и, следовательно, время переноса носителей через базу будет большим. Для повышения быстродействия транзисторов необходимо уменьшить время пролета, а следовательно, увеличить скорость движения инжектированных носителей в базе. Один из способов достижения такого эффекта – переход от диффузионного к дрейфовому механизму переноса в базе.

За счет внешних источников напряжения создать электрическое поле в квазинейтральном объеме барьерных структур не представляется возможным. В дрейфовых транзисторах используется принцип встраивания электрического поля в базу. Этот принцип реализуется путем неоднородного легирования базы.

Рассмотрим неоднородно-легированный полупроводник n -типа, в котором концентрация примеси меняется в зависимости от координаты x . В таком полупроводнике существует градиент концентрации

свободных носителей, который приводит к возникновению диффузионного тока $J_D = qD \frac{dn}{dx}$. Этот ток вызывает перераспределение сво-

бодных носителей, в то время как ионизованные доноры остаются на своих прежних местах. Вследствие этого возникает электрическое поле E , препятствующее дальнейшему перемещению электронов и вызывающее появление дрейфовой компоненты электрического тока.

Один из часто встречающихся случаев – экспоненциальное распределение примеси $N_D = N_0 \exp(-\frac{x}{L_0})$, где $L_0 = 2\sqrt{Dt}$ – характери-

стическая диффузионная длина, описывающая снижение концентрации донорной примеси при удалении от поверхности ($x = 0$); D – коэффициент диффузии примеси; t – время диффузии. В таком случае

$E = \frac{kT}{q} \frac{1}{L_0}$, т.е. благодаря той связи, которая существует между экс-

понентой и ее производной, напряженность встроенного электрического поля во всей области, где распределение примеси можно считать экспоненциальным, имеет постоянную величину, не зависящую от концентрации N_0 на поверхности.

Применим эту ситуацию к биполярному транзистору $p-n-p$ типа. В случае неоднородно-легированной базы (причем вблизи эмиттера база должна быть сильно легирована, а вблизи коллектора – слабо) электрическое поле в базе направлено от эмиттерного перехода к коллекторному. При инжекции неосновных носителей (дырок) они будут ускоренно двигаться в электрическом поле и добавят к диффузионному переносу через базу дрейфовую составляющую (рис. 3.2).

Для того чтобы точно найти распределение инжектированных носителей по базе биполярного транзистора $p(x)$, нужно решить уравнение непрерывности с учетом дрейфовой и диффузионной компонент тока.

Сравним время переноса через базу в биполярном транзисторе при дрейфовом ($t_{др}$) и диффузионном ($t_{диф}$) переносе:

$$t_{др} = \frac{W}{\mu E} = \frac{W \frac{kT}{q} L_0}{D \frac{kT}{q}} = \frac{WL_0}{D}; \quad t_{диф} = \frac{W^2}{2D}.$$

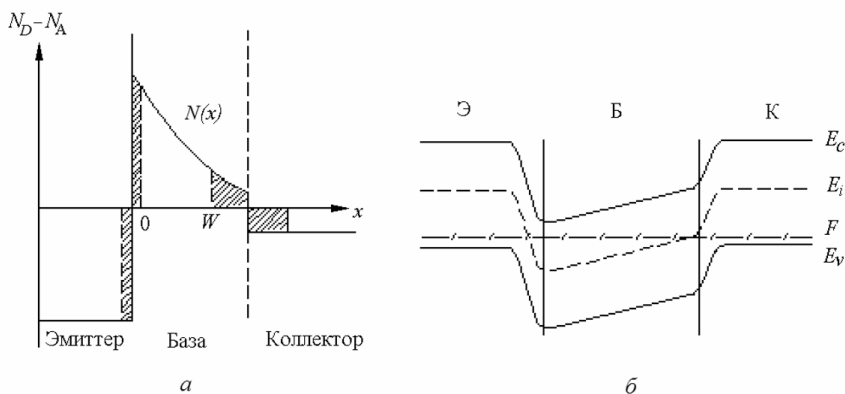


Рис. 3.2. Диаграмма, иллюстрирующая распределение концентрации легирующей примеси дрейфового транзистора, (а) и зонная диаграмма (б)

$$\text{Отношение } \frac{t_{\text{др}}}{t_{\text{диф}}} = \frac{W^2 L_0}{D} \frac{2D}{W^2} = \frac{2L_0}{W}.$$

Для транзисторов с диффузионным переносом носителей через базу определяющее значение имеет время диффузии через базу $t_{\text{диф}}$. При этом критическая частота $f_k \sim 10^7$ Гц. У дрейфовых транзисторов $f_k \sim 10^9$ Гц.

Есть некоторые требования к структуре биполярного транзистора, которые в рамках гомопереходов, составляющих транзистор, выполнить крайне трудно. Прежде всего это касается толщины базы и степени ее легирования. Уменьшение толщины базы, необходимое для сокращения времени переноса, ведет к увеличению сопротивления базы r_b . Увеличение степени легирования базы снижает эффективность эмиттера.

Есть и общие требования к полупроводниковому материалу. Так как поле очень большое, скорость носителей достигает насыщения, и поэтому желательно, чтобы используемый материал имел возможно большую скорость насыщения.

Наконец, поскольку усиление мощности возрастает с увеличением поля в коллекторном переходе, напряжение пробоя материала должно быть как можно большим.

Использование в качестве эмиттера широкозонного материала, т.е. создание гетеропереходного биполярного транзистора, позволяет

легировать базу значительно сильнее и уменьшать ее толщину без увеличения базового сопротивления.

Биполярные гетеропереходные транзисторы имеют и другие потенциальные преимущества перед обычными биполярными транзисторами.

3.2. Принцип действия биполярного гетеротранзистора

Идея использования широкозонного эмиттера для повышения коэффициента инжекции в транзисторе была предложена У.Б. Шокли еще в 1948 г. Однако реализации этой идеи в те времена не было уделено достаточного внимания. Это было связано с большими технологическими трудностями получения гетеропереходов с малым числом дефектов на границе раздела, тогда как технология изготовления транзисторов с гомопереходами в то время развивалась успешно. Однако на пути улучшения параметров биполярных транзисторов (увеличения коэффициента усиления, предельной частоты) существуют определенные ограничения. Одним из них является невозможность применения сильнолегированной базы. С развитием технологии изготовления гетеропереходов снова возник интерес к биполярным транзисторам с гетеропереходами (БГТ). Последнее десятилетие развития полупроводниковой электроники ознаменовалось бурным прогрессом в разработке и выпуске в промышленных масштабах биполярных гетеропереходных транзисторов на основе известных гетеропереходных пар Si/GeSi, AlGaAs/GaAs, InGaP/GaAs, а также попытками разработки БГТ на основе таких новых материалов, как нитриды III группы: GaN/SiC, AlGaN/GaN.

Отличительной особенностью БГТ является эмиттерный гетеропереход, составленный из широкозонного эмиттера и узкозонной базы. Коллекторный переход также может быть гетеропереходом. В настоящее время реализованы гетеропереходные транзисторные структуры на основе различных материалов. В табл. 3.1 приведены примеры таких гетерокомпозиций.

Целью данного раздела является анализ современного состояния разработок биполярных гетеропереходных транзисторов на материалах $A^{III}B^V$, SiGe и нитридах III группы, сравнение и оценка возможностей материалов и приборов на их основе [2].

Таблица 3.1

Гетерокомпозиции, на основе которых созданы биполярные гетеротранзисторы

| Подложка | Эмиттер | База | Коллектор |
|--------------------------------|------------|---|--|
| GaAs | AlGaAs | GaAs или AlGaAs (с градиентом состава) | GaAs или AlGaAs (БГТ с двойным гетеропереходом – ДБГТ) |
| | GaInP | GaAs или AlGaAs (с градиентом состава) | GaAs или GaInP (ДБГТ) |
| InP | InP | InGaAs | InGaAs или InP (ДБГТ) |
| | AllnAs | InGaAs | InGaAs или InP (ДБГТ) |
| | InP | GaAsSb | InP (ДБГТ) |
| Si | Si/поли-Si | Si-Ge (с градиентом состава) | Si |
| Al ₂ O ₃ | AlGaN | GaN | GaN |

У БГТ ширина запрещенной зоны эмиттера больше, чем ширина запрещенной зоны базы. Если для простоты положить, что разность значений ширины запрещенных зон ΔE_g совпадает по величине с разрывом валентной зоны ΔE_v , то зонная диаграмма БГТ будет иметь вид, представленный на рис. 3.3.

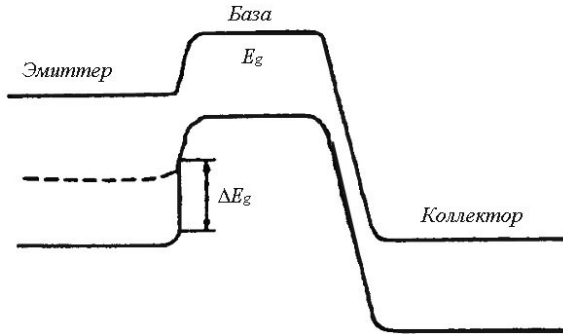


Рис. 3.3. Упрощенная энергетическая диаграмма БГТ; штриховой линией показан край валентной зоны транзистора с гомопереходом

В этом случае выражение для максимального коэффициента усиления по току в схеме с ОЭ можно записать в виде

$$B_{\max} = N_{D_3} / N_{A_6} v_{n_6} / v_{p_3} \exp(\Delta E_g / kT).$$

Таким образом, высокие значения B_{\max} могут быть достигнуты даже при $N_{D_3} < N_{A_6}$. Реальная зонная диаграмма гетеропереходного эмиттера имеет вид, представленный на рис. 3.4.