

Федеральное агентство по образованию

**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)**

Кафедра телевидения и управления (ТУ)

В.М. Ицкович

ЭЛЕКТРОНИКА

Раздел 1

Учебное пособие

**Рекомендовано Сибирским региональным отделением
учебно-методического объединения высших
учебных заведений РФ по образованию в области
радиотехники, электроники, биомедицинской
техники и автоматизации для межвузовского
использования в качестве учебного пособия**

2005

Рецензент: д.т.н., профессор ТПУ Ройтман М.С.

Корректор: Осипова Е.А.

Ицкович В.М.

Электроника: Учебное пособие. В 2-х разделах. — Томск: Томский межвузовский центр дистанционного образования, 2005. — Р.1. — 209 с.

© Ицкович В.М., 2005
© Томский межвузовский центр
дистанционного образования, 2005

СОДЕРЖАНИЕ

Предисловие	5
1. ПОЛУПРОВОДНИКОВЫЕ МАТЕРИАЛЫ	6
1.1 Введение	6
1.2 Структура полупроводниковых материалов	6
1.3 Энергетические зоны твердого тела.....	13
1.4 Зонная структура полупроводников	16
1.5 Концентрации носителей в зонах полупроводника.....	20
1.6 Уровень Ферми	25
1.7 Концентрация носителей в полупроводниках	28
1.8 Подвижность носителей и удельная проводимость	35
1.9 Генерация и рекомбинация носителей в полупроводниках	39
1.10 Плотность тока в полупроводниках	41
1.11 Заряды в полупроводниках	44
1.12 Движение зарядов в полупроводниках	46
2. ДИОДЫ	56
2.1 Введение	56
2.2 Электронно-дырочный переход.....	58
2.3 Контакты металл-полупроводник	74
2.4 Анализ идеализированного диода	77
2.5 Обратная и прямая характеристики реального диода	87
2.6 Переходные характеристики плоскостного диода	96
3. РАЗНОВИДНОСТИ ДИОДОВ.....	107
3.1 Точечные диоды	107
3.2 Полупроводниковые стабилитроны	111
3.3 Туннельные диоды	114
3.4 Диоды Шоттки.....	118
3.5 Фотоприёмники (приёмники оптического излучения)	120
3.6 Фотодиоды	121
4. БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ.....	124
4.1 Введение	124
4.2 Основные процессы в биполярном транзисторе	128
4.3 Статические характеристики транзистора	135
4.4 Статические параметры транзистора	141
4.5 Динамические параметры транзистора.....	148

4.6	Зависимость параметров транзистора от режима и температуры.....	154
4.7	Характеристики и параметры транзистора при включении с общим эмиттером.....	157
4.8	Разновидности эквивалентных схем	166
4.9	Составные транзисторы.....	168
4.10	Допустимая мощность	169
4.11	Дрейфовые транзисторы.....	172
5.	ПОЛЕВЫЕ ТРАНЗИСТОРЫ.....	177
5.1	Полевой транзистор с управляющим p - n переходом	177
5.2	Полевые транзисторы с изолированным затвором	184
6.	ТРАНЗИСТОРНЫЕ КЛЮЧИ	194
6.1	Введение	194
6.2	Статические характеристики ключа ОЭ.....	195
6.3	Переходный процесс в насыщенном ключе при открывании транзистора.....	198
6.4	Методы сокращения времени переходного процесса	205

ПРЕДИСЛОВИЕ

Настоящее учебное пособие написано на основе лекций, которые автор читает длительное время по дисциплинам: «Электронные приборы», «Электроника и микроэлектроника», «Твердотельные электронные приборы и основы микроэлектроники».

В учебном пособии использованы работы И.П. Степаненко «Основы теории транзисторов и транзисторных схем», «Основы электроники» и других авторов.

Для лучшего усвоения теоретического материала в конце каждой главы приведены вопросы для самопроверки.

Пособие написано в соответствии с ГОС для направлений «Радиотехника» и «Телекоммуникации».

При написании учебного пособия автор не ставил перед собой задачи рассмотреть все виды полупроводниковых приборов, которые известны в настоящее время.

1. ПОЛУПРОВОДНИКОВЫЕ МАТЕРИАЛЫ

1.1 Введение

К полупроводниковым материалам относят вещества, которые при комнатной температуре имеют удельное электрическое сопротивление ρ в пределах от 10^{-3} – 10^{-2} до 10^8 – 10^9 Ом см. Вещества с меньшим сопротивлением (10^{-6} – 10^{-4} Ом см) относят к проводникам, а со значительно большим (10^{10} – 10^{18} Ом·см) — к диэлектрикам. Количество полупроводниковых материалов, известных в настоящее время, далеко превышает число металлов и диэлектриков. Полупроводниками являются химические элементы (*Si, Ge, Se*), интерметаллические соединения (*Insb, GaA*), окислы (*Cu₂O, ZnO*), сульфиды (*Cd, Zns*), карбиды (*SiC*) и множество других химических соединений.

Различие между диэлектриками и полупроводниками в основном количественное (удельное сопротивление у диэлектриков значительно выше, больше ширина запрещенной зоны). Различие же между металлами и полупроводниками не только в величине удельных сопротивлений.

В отличие от чистых металлов сопротивление чистых полупроводников (понятие чистый полупроводник в некоторой степени условно) сильно зависит от температуры и, кроме того, с ростом температуры не увеличивается, а уменьшается. При добавлении примеси в чистый металл удельное сопротивление образующегося сплава больше удельного сопротивления каждого из компонентов, тогда как при добавлении примеси в чистый полупроводник его удельное сопротивление сильно уменьшается; например, 10^{-5} % мышьяка в германии снижает сопротивление в 200 раз.

В полупроводниковой технике под удельным сопротивлением понимают сопротивление между гранями куба с длиной ребра 1 см.

1.2 Структура полупроводниковых материалов

Полупроводниковые материалы имеют кристаллическую решетку с регулярной структурой. Каждый кристалл можно раз-

бить на множество повторяющихся однотипных элементарных ячеек — структурных молекул, которые подобно химическим молекулам состоят из нескольких атомов. Кристаллические решетки большинства полупроводниковых материалов, например кремний, арсенид галлий, которые очень широко применяются в настоящее время, называются тетраэдрической или решеткой типа алмаза. Такой тип решетки характерен для всех четырехвалентных элементов.

В основе этой решетки лежит пространственная фигура — тетраэдр, в углах и центре которой расположены атомы (рис. 1.1, *a*). Главная особенность тетраэдрической системы заключается в том, что центральный атом расположен на одинаковых расстояниях от узловых атомов.

Каждый угловой атом в свою очередь служит центральным атомом для других четырех ближайших атомов.

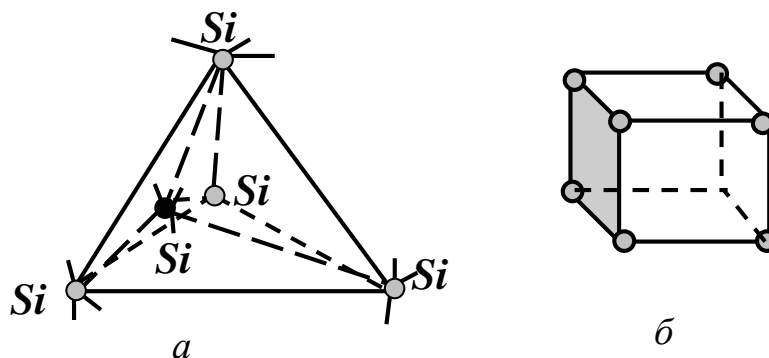


Рис. 1.1 — Тетраэдрическая структура кристаллической решетки
a — элементарный тетраэдр; *б* — элементарная ячейка

Совокупность нескольких тетраэдров образует элементарную ячейку (рис. 1.1, *б*) кубической формы с размером ребра около 5,5 А.

Удобнее пользоваться плоским эквивалентом тетраэдрической структуры (рис. 1.2), в котором сохранена главная особенность решетки типа алмаза — одинаковые расстояния между смежными атомами. Связь атомов в рассматриваемой решетке устанавливается вследствие наличия специфических обменных сил, возникающих в результате попарного объединения валентных электронов. Эти силы отражены на рис. 1.2 сходящимися

стрелками. Объединение электронов следует понимать так: пара электронов, обведенная на рис. 1.2 пунктиром, в равной степени принадлежит обоим атомам или, образно говоря, поочередно примыкает то к одному, то к другому. Соответственно «в среднем» каждый атом обладает устойчивой восьмиэлектронной оболочкой. Такая связь атомов, при которой каждый из них остается нейтральным, называется парно-электронной, ковалентной или просто валентной.

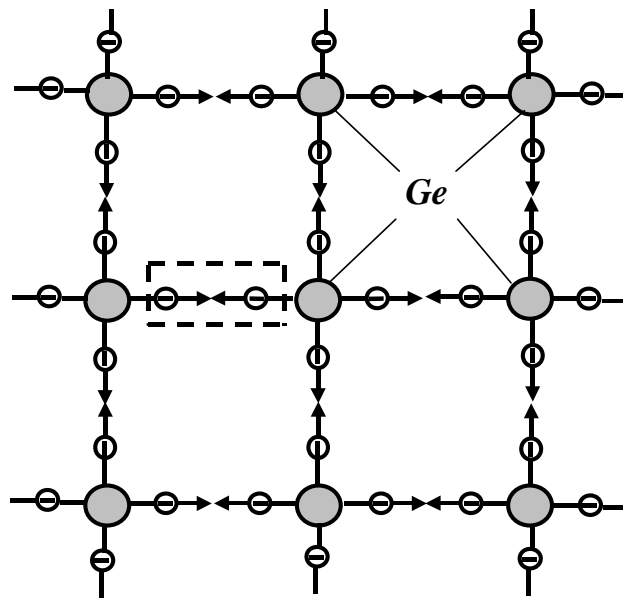


Рис. 1.2 — «Плоский» эквивалент тетраэдрической решетки с валентными связями атомов

Абсолютно однородная структура решетки (рис. 1.2) бывает у кристалла только при температуре абсолютного нуля. По мере нагревания полупроводника часть валентных связей нарушается под действием тепловых колебаний атомов в узлах кристаллической решетки. В корпускулярной интерпретации носителями энергии механических колебаний атомов являются фононы — акустические аналоги световых фотонов. Поэтому можно сказать, что при нагреве кристалла валентные связи нарушаются фононами, число и энергия которых растут с ростом температуры.

Нарушение валентных связей приводит к одновременному образованию свободных электронов и пустых мест — дырок в атомах, от которых оторвались электроны (рис. 1.3). Такая дырка ведет себя подобно частице с положительным элементарным за-

рядом. Она, так же как электрон, совершает движение в течение некоторого времени после своего появления (времени жизни), а затем рекомбинирует с одним из свободных электронов.

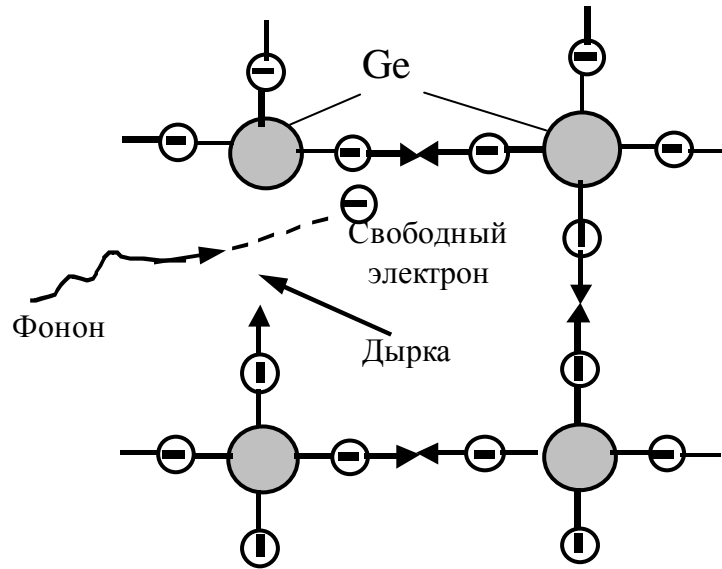


Рис. 1.3 — Процесс образования пары электрон-дырка в решетке под действием фонона (или фотона)

На рис. 1.4 показана возможная схема движения дырки в решетке полупроводника как результат последовательного замещения пустых мест электронами, принадлежащими разным атомам. Допустим, в результате воздействия фонона образовалась исходная дырка в атоме 1. Образовавшаяся дырка замещается одним из электронов атома 2 и тем самым «переходит» к атому 2; затем дырка в атоме 2 замещается одним из электронов атома 3 и тем самым «переходит» к атому 3 и т.д. Последовательное замещение пустых мест эквивалентно хаотическому движению заряда с положительным единичным зарядом. Таким образом, можно считать, что в полупроводниках имеются два типа подвижных носителей заряда — электроны и дырки. При нагревании абсолютно чистого и однородного полупроводника, который называют собственным, свободные электроны и дырки всегда образуются парами, что хорошо видно из рис. 1.3. Число этих пар в стационарном режиме определяется процессами термогенерации и рекомбинации носителей. Например, с повышением температуры растет число фононов с энергией, достаточной, чтобы увеличи-

лось число свободных электронов, а следовательно, и дырок, т.е. растет скорость термогенерации. С увеличением числа свободных электронов и дырок растет вероятность встречи электрона со свободной дыркой, что приводит к увеличению скорости рекомбинации. В системе обеспечивается термодинамическое равновесие. Проводимость собственного полупроводника, обусловленную парными носителями теплового происхождения, называют собственной. Проводимость, обусловленную наличием примесных атомов, нарушающих структуру кристаллической решетки, называют примесной проводимостью.

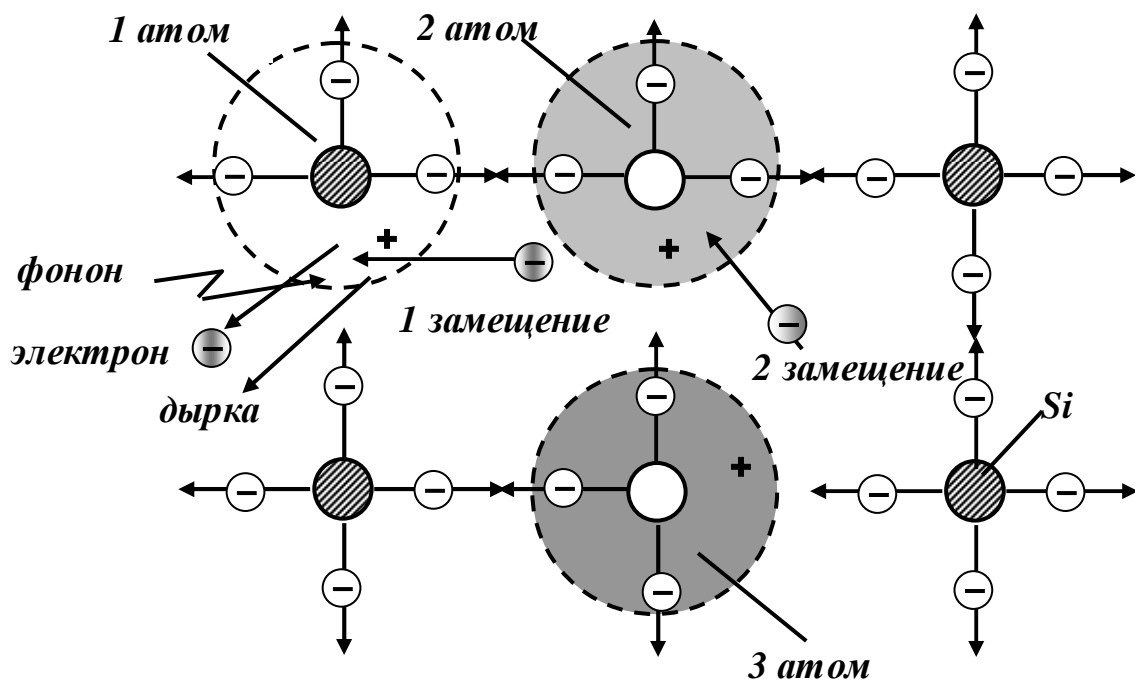


Рис. 1.4 — Схема движения свободной дырки в кристаллической решетке

В ряде случаев такие же последствия, как наличие примесных атомов, могут вызывать различные дефекты решетки: избыток одного из основных компонентов вещества, смещение некоторых атомов в узлах кристаллической решетки и др. Поэтому иногда используется понятие — дефектная проводимость, т.е. проводимость, которая возникает в полупроводниковом материале вследствие нарушения регулярности в кристаллической решетке.

Тепловые колебания решетки и образующиеся при этом фо-

ноны делятся на два типа: акустические (низкочастотные) и оптические (высокочастотные) Первые являются результатом синфазных, а вторые — противофазных колебаний атомов в узлах кристаллической решетки.

Атомы примеси могут по-разному располагаться в решетке полупроводника. В полупроводнике примесные атомы замещают основные атомы в узлах решетки.

Результаты такого замещения зависят от типа примеси. Если ввести в германий атом пятивалентной сурьмы, то четыре из его валентных электронов вступают в связь с четырьмя электронами соседних атомов германия (рис. 1.5, *a*) и образуют устойчивую оболочку из восьми электронов. Девятый электрон в этой комбинации оказывается слабо связанным с ядром, легко отрывается фононами с энергией 10^{-2} эВ и делается свободным.

Примесный атом превращается в **неподвижный ион с положительным единичным зарядом. Нельзя путать ион с дыркой, т.к. ион неподвижен и не может изменить проводимость полупроводникового материала.** Свободные электроны примесного происхождения добавляются к собственным свободным электронам, порожденным термогенерацией, и если число примесных атомов значительно превышает количество свободных электронов, проводимость полупроводника делается преимущественно электронной. Такие полупроводники называются электронными или типа *n*.

Если ввести в германий атом трехвалентного индия, результат будет другим. Для валентной связи индия с четырьмя соседними атомами германия требуется образование устойчивой восьмиелектронной оболочки, т.е. нужен дополнительный электрон. Этот электрон отбирается из основной решетки (рис. 1.5, *б*) и, будучи связанным, превращает атом индия в **неподвижный отрицательный ион.**

На том месте, откуда пришел электрон, образуется свободная дырка, которая добавляется к собственным дыркам, порожденным термогенерацией.

Полупроводники, в которых число дырок превышает количество электронов, т.е. имеют преимущественно дырочную проводимость, называются дырочными или типа *p*, а соответствующие примеси — акцепторными «принимающими» электроны.

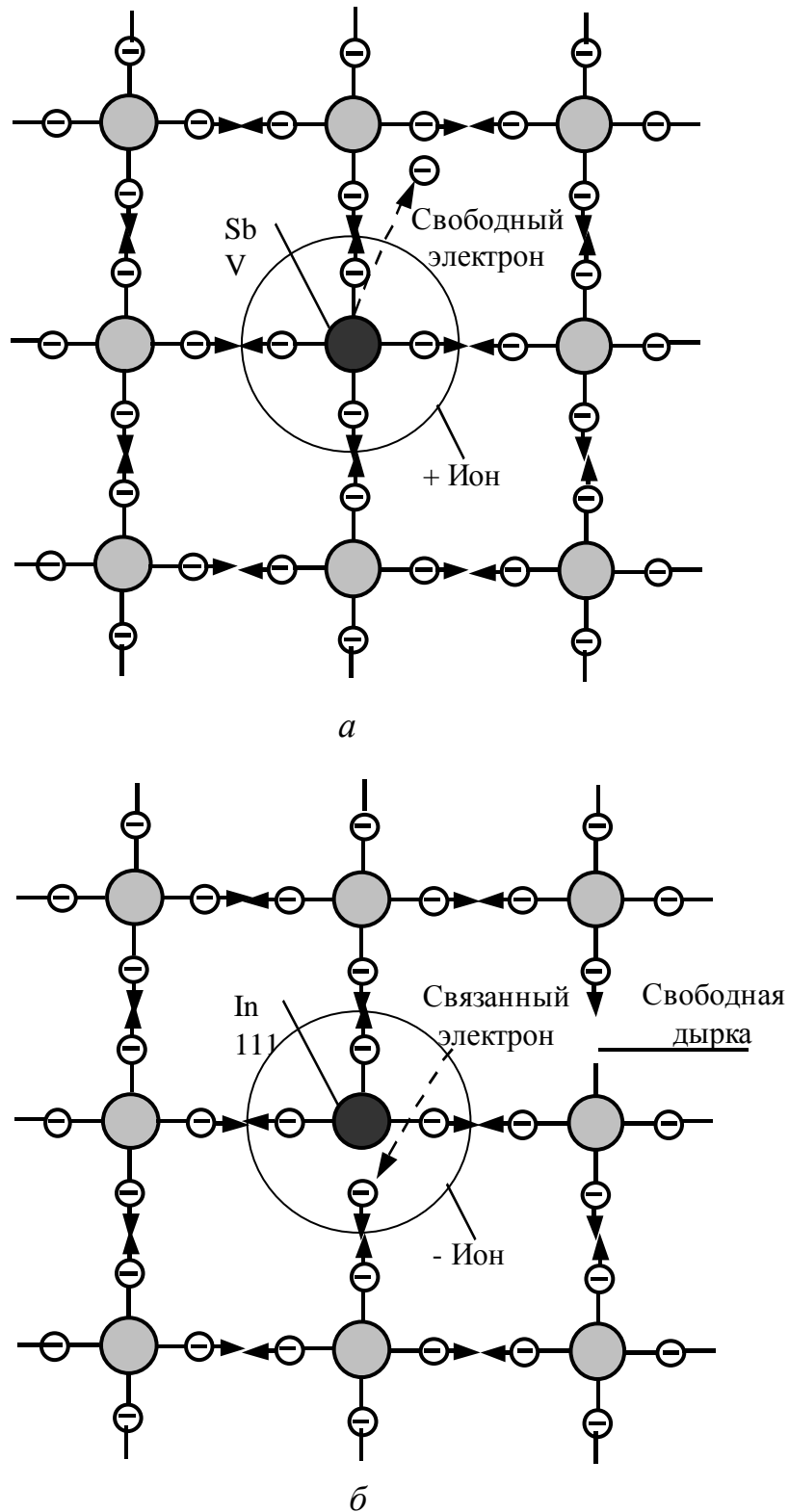


Рис. 1.5 — Замещение примесными атомами основных атомов в решетке: *a* — донорная примесь (образуются свободный электрон и неподвижный положительный ион), *б* — акцепторная примесь (образуются свободная дырка и неподвижный отрицательный ион)

Отрыв «лишнего» электрона от донора и «недостающего» электрона для акцептора требует затраты некоторой энергии (энергия ионизации или активации примеси). При нулевой температуре ($T = 0^\circ K$) ионизация не может иметь места, и проводимость равна нулю. С ростом температуры все большая часть примесных атомов ионизирует. При относительно не высоких температурах все атомы примесей оказываются ионизированными, и дальнейшее увеличение температуры приводит к увеличению только собственных электронов и дырок, что приводит к увеличению собственной проводимости.

В примесных полупроводниках количество одного типа подвижных зарядов значительно больше другого. Носители подвижного заряда, которые составляют большинство, называют основными, а те, которые составляют меньшинство, — неосновными. Таким образом, основными носителями могут быть как электроны, так и дырки.

Так, в полупроводнике p типа основные носители — дырки, неосновные — электроны.

1.3 Энергетические зоны твердого тела

Количественный анализ полупроводников и полупроводниковых приборов базируется на зонной теории твердого тела.

В невозбужденном состоянии атома его верхние уровни всегда свободны. Твердое тело представляет собой множество атомов, сильно взаимодействующих благодаря малым межатомным расстояниям. Поэтому всю совокупность атомов в некотором объеме твердого тела следует рассматривать как единое целое, как гигантскую молекулу, которая подобно атому характеризуется некоторым единым для всего тела энергетическим спектром.

Для кристалла с межатомным расстоянием d_0 получается зонная диаграмма, показанная на рис. 1.6, в которой разрешенные зоны чередуются с запрещенной зоной. Ширина тех и других обычно не превышает нескольких электрон-вольт и не зависит от числа атомов в твердом теле, т.е. от его размеров. Строго говоря, разрешенные зоны имеют дискретную структуру и состоят из стольких уровней, сколько атомов имеется в данном теле.

Количество атомов даже в микроскопических объемах настолько велико, что энергетические расстояния между уровнями зоны в реальных случаях чрезвычайно малы, т.е. разрешенные зоны практически можно считать сплошными. Энергетические «расстояния» между разрешенными зонами (т.е. ширина запрещенных зон) определяются энергией связи электронов с атомами решетки. Граничные энергетические уровни, образующие «дно» и «потолок» каждой разрешенной зоны, соответствуют потенциальной энергии электронов, их неподвижному состоянию.

Любой уровень, расположенный внутри разрешенной зоны, соответствует сумме потенциальной и кинетической энергий. **Иначе говоря, кинетическая энергия электронов возрастает по мере удаления от границы в глубь зоны и достигает максимума в ее средней части.** Проводимость в твердом теле возможна лишь тогда, когда происходит переход электрона на другой энергетический уровень. Значит, в проводимости могут участвовать электроны только тех зон, в которых есть свободные уровни. Такие свободные уровни всегда имеются в верхней разрешенной зоне. Верхнюю зону твердого тела, не заполненную (или не полностью заполненную) электронами при нулевой абсолютной температуре, называют зоной проводимости, а ближайшую к ней разрешенную зону называют валентной.

При температуре абсолютного нуля валентная зона полностью заполнена, и, следовательно, электроны этой зоны не влияют на проводимость. Но при температуре, отличной от нуля, в верхней части валентной зоны образуются свободные уровни, и эта зона также может обусловить проводимость. Следовательно, все основные процессы в полупроводниковых материалах и приборах можно изучить, анализируя процессы, которые протекают между двумя смежными зонами: зоной проводимости и валентной.

Зонная структура твердого тела лежит в основе классификации металлов, полупроводников и диэлектриков (рис. 1.6).

У металлов, как видно из рис. 1.6, зона проводимости и валентная зона взаимно перекрываются, поэтому даже при температуре абсолютного нуля в зоне проводимости находится значительное количество электронов и, следовательно, имеет место проводимость. У полупроводников и диэлектриков при $T=0^0$ К зона проводимости пуста и проводимость отсутствует; в этом их

качественное отличие от металлов. Различия же между полупроводниками и диэлектриками в основном количественные и обусловлены значительно большей шириной запрещенной зоны у диэлектриков. Поскольку зона проводимости практически сплошная, энергия электронов в этой зоне может меняться непрерывно, как у изолированных электронов в вакууме, поэтому электроны в зоне проводимости называются свободными. Термин «свободный» характеризует возможность перемещения электрона внутри твердого тела, но не возможность его вылета за пределы кристалла.

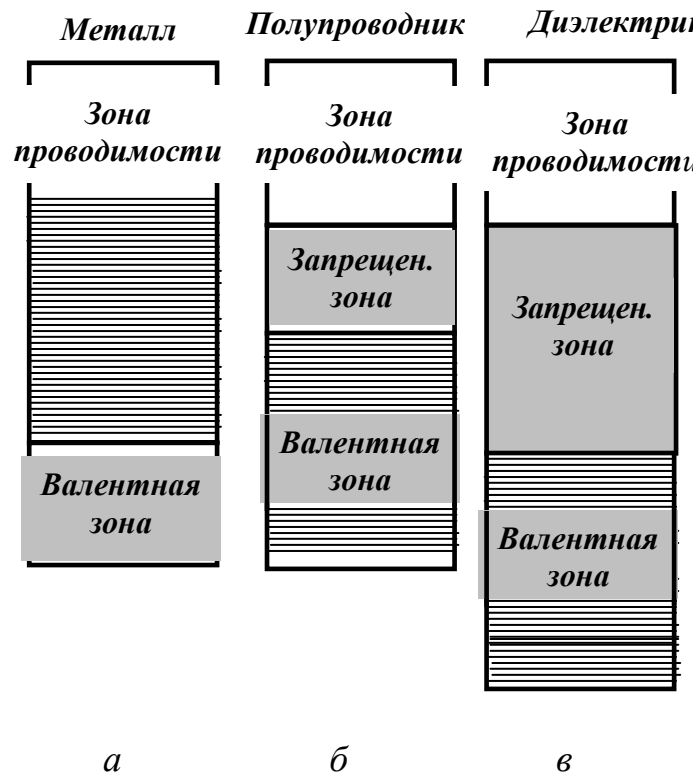


Рис. 1.6 — Зонная структура при $T = 0^\circ \text{K}$:
a — металла; *б* — полупроводника; *в* — диэлектрика

Совокупность электронов в твердом теле можно уподобить электронному газу, который заключен в «сосуд», образованный внешними гранями кристалла. Такое сравнение является чисто условным, т.к. характеристики электронов в вакууме и в твердом теле принципиально отличаются, например, дрейфовые скорости электронов при одной и той же напряженности электрического поля в вакууме и твердом теле существенно отличаются.

Масса электрона в кристалле не совпадает с его массой m в вакууме.

Поэтому в теории твердого тела пользуются понятием эффективной массы m^* , которая зависит от ряда факторов (полная энергия, направление движения и др.). Различие между значениями m и m^* может быть не только количественным, но и качественным.

Понятие дырки с точки зрения зонной теории

Предположим, что к полупроводнику приложено внешнее напряжение. Наличие электрического поля приводит в движение не только свободные электроны зоны проводимости, но и всю совокупность электронов валентной зоны (поскольку в ней имеются незаполненные верхние уровни, т.е. отсутствует часть электронов с отрицательной эффективной массой). Оказывается, что движение огромного количества электронов валентной зоны с различными эффективными массами эквивалентно движению ограниченного количества фиктивных частиц с положительным зарядом.

Рациональность понятия дырок состоит в том, что сложный анализ движения электронов в почти полной валентной зоне удастся свести к сравнительно простому анализу движения дырок в почти пустой (**по отношению к дыркам!**) валентной зоне. При этом поведение дырок оказывается аналогичным поведению свободных электронов в почти пустой (**по отношению к электронам!**) зоне проводимости. Необходимо отметить, что понятие эффективной массы допустимо по отношению к подвижным носителям; его нельзя отождествлять с понятием массы в обычном смысле этого слова. Так, различие эффективных масс не означает различие силы тяжести.

1.4 Зонная структура полупроводников

При нулевой абсолютной температуре проводимость полупроводника, в котором отсутствуют дефекты в кристаллической решетке (такой полупроводник называют собственным) равна нулю, потому что зона проводимости пуста, а валентная зона заполнена. При любой температуре, отличной от нуля, в кристалле появляются фононы, энергетический спектр которых непрерыв-

вен. Фононы с энергией, превышающей энергию ширины запрещенной зоны, переводят часть электронов из валентной зоны в зону проводимости (рис. 1.7, *a*). В результате в зоне проводимости появляются свободные электроны, а в валентной зоне — незаполненные уровни-дырки. Те и другие образуются одновременно и в равных количествах. Теперь заряды обеих зон могут двигаться в электрическом поле, обеспечивая проводимость кристалла. Перевод электрона из валентной зоны в зону проводимости аналогичен возбуждению электрона в изолированном атоме, т.е. переводу его на один из верхних уровней. Такое возбужденное состояние, как известно, является временным: через некоторое время электрон в атоме возвращается на исходный уровень, а в твердом теле — в валентную зону. Возвращение электрона из зоны проводимости в валентную зону приводит к рекомбинации электрона и дырки. Время, в течение которого электрон находился в зоне проводимости, т.е. был свободным, называют временем жизни. С увеличением температуры увеличивается число фононов с повышенной энергией и возрастает количество электронно-дырочных пар.

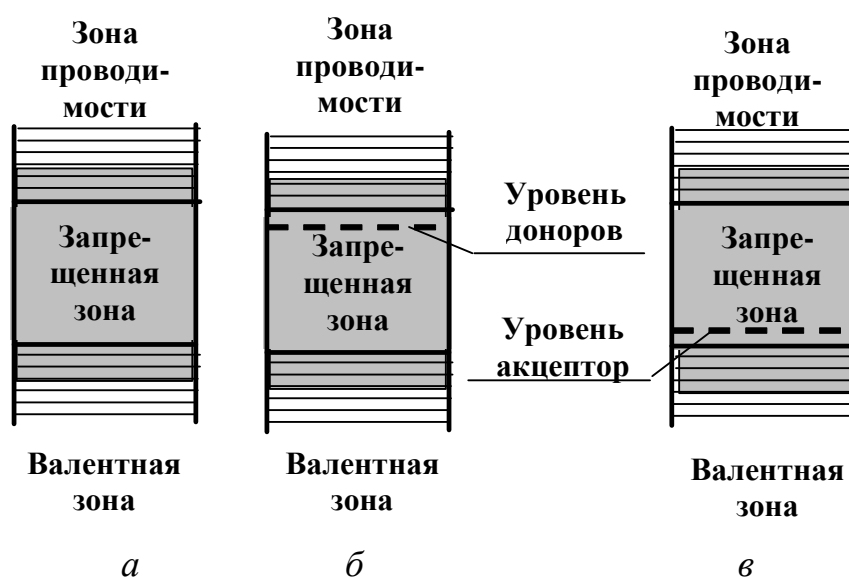


Рис. 1.7 — Зонные структуры полупроводников:
a — собственный полупроводник; *б* — электронный полупроводник; *в* — дырочный полупроводник

Зонные диаграммы для примесных полупроводников показаны на рис. 1.7, б и в. Обычно процентное содержание примесных атомов очень мало по сравнению с атомами основного материала, а, следовательно, расстояния между ними относительно велики. Поэтому их можно рассматривать как изолированные атомы, уровни которых не могут образовать зоны.

Примесные уровни изображены на диаграмме штрихами; каждый штрих соответствует атому примеси.

Донорная примесь выбирается из соображений, чтобы ее валентные уровни располагались ближе к «дну» зоны проводимости (рис. 1.7, б); все эти уровни заполнены при температуре абсолютного нуля. Акцепторная примесь выбирается из соображений, чтобы ее валентные уровни располагались ближе к «потолку» валентной зоны (рис. 1.7, в); все эти уровни пусты при температуре абсолютного нуля. Переходы электронов с донорного уровня в зону проводимости, а из валентной зоны на акцепторный уровень требуют сравнительно небольшой энергии. Поэтому при нагреве число фононов, способных ионизировать донор или акцептор и тем самым создать свободный электрон или свободную дырку, будет гораздо больше числа фононов, способных перевести электрон через всю запрещенную зону и образовать электронно-дырочную пару.

Соответственно, проводимость, обусловленная примесями, оказывается гораздо больше собственной. Однако этот вывод справедлив только до некоторой температуры. Дело в том, что по мере повышения температуры собственная проводимость растет непрерывно, а примесная имеет предел, соответствующий ионизации всех наличных примесных атомов. С повышением температуры увеличивается собственная концентрация электронов и дырок, а следовательно, увеличивается и собственная проводимость полупроводника. При высоких температурах проводимость всегда является собственной. Критерий вырождения примесного полупроводника в собственный будет рассмотрен позже.

Увеличение концентрации примесей приводит к уменьшению расстояния между примесными атомами, и их энергетические уровни постепенно превращаются в зоны. Образовавшаяся примесная зона, расширяясь, в конце концов сливается с бли-

жайшей разрешенной.

Обычно в полупроводниках одновременно присутствуют донорная и акцепторная примеси, рис. 1.8, но чаще — в разных концентрациях ($N_D > N_a$ или $N_a > N_D$).

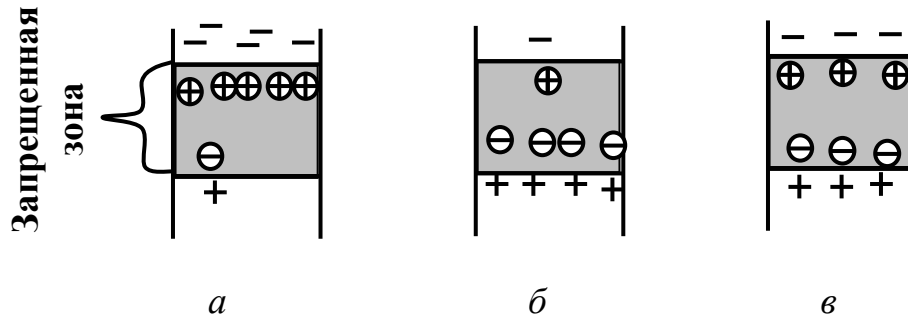


Рис. 1.8. Диаграммы энергетических уровней при различных концентрациях донорных и акцепторных атомов для $T=0$ К°. *а* — превалирует донорная примесь; *б* — превалирует акцепторная примесь; *в* — количество обеих примесей одинаково

Рассматривая этот случай, следует учесть, что при нулевой температуре система должна обладать минимальной энергией.

Пусть $N_D > N_a$, тогда при нулевой температуре (здесь и в дальнейшем будем иметь в виду абсолютный ноль) все акцепторные уровни заполнятся электронами, и число заполненных донорных уровней уменьшится.

При повышении температуры с донорных уровней будет происходить переход электронов в зону проводимости основного полупроводника, так как для перехода электронов с акцепторных уровней требуется гораздо больше энергии.

В дальнейшем под N_D и N_a будем понимать эффективные концентрации. Если количество обоих типов примеси одинаково ($N_D = N_a$), то эффективные концентрации доноров и акцепторов равны нулю. Такой полупроводник называется компенсированным. Он имеет такую же удельную проводимость, как и собственный, но отличается от последнего (по времени жизни носителей и некоторым другим параметрам), поскольку наличие значительного количества примесей нарушает однородность кристаллической решетки.

1.5 Концентрации носителей в зонах полупроводника

Проведем анализ полупроводниковых материалов с тем, чтобы получить качественные и количественные оценки для концентраций электронов и дырок в зависимости от концентраций примесей, изменений температуры и некоторых других внешних воздействий, т.к. величины концентраций свободных электронов и дырок определяют удельную проводимость полупроводника, а следовательно, его электрические свойства.

В теории твердого тела доказано, что энергетические уровни распределены по высоте разрешенной зоны неравномерно: плотность их меняется от границы в глубь зоны. Таким образом, каждому уровню с энергией W соответствует определенная плотность $P(W)$, т.е. число уровней, отнесенное к единице энергии и единице объема материала. Вблизи «дна» и «потолка» разрешенных зон плотность уровней для узких интервалов энергии можно выразить формулой:

$$P(W) = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m^*}{h^2} \right)^{3/2} \sqrt{W - W_{zp}} \left(\frac{\text{уровней}}{\text{дж}^* \cdot \text{см}^3} \right). \quad (1.1a)$$

Здесь h — постоянная Планка; m^* — эффективная масса; W — энергия, которая отсчитывается от граничного уровня внутрь разрешенных зон.

Вероятность нахождения электрона или дырки на том или ином уровне в зависимости от их энергии описывается распределением Ферми-Дирака:

$$F(W) = \frac{1}{e^{kT} + 1}, \quad (1.16)$$

где k — постоянная Больцмана; T — абсолютная температура; W_F — энергия, называемая уровнем Ферми. С формальной точки зрения энергия W_F соответствует такому энергетическому уровню, **вероятность заполнения которого электронами или дырками равна 0.5.**

В статистической физике энергию принято выражать не в джоулях, а в **электрон-вольтах** или **просто в вольтах** (численно

эти величины одинаковы). Чтобы перейти от одной размерности к другой, достаточно разделить энергии W и kT на элементарный заряд.

Сделав такую замену в формулах (1.1), получим:

$$P(\varphi) = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi q m^*}{h^2} \right)^{3/2} \sqrt{\varphi - \varphi_{zp}} \left(\frac{\text{уровней}}{B^* \text{см}^3} \right); \quad (1.2a)$$

$$F_n(\varphi) = \frac{1}{e \frac{\varphi - \varphi_F}{\varphi_T} + 1}, \quad (1.2b)$$

где φ — потенциал, характеризующий энергию уровня; φ_F — **уровень Ферми (потенциал Ферми) в вольтах**; φ_T — **температурный потенциал**

$$\varphi_T = \frac{kT}{q}. \quad (1.3)$$

Название «температурный потенциал» для величины φ_T вполне оправдано, в силу значительной зависимости от температуры.

Полезно запомнить значение φ_T , при температуре $T=300^\circ\text{K}$ (которую принято называть «комнатной») равен

$$\varphi_T(300^\circ\text{K}) = 25 \text{ мВ}.$$

На зонной диаграмме (рис. 1.9) функции $P(\varphi)$ и $F_n(\varphi)$ показаны для собственного полупроводника. Дальнейшие выводы будут справедливы и по отношению к примесным полупроводникам. **В невырожденных полупроводниках уровень Ферми φ_F всегда лежит в запрещенной зоне.** Понятие **невырожденный полупроводник** рассмотрим позже. Глубину залегания уровня Ферми можно характеризовать «расстоянием» от одной из разрешенных зон, выраженным в единицах температурного потенциала.

Для невырожденных полупроводников чаще соблюдаются неравенства:

$$\varphi_c - \varphi_F \gg \varphi_T; \quad (1.4a)$$

$$\varphi_F - \varphi_V \gg \varphi_T, \quad (1.4b)$$

где φ_c и φ_v — потенциалы «дна» зоны проводимости и «потолка» валентной зоны соответственно.

С физической точки зрения температурный потенциал есть выраженная в электрических единицах статистическая температура или близкая к ней средняя кинетическая энергия свободного электрона в электронном газе.

При температуре $T=0^\circ$ К функция $F(\varphi)$ (рис. 1.9) имеет ступенчатый характер, это соответствует уже известным фактам: валентная зона полностью заполнена ($F_n=1$), зона проводимости пуста ($F_n=0$). При температуре $T \neq 0^\circ$ К ступенька функции $F(\varphi)$ сглаживается и получается конечная (хотя и крайне малая) вероятность нахождения электронов в зоне проводимости.

Одновременно вероятность нахождения электронов в валентной зоне делается меньше единицы. В последнем случае удобнее пользоваться вероятностью отсутствия электронов на уровнях или, что то же самое, вероятностью наличия дырок. Из рис. 1.9 следует, что максимальные концентрации электронов в зоне проводимости и дырок в валентной зоне имеют место на границах зон. Плотность состояний на границах зон равна нулю и увеличивается в глубь разрешенных зон.

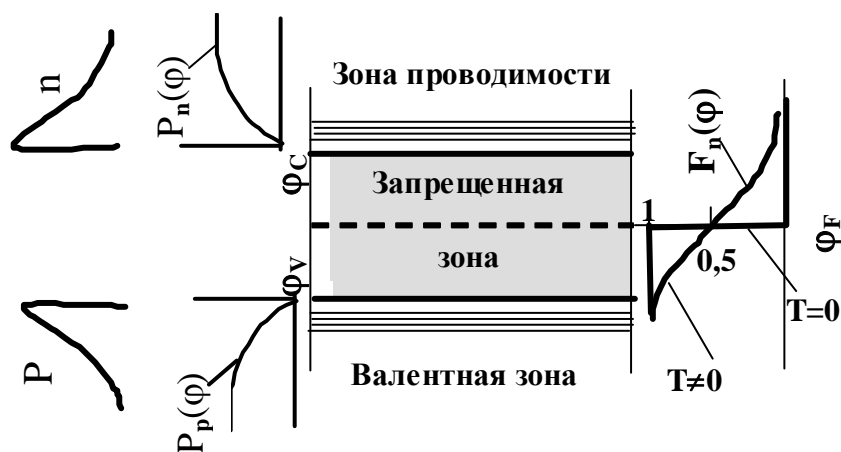


Рис. 1.9 — Плотность уровней энергии, функция вероятности и концентрация носителей в собственном полупроводнике

$$F_p = 1 - F_n = \frac{1}{e^{\frac{\varphi_F - \varphi}{\varphi_T}} + 1} \quad (1.5)$$

Учитывая неравенство (1.4а), можно записать:

$$F_n = e^{-\frac{\varphi - \varphi_F}{\varphi_T}}. \quad (1.6a)$$

Аналогично, учитывая неравенство (1.4б), убеждаемся, что и в валентной зоне, где $\varphi - \varphi_F < 0$, экспонента в выражении (1.5) много больше единицы

$$F_p \approx e^{-\frac{\varphi_F - \varphi}{\varphi_T}}. \quad (1.6б)$$

Функции (1.6), которые являются частным случаем распределения Ферми-Дирака (для области энергий, значительно превышающих φ_F), называются распределением Максвелла–Больцмана.

Это распределение представляет собой основу теории полупроводников, т.к. при этом существенно упрощается количественный анализ полупроводниковых материалов и приборов на их основе.

Концентрация свободных электронов в зоне проводимости с учетом проведенного выше анализа запишется в виде:

$$n = 2 \int_{\varphi_c}^{\infty} P(\varphi - \varphi_c) F_n(\varphi) d\varphi,$$

где подынтегральное выражение есть количество заполненных уровней в элементарном интервале энергии, а множитель 2 означает, что на каждом уровне могут (по принципу Паули) находиться два электрона. Подставив (1.2а) и (1.6а) под знак интеграла, после преобразований получим:

$$n = N_c e^{-\frac{\varphi_c - \varphi_F}{\varphi_T}}; \quad (1.7a)$$

$$N_c = 2 \left(\frac{2\pi m_n q \varphi_T}{h^2} \right)^{3/2},$$

где N_c — эффективная плотность состояний (в 1 см^3) в зоне проводимости. Из (1.7а) следует, что N_c есть **максимально возможная концентрация электронов в невырожденном полупроводнике.**

Концентрация свободных дырок в валентной зоне определяется выражением:

$$p = 2 \int_{\Phi_V}^{-\infty} P[-(\Phi - \Phi_V)] F_p(\Phi) (-d\Phi).$$

Подставив (1.2а) и (1.6б), после преобразований получим:

$$p = N_V e^{-\frac{\Phi_F - \Phi_V}{\Phi_T}}; \quad (1.76)$$

$$N_V = 2 \left(\frac{2\pi m_p q \Phi_T}{h^2} \right)^{3/2},$$

где N_V — эффективная плотность состояний в валентной зоне.

Легко убедиться, что произведение концентраций np не зависит от положения уровня Ферми и определяется только температурой и шириной запрещенной зоны:

$$np = N_C N_V e^{-\frac{\Phi_Z}{\Phi_T}}, \quad (1.8)$$

где $\Phi_Z = \Phi_C - \Phi_V$ — ширина запрещенной зоны.

Из выражения (1.8) следует, учитывая, что эффективные плотности состояний относительно слабо зависят от типа полупроводникового материала при постоянной температуре, произведение концентраций зависит в основном от ширины запрещенной зоны, т.е. характеристики полупроводника. Для кремния ширина запрещенной зоны примерно в два раза больше чем у германия, в результате чего произведение концентраций электронов — дырок у кремния на три порядка меньше. Следовательно, электрические параметры и характеристики кремния и германия должны существенно отличаться.

Из формул (1.7) отношение концентраций получается в следующем виде:

$$\frac{n}{p} = \frac{N_C}{N_V} e^{-2(\Phi_E - \Phi_F)/\Phi_T}; \quad (1.9a)$$

$$\Phi_E = \frac{\Phi_C + \Phi_V}{2},$$

где Φ_E — потенциал середины запрещенной зоны, который иногда называют электростатическим потенциалом полупроводника. Если принять, что выполняется условие $N_C = N_V$, выражение (1.9а), можно записать в виде:

$$\frac{n}{p} = e^{-\frac{2(\varphi_E - \varphi_F)}{\varphi_T}}. \quad (1.96)$$

Анализируя выражение (1.9а), можно сделать следующие выводы:

1. Если уровень Ферми совпадает с электростатическим потенциалом, тогда $\frac{n}{p} = 1$, т.е. полупроводник является собственным.

2. При выполнении условия $\varphi_F > \varphi_E$ $\frac{n}{p} > 1$, т.е. полупроводник электронный.

3. При выполнении условия $\varphi_F < \varphi_E$ $\frac{n}{p} < 1$, т.е. полупроводник дырочный.

4. Из выражения (1.9б) следует, что изменение уровня Ферми должно приводить к изменению концентрации примесей. Физически можно изменять концентрацию электронов или дырок, вводя разное количество примесей в основной полупроводник, а уровень Ферми рассчитывать.

1.6 Уровень Ферми

При анализах, которые мы провели выше, считалось, что уровень Ферми нам известен, и с его помощью вычислялись концентрации свободных носителей заряда. На самом же деле уровень Ферми является функцией этих концентраций, так как изменить уровень Ферми можно только, изменив концентрацию доноров или акцепторов, а концентрации носителей оценить из тех или иных соображений или условий.

Используя выражения (1.2б) и (1.5), интегралы легко привести к виду

$$\frac{2}{\sqrt{\tau}} \int_0^{\infty} \frac{\sqrt{\eta} d\eta}{e^{\eta - (\chi/\varphi_T)} + 1} = v, \quad (1.10)$$

где в случае электронного и дырочного полупроводников соответственно запишем:

$$\eta_n = (\varphi - \varphi_C) / \varphi_T; \chi_n = \varphi_F - \varphi_C; v = \frac{n}{N_C};$$

$$\eta_p = (\varphi_V - \varphi) / \varphi_T; \chi_p = \varphi_V - \varphi_F; v = \frac{p}{N_V}.$$

Величину χ , входящую в интегральное уравнение (1.10), в статистической физике называют химическим потенциалом.

Химический потенциал является функцией концентрации соответствующих частиц. О наличии химического потенциала можно говорить только в том случае, если концентрации электронов и дырок в полупроводнике различные, т.е. полупроводник примесный. При равенстве концентраций электронов и дырок полупроводник собственный и химический потенциал равен нулю. В компенсированном полупроводнике имеет место наличие химических потенциалов дырок и электронов, но они равны по абсолютной величине и направлены навстречу друг другу, в результате чего результирующий химический потенциал равен нулю. **Разность химических потенциалов означает наличие разности концентраций электронов или дырок в разных объемах полупроводника (полупроводников), что, естественно, вызывает перемещение — диффузию частиц в направлении от большей концентрации к меньшей.** Таким образом, разность химического потенциала характеризует возможность диффузии сводных частиц (заряженных или незаряженных), подобно тому, как электрический потенциал характеризует возможность дрейфа свободных заряженных частиц. Потенциал Ферми можно записать в виде алгебраической суммы электрического и химического потенциалов:

$$\varphi_F = \varphi_C + \chi_n; \quad (1.11a)$$

$$\varphi_F = \varphi_V - \chi_p. \quad (1.11b)$$

Отсюда следует еще одно название величины φ_F — электрохимический потенциал. **Градиент потенциала Ферми, будучи суммой градиентов электрического и химического потенциалов, позволяет одновременно характеризовать оба типа движения носителей — диффузию и дрейф.**

В условиях равновесия, когда направленного движения носителей нет, должно выполняться условие $grad \varphi_F = 0$, т.е.

$$\varphi_F = const.$$

Постоянство («горизонтальность») уровня Ферми в равновесной системе является одним из фундаментальных соотношений теории твердого тела. Заметим, что условие $\varphi_F = const$ не означает постоянства каждого из слагаемых φ_c и φ_F . Иначе говоря, в равновесной системе могут иметь место градиенты электрического и химического потенциалов и соответственно дрейфовые и диффузионные потоки носителей. Эти потоки должны взаимно уравновешиваться, т.е. например, диффузионный поток электронов равен дрейфовому потоку и потоки направлены навстречу друг другу.

Для определения химических потенциалов χ_n , χ_p и потенциала Ферми через концентрации n и p нужно решить интегральное уравнение (1.10). В общем виде аналитического решения интеграла нет, однако в двух частных случаях, принципиально важных для практики, получаются аналитические решения.

1. Положим $\chi < 0$ и $|\chi| \gg \varphi_T$. Тогда, пренебрегая единицей в знаменателе подынтегрального выражения, т.е. используя распределение Максвелла–Больцмана, получаем уравнение

$$e^{\chi/\varphi_T} = v,$$

решением которого будут химические потенциалы:

$$\chi_n = \varphi_T \ln \frac{n}{N_C}; \quad \chi_p = \varphi_T \ln \frac{p}{N_V}. \quad (1.12)$$

Учитывая, что решение уравнения получено при условии $\chi < 0$, приходим к выводу, что полученные решения справедливы при условии $v \ll 1$. Полупроводники, у которых выполняется это условие, т.е. концентрация свободных носителей меньше эффективной плотности состояний, называют **невырожденными**.

Потенциал Ферми для невырожденных полупроводников можно записать в виде:

$$\varphi_F = \varphi_C + \varphi_T \ln \frac{n}{N_C}; \quad (1.13a)$$

$$\varphi_F = \varphi_V - \varphi_T \ln \frac{p}{N_V}. \quad (1.13b)$$

Из выражений (1.13а) и (1.13б) следует, что потенциал Ферми в невырожденных полупроводниках лежит в запрещенной зоне.

Вычитая или складывая выражения (1.13), легко получить соответственно выражения (1.8) и (1.9). Также легко видеть, что у невырожденных полупроводников потенциал Ферми всегда лежит в запрещенной зоне, поскольку логарифмы в обоих выражениях (1.13) отрицательны.

2. Положим $\chi > 0$ и $|\chi| \gg \phi_T$. Тогда при выполнении условия $\eta > \chi / \phi_T$ подынтегральное выражение быстро приближается к нулю, и интегрирование в этом диапазоне не имеет смысла. Поэтому примем в качестве верхнего предела интегрирования $\eta = \chi / \phi_T$. Выполнив интегрирование, получим простое уравнение

$$\chi_n = \left(\frac{3\sqrt{\pi}}{4} \right)^{2/3} \phi_T \left(\frac{n}{N_C} \right)^{2/3}; \quad \chi_n = \left(\frac{3\sqrt{\pi}}{4} \right)^{2/3} \phi_T \left(\frac{n}{N_C} \right)^{2/3}.$$

Эти решения действительны практически при $\eta \geq 3$.

Полупроводники, у которых соблюдается условие $v > 1$, т.е. концентрация свободных носителей существенно превышает эффективную плотность состояний в разрешенной зоне, называют вырожденными или полуметаллами.

Для них распределение Максвелла–Больцмана недействительно.

Критерии вырождения можно записать в виде

$$n > N_C; \quad (1.14а)$$

$$p > N_V. \quad (1.14б)$$

В неравновесном состоянии значения потенциалов Ферми в (1.13а) и (1.13б), вообще говоря, различны, т.е. происходит расщепление потенциала Ферми, тогда их называют соответственно квазиуровнями Ферми для электронов и дырок.

1.7 Концентрация носителей в полупроводниках

Рассмотрим концентрации носителей для различных типов полупроводников.

В собственном полупроводнике концентрации свободных электронов и дырок одинаковы: $n_i = p_i$. Тогда из формул (1.9) и (1.13) следует, что при любой температуре уровень Ферми собственного полупроводника расположен вблизи середины запрещенной зоны. Подставляя $n_i = p_i$ в формулу (1.8), получаем концентрации свободных электронов и дырок в собственном полупроводнике:

$$n_i = p_i = \sqrt{N_C N_V} e^{-\frac{\varphi_3}{\varphi_T}} \approx N_C e^{-\varphi_3 / \varphi_T} \approx N_V e^{\varphi_3 / \varphi_T}. \quad (1.15)$$

Зависимость собственных концентраций n_i и p_i от температуры очень сильна и обусловлена в основном величиной температурного потенциала, т.е. зависит от энергии кристаллической решетки. Изменение энергии кристаллической решетки приводит к изменению колебаний атомов в узлах кристаллической решетки, следовательно, изменяется вероятность нахождения электронов (дырок) на уровнях в разрешенной зоне.

Столь же сильно зависит собственная концентрация от ширины запрещенной зоны при постоянной температуре. Так, сравнительно небольшое различие в величине φ_3 у германия и кремния (0,67 эВ и 1,11 эВ) приводит к различию собственных концентраций при комнатной температуре на 3 и более порядков. Сравнивая (1.8) и (1.15), соотношение (1.8) можно записать в более компактной форме:

$$np = n_i^2. \quad (1.16)$$

Соотношение (1.16) говорит о том, что увеличение концентрации одного типа носителей сопровождается уменьшением концентрации другого типа носителей. Из выражения (1.16) следует, что в полупроводниках с различной шириной запрещенной зоны, а следовательно, с различной собственной концентрацией, при одинаковой концентрации электронов (дырок) концентрация дырок (электронов) будет различной.

Рассмотрим сказанное на конкретном примере.

Используя выражение (1.16), запишем очевидное соотношение для полупроводников из германия и кремния: $\frac{n_{\Gamma} p_{\Gamma}}{n_K p_K} = \frac{n_{i\Gamma}^2}{n_{iK}^2}$.

Используя соотношение (1.15), можно записать

$$\frac{n_{\Gamma} p_{\Gamma}}{n_K p_K} = e^{\frac{\varphi_{3\Gamma}}{2\varphi_{\Gamma}}} \approx 10^6.$$

В кремниевом полупроводнике выполним условие $n_K = 10p_K$, т.е. кремниевый полупроводник является ярко выраженным электронным. Потребуем, чтобы концентрация электронов в германиевом и кремниевом полупроводниках была одинаковой. При выполнении данного условия можно записать: $\frac{p_{\Gamma}}{n_{\Gamma}} = 10^5$, т.е. германиевый полупроводник должен быть дырочным.

Используя формулы (1.16) и (1.7) и полагая $N_c = N_v$, трудно выразить концентрации n и p через собственную концентрацию n_i

$$n = n_i e^{-\varphi_E - \varphi_F / \varphi_T}; \quad (1.17a)$$

$$p = n_i e^{-\varphi_F - \varphi_E / \varphi_T}. \quad (1.17b)$$

Отсюда легко получить потенциал Ферми в двух формах

$$\varphi_F = \varphi_E + \varphi_T \ln \frac{n}{n_i}; \quad (1.18a)$$

$$\varphi_F = \varphi_E + \varphi_T \ln \frac{n}{n_i}. \quad (1.18b)$$

Для того чтобы определить потенциал φ_F по формулам (1.13) или (1.18), нужно знать концентрации свободных носителей.

На рис. 1.10 и 1.11 приведены зонные диаграммы для электронного и дырочного полупроводников.

При оценке величин n и p используют условие нейтральности полупроводника.

Это важное условие формулируется следующим образом: в однородном полупроводнике не может быть существенных не скомпенсированных объемных зарядов ни в равновесном состоянии, ни при наличии тока.

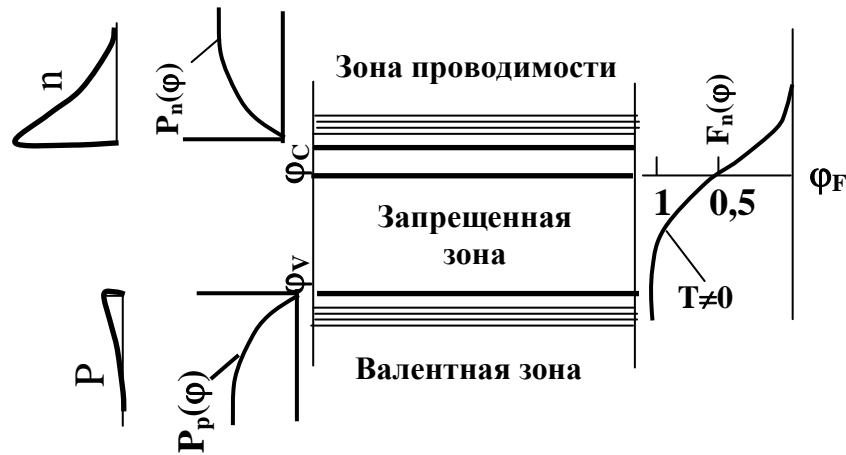


Рис. 1.10 — Плотность уровней энергии, функция вероятности и концентрации носителей в примесном полупроводнике n типа

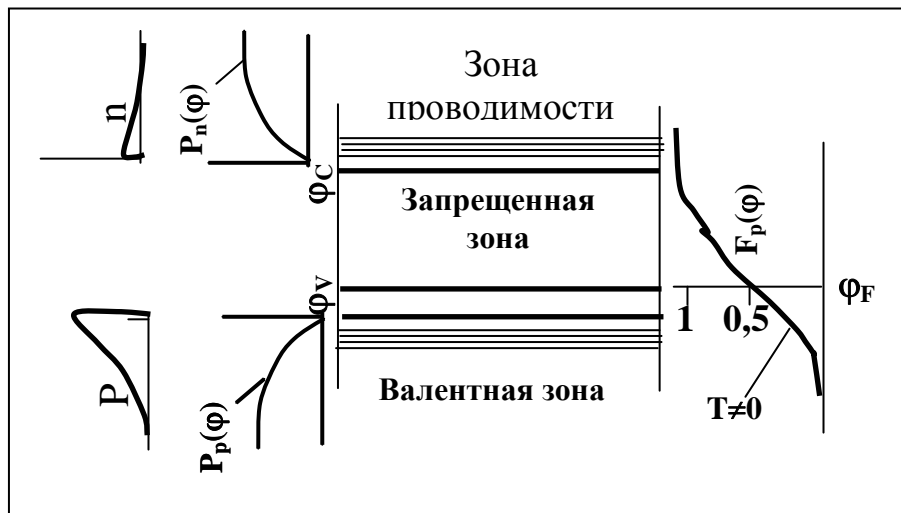


Рис. 1.11 — Плотность уровней энергии, функция вероятности и концентрации носителей в примесном полупроводнике p типа

В общем, виде условие нейтральности для единичного объема можно записать в виде

$$p + N_D^* - (n + N_a^*) = 0, \quad (1.19a)$$

где N_D^* и N_a^* — концентрации ионизированных доноров и акцепторов. Уравнение (1–19а) говорит о том, что концентрация частиц, несущих положительный заряд (дырки и ионизированные доноры), равна концентрации электронов и ионизированных акцепторов.

Для электронных полупроводников, не содержащих акцепторов,

$$n = N_D^* + p. \quad (1.19б)$$

Для дырочных полупроводников, не содержащих доноров,

$$p = N_a^* + n. \quad (1.19в)$$

Перейдем непосредственно к оценке концентраций свободных носителей в разных типах полупроводников. Рассмотрим сначала электронные полупроводники. В типичном электронном полупроводнике выполняется неравенство $n > p$. Кроме того, будем считать, что в рабочем диапазоне температур донорные атомы практически полностью ионизированы, т.е. $N_D^* = N_D$. Допустим, что собственной концентрацией электронов в силу того, что их концентрация много меньше примесных, можно пренебречь. Тогда из соотношения (1.19б) получаем концентрацию свободных электронов

$$n = N_D. \quad (1.20а)$$

Как видим из 1.20а, концентрация электронов определяется концентрацией примеси и, следовательно, не зависит от температуры. С помощью (1.16) легко получаем концентрацию свободных дырок, неосновных носителей, которая зависит от температуры в силу сильной зависимости от температуры собственной концентрации электронов (дырок).

$$p = \frac{n_i^2}{N_D}. \quad (1.20б)$$

Наконец, из (1.18а) или (1.13а) находим уровень Ферми в типичном электронном полупроводнике:

$$\varphi_F = \varphi_E + \varphi_T \ln \frac{N_D}{n_i} = \varphi_C + \varphi_T \ln \frac{N_D}{N_C}. \quad (1.20в)$$

Выражение (1.20в) позволяет сделать вывод, что уровень Ферми лежит тем выше, чем больше концентрация доноров и чем ниже температура. Вывод можно считать справедливым до тех пор, пока снижение температуры не приводит к уменьшению числа ионизированных доноров, т.к. при $T = 0^0 K$ уровень Ферми равен $\varphi_F = \varphi_E$. Увеличение температуры приводит к росту собственной концентрации носителей, которая становится сравнимой с концентрацией примесей, затем и превышает её. Проводимость становится преимущественно собственной. С изменением типа

проводимости значительно изменяются свойства полупроводниковых материалов и приборов, выполненных на их основе. Температура, при которой проводимость становится преимущественно собственной, определяет допустимую температуру работы полупроводникового прибора.

Необходимо помнить, что выражение (1.20в) получено при допущении $N_D \gg n_i$.

Зависимость от температуры обусловлена в основном φ_T , т.к. отношение концентрации доноров к эффективной плотности состояний остается практически постоянным. Простые формулы (1.20) широко используются на практике. Однако следует иметь в виду, что они действительны в ограниченном температурном диапазоне: с понижением температуры степень ионизации доноров уменьшается и принятое равенство $N_D = N_D^*$ становится менее строгим.

Случай дырочного полупроводника, в котором $p > n$, нет необходимости рассматривать столь же подробно. Главные особенности акцепторной примеси видны из рис. 1.11. Если акцепторы полностью ионизированы и температура ниже критической, то вместо формул (1.20) получаем аналогичные соотношения, характерные для ярко выраженного дырочного полупроводника:

$$p = N_a; \quad (1.21a)$$

$$n = \frac{n_i^2}{N_a}; \quad (1.21б)$$

$$\varphi_F = \varphi_E - \varphi_T \ln \frac{N_a}{n_i} = \varphi_V - \varphi_T \ln \frac{N_a}{N_V}. \quad (1.21в)$$

Весь предыдущий анализ и зонные диаграммы соответствовали однородным полупроводникам, в которых примеси распределены равномерно. Разумеется, однородный полупроводник является некоторой идеализацией. Более того, часто специально создают неоднородность внутри кристалла в виде градиента концентрации примесей, что придает полупроводнику новые свойства, необходимые для работы ряда приборов. Посмотрим, каковы особенности неоднородных полупроводников. Пусть, например, в полупроводнике типа n концентрация доноров изменяется от N_{D1} до $N_{D2} < N_{D1}$.

Поскольку в равновесной системе уровень Ферми во всех ее частях одинаков, зонная диаграмма должна иметь вид, как показано на рис. 1.12, *а*. Учитывая, что концентрация доноров, а следовательно, и концентрация электронов слева выше, уровень Ферми у левой границе полупроводника должен быть ближе к «дну» зоны проводимости. Это возможно только при искривлении зон, как показано на рис. 1.12, *а*. Электростатический потенциал φ_E вдоль неоднородного полупроводника меняется, т.к. по определению это потенциал середины запрещенной зоны. Следовательно, в неоднородных полупроводниках имеются внутренние электрические поля, в которых возможен дрейф носителей. Однако в отсутствие внешнего поля дрейфовые потоки носителей равны противоположно направленным диффузионным потокам, которые обусловлены градиентом концентрации тех же носителей. Поэтому результирующий поток отсутствует и соблюдается больцмановское равновесие.

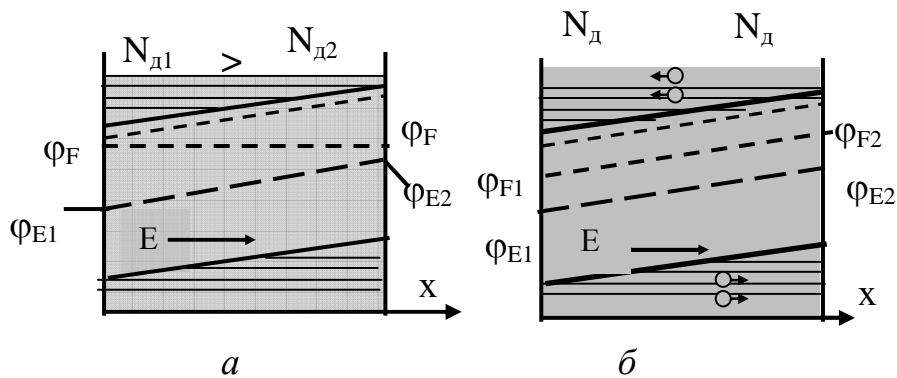


Рис 1.12 — Зонные диаграммы неоднородного полупроводника (*а*) и однородного полупроводника при наличии внешнего электрического поля (*б*)

Для сравнения на рис. 1.12, *б* показана зонная диаграмма однородного полупроводника при наличии внешнего электрического поля (напряженность $E = \frac{d\varphi_E}{dx}$ та же, что и на рис. 1.12, *а*). Если в силу условия квазинейтральности принять концентрации носителей неизменными вдоль оси x , то будут неизменными и химические потенциалы, т.е. «расстояния» уровня φ_F от краев разрешенных зон. Тогда согласно (1.18) имеет место градиент уровня

Ферми, обусловленный нарушением равновесия (внешнее электрическое поле вызывает протекание тока). Наличие градиента потенциала Ферми обуславливает принципиальное отличие данного случая от предыдущего (рис. 1.12, *a*), и прежде всего во втором результирующий ток не равен нулю, несмотря на внешнее сходство «перекошенных» зонных диаграмм. Заметим, что наличие внутреннего электрического поля в общем случае не означает нарушения условия квазинейтральности, так как постоянное или почти постоянное поле не связано с существенными объемными зарядами.

Все сказанное действительно и для полупроводника типа *p* с учетом расположения уровня Ферми.

1.8 Подвижность носителей и удельная проводимость

Удельная проводимость любого тела зависит не только от концентрации носителей, но и от их подвижности в электрическом поле. Подвижность носителей по определению есть их средняя направленная скорость в электрическом поле с напряженностью 1 В/см. Соответственно дрейфовую скорость можно записать в виде $v_{др} = \mu E$. Постоянство дрейфовой скорости носителей в однородном поле ($E = const$) специфично для твердого тела, где имеются различные «препятствия» движению. В вакууме, где таких препятствий нет, движение заряженных частиц в однородном поле равноускоренное, т.е. понятие дрейфовой скорости отсутствует.

Направленное движение носителей в твердом теле под действием поля сочетается с их хаотическим (тепловым) движением.

$$V_T = \sqrt{3kT / m^*}. \quad (1.22)$$

Из формулы (1.22) видно, что тепловая скорость зависит от температуры.

При слабых электрических полях дрейфовая скорость намного меньше тепловой или, как говорят, температура носителей определяется температурой кристаллической решетки. В этих условиях подвижность можно выразить формулой:

$$\mu = \frac{q}{m^*} t_{CP} = \frac{q}{m^*} \frac{l_{CP}}{V_T}, \quad (1.23)$$

где t_{CP} — среднее время свободного пробега частицы;

l_{CP} — средняя длина ее свободного пробега.

Величины t_{CP} и l_{CP} характеризуют частоту столкновений носителей с теми или иными «препятствиями». В результате таких столкновений происходит изменение скорости и направления движения носителей, т.е. их рассеяние. При нулевой абсолютной температуре в идеальной кристаллической решетке рассеяние не имеет места. Иначе говоря, атомы решетки как таковые не являются препятствиями на пути движения носителей. Истинными препятствиями являются лишь колеблющиеся атомы решетки (в корпускулярной интерпретации — фононы), а также атомы примесей и дефекты структуры. В области очень низких температур интенсивность тепловых колебаний мала, количество фононов незначительно и рассеяние обусловлено главным образом дефектами решетки и нейтральными примесями (нейтральность примесей является следствием низкой температуры, меньшей температуры ионизации).

Из выражений 1.22 и 1.23 следует, что величина подвижности сильно зависит от температуры кристалла, т.е. от тепловой скорости, которая приводит к хаотическому движению заряженных частиц.

С уменьшением напряженности поля дрейфовая составляющая тока, которая является информационной, будет уменьшаться, что приводит к увеличению случайной составляющей тока. Если дрейфовая составляющая скорости становится сравнимой с тепловой скоростью, выделить информационную составляющую тока становится сложно. Обычно в радиотехнике тепловую составляющую тока называют шумовой, которая определяет предельную чувствительность электронных устройств.

В области рабочих температур рассеяние обусловлено главным образом фононами при ионизации примесей. В этой области подвижность можно представить состоящей из двух компонентов μ_l и μ_L , связанных с результирующей подвижностью простой зависимостью

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I}. \quad (1.24)$$

Очевидно, что результирующая подвижность определяется наименьшим из слагаемых. Тогда из (1.23) с учетом (1.22) следует соотношение

$$\mu_L \cong (m^*)^{-5/2} T^{-3/2}. \quad (1.25a)$$

При ионном рассеянии получается соотношение

$$\mu_L \cong (m^*)^{-1/2} T^{-3/2} N^{-1}, \quad (1.25b)$$

где N — концентрация ионизированной примеси

Если воспользоваться общепринятым обозначением

$$b = \frac{\mu_n}{\mu_p}, \quad (1.26)$$

то получим для кремния $b=2,8$; германия $b=2,1$.

Соотношения (1.25) позволяют исследовать зависимость подвижности от температуры и концентрации примеси.

Зависимость $\mu(E)$ оказывает непосредственное влияние на вольт-амперную характеристику полупроводника. Например, в простейшем случае (однородный кристалл с постоянным сечением) ток I пропорционален скорости, а напряжение U — напряженности электрического поля. Поэтому форма кривой $I(U)$ такая же, как кривой $V(E)$ (рис. 1.13).

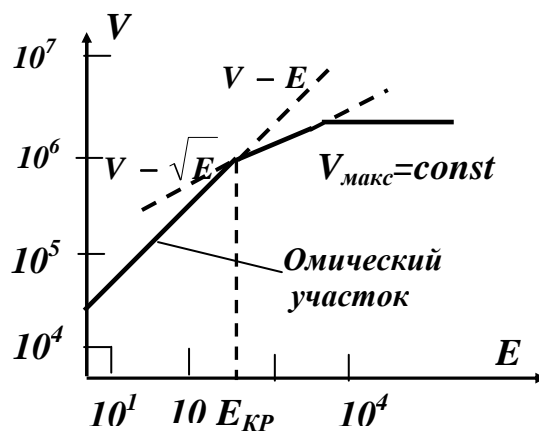


Рис. 1.13 — Зависимость скорости носителей от напряженности электрического поля

В общем виде удельную проводимость можно выразить формулой:

$$\sigma = \frac{1}{\rho} = qn\mu_n + qp\mu_p, \quad (1.27)$$

где μ_n и μ_p — подвижности электронов и дырок; ρ — удельное сопротивление.

Для частных случаев собственного, электронного и дырочного полупроводников получаем соответственно:

$$\sigma_i = \frac{1}{\rho_i} = qn_i(\mu_n + \mu_p); \quad (1.28)$$

$$\sigma_n = \frac{1}{\rho_n} \approx qn_n\mu_n; \quad (1.29)$$

$$\sigma_p = \frac{1}{\rho_p} \approx qp_p\mu_p. \quad (1.30)$$

Температурную зависимость удельной проводимости или удельного сопротивления можно получить, зная температурную зависимость концентрации и их подвижности.

На рис. 1.14 показаны зависимости удельных сопротивлений разных типов полупроводников при изменении температуры.

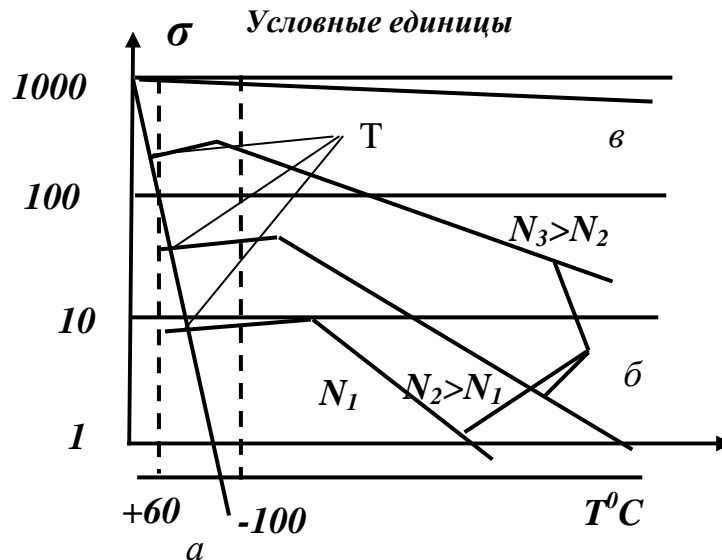


Рис. 1.14 — Зависимость удельной проводимости полупроводника от температуры:
а — собственный полупроводник; *б* — примесные полупроводники; *в* — полуметалл

При большой концентрации примесей полупроводник превращается в полуметалл с высокой проводимостью, мало зависящей от температуры. На рис. 1.14 для наглядности показана зависимость удельного сопротивления от температуры в линейном масштабе.

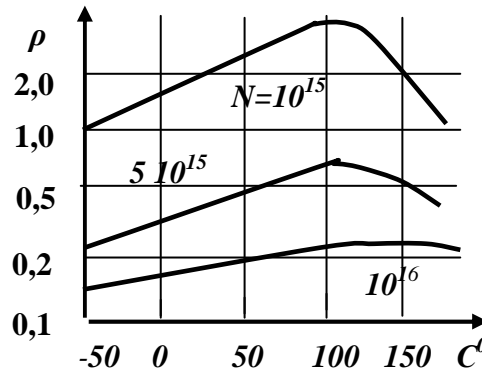


Рис. 1.15 — Зависимость удельного сопротивления *n*-германия от температуры при различных концентрациях доноров

1.9 Генерация и рекомбинация носителей в полупроводниках

Процессы генерации и рекомбинации носителей заряда неотъемлемы друг от друга и в то же время противоположны по содержанию. Генерация является ведущим началом в этом единстве и связана с воздействием таких внешних факторов, как нагрев, освещение или облучение. Рекомбинация представляет собой внутреннюю реакцию системы на появление и возрастание числа носителей. Именно рекомбинация, противодействуя накоплению носителей, обуславливает их равновесные концентрации. Поэтому изучение механизма рекомбинации и ее количественных закономерностей необходимо для понимания и использования многих важнейших явлений в полупроводниках и полупроводниковых приборах.

Непосредственная рекомбинация свободного электрона со свободной дыркой — сравнительно редкое событие. Время жизни носителей, вычисленное исходя из непосредственной рекомбинации, на несколько порядков больше наблюдаемых величин. По-

этому главную роль следует отвести механизму рекомбинации с помощью центров рекомбинации, которые часто называют «ловушками». Напомним, что ловушка описывается совокупностью энергетических уровней, расположенных глубоко в запрещенной зоне, близко к ее середине. Переход электрона из зоны проводимости на уровень ловушки и затем в валентную зону гораздо более вероятен, чем непосредственный переход через запрещенную зону при непосредственной рекомбинации.

На рис. 1.16 показаны две возможные последовательности процессов при рекомбинации на ловушках. Поскольку ловушка в равной степени облегчает переход электронов как из зоны проводимости в валентную зону, так и в обратном направлении, она представляет собой не только центр рекомбинации, но и центр генерации носителей, как и должно быть из общих соображений о равновесии. С физической точки зрения понятие ловушек столь же широко, как и понятие примесей; это могут быть и посторонние атомы, и различные подвижные частицы, и дефекты кристаллической решетки. Для изготовления германиевых и кремниевых диодов и транзисторов, как правило, используют монокристаллы с регулярной структурой, у которых время жизни обычно лежит в пределах 10–100 мксек. До сих пор мы имели в виду рекомбинацию в объеме полупроводника и объемное время жизни. Однако не менее, а часто более существенное значение имеет рекомбинация на поверхности.

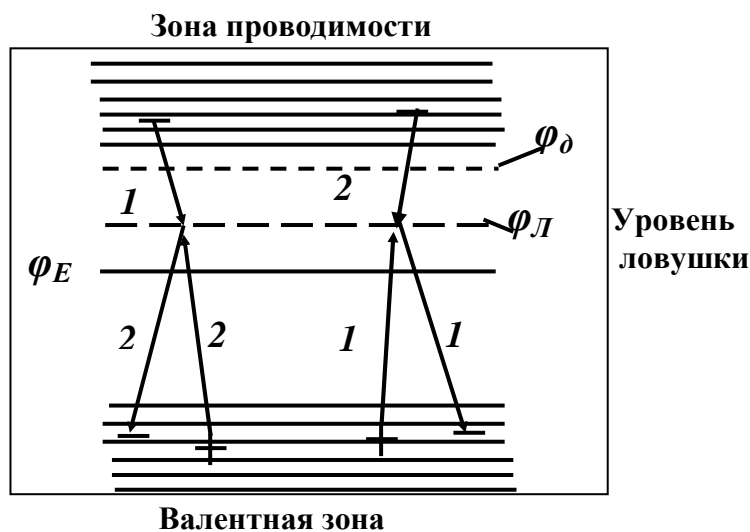


Рис. 1.16 — Возможные варианты рекомбинации носителей на ловушках 1 и 2

Это объясняется тем, что дефекты решетки неизбежны в первую очередь на поверхности кристалла, где нарушена симметрия связей атомов, а также наиболее вероятно наличие разного рода пленок. Поэтому поверхность полупроводника представляет собой особую, весьма активную область, содержащую большое число энергетических уровней. Поверхностные уровни, расположенные в запрещенной зоне, могут играть роль ловушек, и тогда поверхность становится областью интенсивной рекомбинации и генерации носителей. Обычно эти поверхностные процессы характеризуются не временем жизни, а скоростью поверхностной рекомбинации.

Для анализа и расчета полупроводниковых приборов чаще всего используют единый параметр — так называемое эффективное время жизни τ , которое характеризует совместное влияние объемной и поверхностной рекомбинаций и определяется соотношением

$$\frac{1}{\tau} = \frac{1}{\tau_V} + \frac{1}{\tau_S}. \quad (1.31)$$

1.10 Плотность тока в полупроводниках

В общем случае движение носителей заряда в полупроводниках обусловлено двумя процессами: диффузией под действием градиента концентрации (разностью химического потенциала) и дрейфом под действием градиента потенциала в электрическом поле. Поскольку в полупроводниках мы имеем дело с двумя типами носителей — дырками и электронами, полный ток состоит из четырех составляющих:

$$j = (j_p)_{диф} + (j_p)_{др} + (j_n)_{диф} + (j_n)_{др}, \quad (1.32)$$

где индексы «диф.» и «др.» относятся соответственно к диффузионной и дрейфовой составляющим токов. Плотности дрейфовых составляющих тока пропорциональны градиенту электрического потенциала φ , т.е. напряженности электрического поля E . В одномерном случае, когда движение носителей происходит только вдоль оси x , без отклонения в стороны, что является приближением, имеем:

$$(j_p)_{\partial p} = -q\rho\mu_p \frac{\partial\varphi}{\partial x} = q\rho\mu_p E; \quad (1.33a)$$

$$(j_n)_{\partial p} = -qn\mu_p \frac{\partial\varphi}{\partial x} = qn\mu_n E. \quad (1.33b)$$

Следует заметить, что в обычном соотношении φ , использованном в формулах (1.33), потенциал φ относится к положительным зарядам, тогда как на зонных диаграммах величина φ характеризует энергию отрицательных зарядов — электронов.

Поэтому для зонных диаграмм действительно соотношение $E = d\varphi/dx$, т.е. положительной напряженности поля соответствует положительный потенциал.

При анализе обычно удобнее пользоваться не токами I , а плотностями токов j , что и сделано в формуле (1.33). Там, где это не вызывает недоразумений, мы будем называть величину j током.

Плотности диффузионных составляющих тока пропорциональны градиентам химических потенциалов χ_n и χ_p , которые для невырожденных полупроводников определяются формулами (1.12). Поэтому в одномерном случае имеем:

$$(j_p)_{\text{диф}} = -q\rho\mu_p \frac{d\chi_p}{dx} = -qD_p \frac{dp}{dx}; \quad (1.34a)$$

$$(j_n)_{\text{диф}} = -qn\mu_n \frac{d\chi_n}{dx} = -qD_p \frac{dn}{dx}. \quad (1.34b)$$

Здесь D_n , D_p — коэффициенты диффузии дырок и электронов, связанные с подвижностями тех же носителей формулой Эйнштейна:

$$D = \varphi_T \mu. \quad (1.35)$$

Из выражений (1.34) следует, что в невырожденных полупроводниках, для которых действительны использованные значения химических потенциалов, диффузионные токи пропорциональны градиенту концентраций носителей, а коэффициенты диффузии не зависят от этих концентраций. Подставляя (1.34) и (1.33) в формулу (1.32), получаем зависимость плотности полного тока:

$$j = -qD_p \frac{\partial p}{\partial x} + q\mu_p E + qD_n \frac{\partial n}{\partial x} + q\mu_n E. \quad (1.36)$$

Как видим, для определения тока необходимо знать концентрации носителей, их распределение и напряженность поля. В общем случае концентрации p и n зависят от двух переменных: координаты x и времени t . Поэтому для определения токов нужно предварительно найти функции $p(x,t)$ и $n(x,t)$. Эти функции являются решениями так называемых уравнений непрерывности потока, которым в любой момент времени подчиняется движение носителей.

Знак минус в формуле (1.34а) имеет следующий физический смысл: диффузия всегда происходит в направлении убывания концентрации, а поскольку дырки несут положительный заряд, ток $(j_p)_{\text{диф}}$ должен быть положительным при $dp/dx < 0$.

Для дырок и электронов уравнения непрерывности записываются в следующем виде:

$$\frac{\partial p}{\partial t} = \Delta g_p - \frac{p - p_0}{\tau_p} - \frac{1}{q} \operatorname{div}(j_p); \quad (1.37a)$$

$$\frac{\partial n}{\partial t} = \Delta g_n - \frac{n - n_0}{\tau_n} - \frac{1}{q} \operatorname{div}(j_n), \quad (1.37b)$$

где $p - p_0 = \Delta p$ и $n - n_0 = \Delta n$ — избыточные концентрации; Δg_p и Δg_n — скорости генерации под действием внешних факторов, например света, τ_p и τ_n — времена жизни дырок и электронов.

Слагаемые в правых частях (1.37) соответствуют возможным причинам изменения концентрации носителей во времени. В частности, последние слагаемые можно рассматривать как скорости накопления или рассасывания носителей, связанные с неравенством потоков, втекающих и вытекающих из некоторого элементарного объема. Такой небаланс потоков характеризуется дивергенцией вектора плотности потока. В нашем случае плотность потока есть j/q .

Дивергенция этого вектора для одномерного случая равна:

$$\operatorname{div} \frac{j}{q} = \frac{1}{q} \frac{\partial}{\partial x} [(j)_{\text{ДИФ}} + (j)_{\text{ДР}}].$$

Подставляя сюда соотношения (1.33) и (1.34), получаем:

$$\begin{aligned} \frac{1}{q} \operatorname{div}(j_p) &= -D_p \frac{\partial^2 p}{\partial x^2} + \mu_p E \frac{\partial p}{\partial x} + p \mu_p \frac{\partial E}{\partial x}; \\ \frac{1}{q} \operatorname{div}(j_n) &= -D_n \frac{\partial^2 n}{\partial x^2} + \mu_n E \frac{\partial n}{\partial x} + n \mu_n \frac{\partial E}{\partial x}. \end{aligned}$$

С учетом этих выражений, а также при отсутствии внешних факторов (свет, радиация и т.п.) уравнения непрерывности (1.37) принимают вид

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau_p} + D_p \frac{\partial^2 p}{\partial x^2} - \mu_p E \frac{\partial p}{\partial x} - p \mu_p \frac{\partial E}{\partial x}; \quad (1.38a)$$

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau_n} + D_n \frac{\partial^2 n}{\partial x^2} - \mu_n E \frac{\partial n}{\partial x} - n \mu_n \frac{\partial E}{\partial x}. \quad (1.38b)$$

В том случае, когда поле отсутствует или когда его ролью можно пренебречь, что имеет место в однородном полупроводнике, полагаем $E=0$. В этом случае выражения (1.38) существенно упрощаются и принимают вид:

$$\frac{\partial p}{\partial t} = -\frac{p - p_0}{\tau_p} + D_p \frac{\partial^2 p}{\partial x^2}; \quad (1.39a)$$

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau_n} + D_n \frac{\partial^2 n}{\partial x^2}. \quad (1.39b)$$

Уравнения 1.39а и 1.39б называются уравнениями диффузии.

Уравнения (1.39) позволяют достаточно строго анализировать многие типы полупроводниковых приборов.

1.11 Заряды в полупроводниках

Диэлектрическая релаксация. Пусть в ограниченном объеме полупроводника удалось сосредоточить избыточные концентрации электронов или дырок, так что образовался объемный за-

ряд с плотностью λ . Под действием возникшего поля заряд будет рассасываться, т.е. носители будут покидать тот начальный объем, в котором они были сосредоточены. Такое рассасывание заряда под действием собственного поля носит название диэлектрической релаксации, или релаксации Максвелла. При анализе диэлектрической релаксации пренебрегают рекомбинацией носителей и их диффузией, чтобы выделить явление релаксации в чистом виде. Следовательно, в правых частях (1.38) можно опустить все члены, кроме слагаемых, которые зависят от производной напряженности поля:

$$\frac{dp}{dt} = -p\mu_p \frac{dE}{dx}, \quad \frac{dn}{dt} = n\mu_n \frac{dE}{dx}.$$

Вычитая второе уравнение из первого для приращения концентраций, получаем

$$\frac{d(\Delta p - \Delta n)}{dt} = -\frac{\sigma}{\xi\xi_0}(\Delta p - \Delta n). \quad (1.40)$$

Решением является экспоненциальная функция

$$\Delta p - \Delta n = [\Delta p(0) - \Delta n(0)]e^{-t/\tau_\xi}, \quad (1.41)$$

где $\Delta p - \Delta n$ — избыточная начальная концентрация

$$\tau_\xi = \xi_0\xi/\sigma, \quad (1.42)$$

где τ_ξ — время диэлектрической релаксации.

Величина τ_ξ характеризует время, в течение которого нарушена нейтральность полупроводника: через $(3-4)\tau_\xi$ избыточный объемный заряд компенсируется и нейтральность восстанавливается. Величина $\tau_\xi = (10^{-11} - 10^{-12})$ с. Такая малая величина типична для процессов диэлектрической релаксации и является одной из основных условий квазинейтральности полупроводников.

Изменение удельного сопротивления полупроводника на том участке, где по тем или иным причинам скапливаются избыточные носители, носит название эффекта модуляции проводимости. Этот эффект играет значительную роль в полупроводниковых приборах, особенно при больших сигналах.

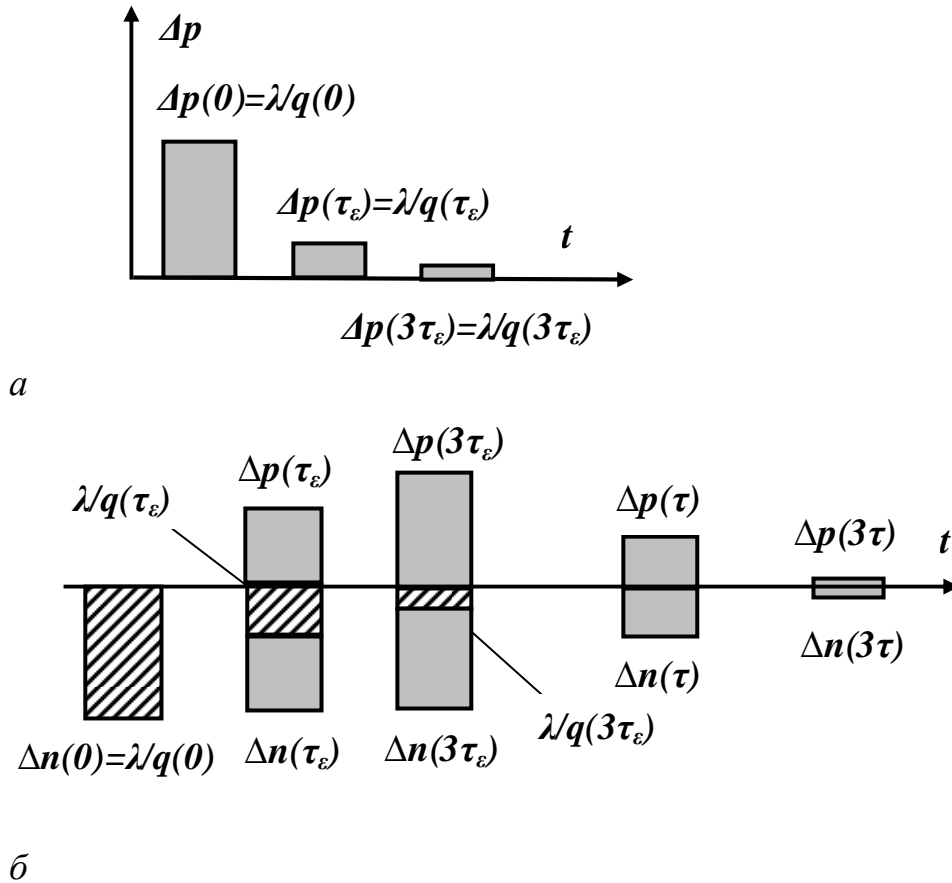


Рис 1.17 — Схематическое изображение процессов диэлектрической релаксации в примесном (дырочном) полупроводнике:

- а* — исходное возмущение вызвано основными носителями,
б — исходное возмущение вызвано неосновными носителями

1.12 Движение зарядов в полупроводниках

Биполярная диффузия. Пусть на поверхность полупроводника падает рассеянный пучок света (рис. 1.18). Тогда в тонком приповерхностном слое, в который проникает свет, будут генерироваться электронно-дырочные пары со скоростью Δg , пар/(см³/сек). Между поверхностью и основным объемом возникнут градиенты концентрации электронов и дырок, и избыточные носители начнут диффундировать в глубь полупроводника.

Такое совместное движение обоих типов носителей называют биполярной диффузией. Если бы подвижности (а значит, и коэффициенты диффузии) у электронов и дырок были одинаковы, то они двигались бы в виде единого потока с нейтральным заря-

дом, который можно было бы анализировать с помощью одного из уравнений (1.40). На самом деле подвижности дырок и электронов различны, поэтому у электронного потока будет тенденция «обогнать» дырочный поток.

В результате взаимного сдвига потоков образуются объемный заряд и соответствующее электрическое поле, которое тормозит поток электронов и ускоряет поток дырок. В результате этого устанавливается стационарный режим, при котором избыточные электроны и дырки распределены в виде сдвинутых друг относительно друга «облачков». Эти «облачка» двигаются синхронно, так что результирующий ток отсутствует. По мере движения происходит рекомбинация электронов и дырок, что приводит к уменьшению их избыточных концентраций. Описанные явления известны под названием эффекта Дембера, а электрическое поле и разность потенциалов, свойственные этому эффекту, называют демберовским полем и демберовским напряжением. Величины демберовского поля и демберовского напряжения можно оценить из условия нулевого результирующего тока. Полагая $j = 0$ в формуле (1.37), подставляя значения токов из (1.38) (1.33) и используя соотношение (1.32), получаем:

$$E_{ДЕМ} = \frac{q}{\sigma} \left(D_p \frac{dp}{dx} - D_n \frac{dn}{dx} \right), \quad (1.43a)$$

если учесть, что $\frac{dp}{dx} = \frac{dn}{dx}$, получим

$$E_{ДЕМ} = \frac{q}{\sigma} (D_p - D_n) \frac{Dp}{dx}. \quad (1.43b)$$

Учитывая соотношение $-\frac{du}{dx} = E_{Демб}$, проинтегрируем (1.43a) в пределах $x = 0$ до $x = \infty$ и положим $p(\infty) = p_0$ и $n(\infty) = n_0$. Тогда

$$U_{ДЕМ} = \frac{q}{\sigma} [D_p \Delta p(0) - D_n \Delta n(0)] \quad (1.44a)$$

или, если принять условие квазинейтральности $\Delta n = \Delta p$,

$$U_{ДЕМ} = \frac{q}{\sigma} \Delta p (D_p - D_n). \quad (1.44b)$$

В заключение отметим, что биполярная диффузия, несмотря на отсутствие результирующего тока, является неравновесным

процессом, так как имеются потоки носителей. Поэтому уровень Ферми не будет в данном случае постоянным, во всяком случае, в той области, где потоки еще существенны и где повышены концентрации обоих типов носителей. В этих областях уровень Ферми «расщепляется» на два квазиуровня — один для электронов, другой для дырок.

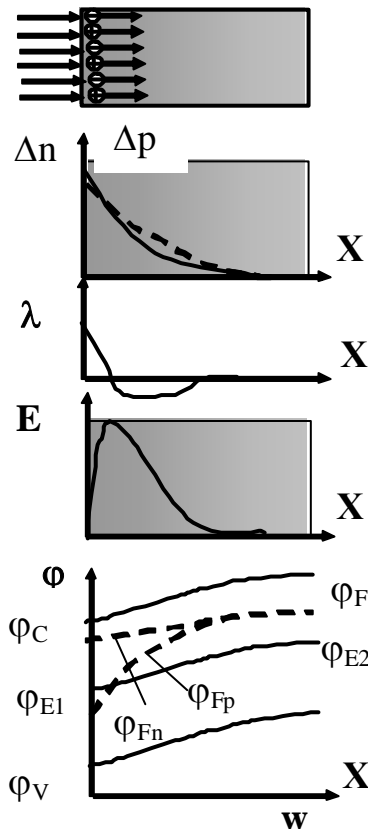


Рис 1.18 — Биполярная диффузия под действием света, генерирующего электронно-дырочные пары

Монополярная диффузия

При анализе биполярной диффузии предполагалось, что оба типа носителей образуются одновременно и в равных количествах. Поэтому «автоматически» обеспечивалась квазинейтральность, а результирующий ток отсутствовал. Предположим теперь, что в приповерхностный слой полупроводника вводится (инжектируется) только один тип носителей — неосновных (рис. 1.19). Пусть для определенности осуществляется инжекция дырок в электронный полупроводник. Инжектированные дырки благодаря градиенту концентрации будут диффундировать в глубь кри-

сталла, т.е. появится дырочный ток. Заряд дырок практически мгновенно (со временем диэлектрической релаксации) будет компенсирован таким же зарядом электронов, притягиваемых из глубоких слоев (необходимое количество дополнительных электронов поступает из внешней цепи, по которой протекает ток). В результате вблизи инжектирующей поверхности образуется квазинейтральное электронно-дырочное «облачко», почти такое же, как при биполярной диффузии (см. рис. 1.18). Образуется также демберовское поле. Процесс, при котором диффундируют носители одного типа, а носители другого типа лишь обеспечивают квазинейтральность, называют монополярной диффузией.

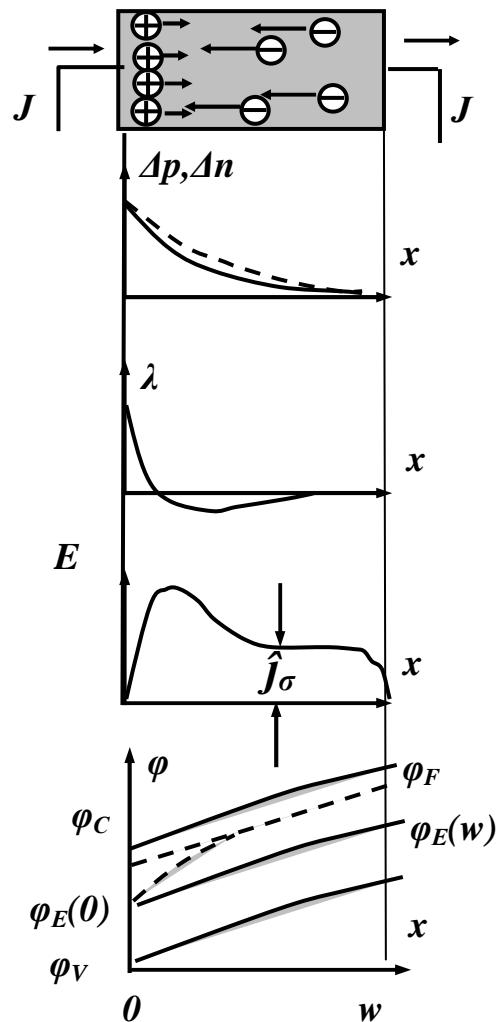


Рис. 1.19 — Монополярная диффузия в результате инжекции дырок в электронный полупроводник

Несмотря на внешнее сходство в распределении носителей, монополярная диффузия принципиально отличается от биполярной. Отличия состоят в следующем.

1. Имеется результирующий ток; поэтому полупроводник должен быть элементом замкнутой цепи. Соответственно, помимо демберовского поля, сосредоточенного вблизи инжектирующей поверхности, во всей толще полупроводника действует омическое поле, обусловленное приложенным напряжением.

2. Потоки дырок и электронов направлены в разные стороны: дырки двигаются в глубь кристалла, а электроны — в сторону инжектирующей поверхности, в район электронно-дырочного «облачка», где происходит интенсивная рекомбинация и необходимо пополнение основных носителей. Результирующий ток является суммой дырочной и электронной составляющих.

3. В связи с постоянством результирующего тока дырочная и электронная составляющие меняются в разные стороны: убывание дырочного тока от поверхности в глубь кристалла сопровождается соответствующим ростом электронной составляющей. На самой инжектирующей поверхности электронный ток равен нулю, так как в элементарном приповерхностном слое убыль электронов за счет рекомбинации ничтожна и нет необходимости в их пополнении потоком электронов из глубинных слоев. Следовательно, в непосредственной близости от инжектирующей поверхности ток обусловлен только дырками. В глубине кристалла, где дырочный ток благодаря рекомбинации делается равным нулю, обращается в нуль и диффузионная составляющая электронного тока. Следовательно, вдали от инжектирующей поверхности ток обусловлен только электронами и имеет чисто дрейфовый характер: электроны двигаются в омическом поле, созданном внешним напряжением. Указанные отличия принципиальны с физической точки зрения. С математической же точки зрения распределения $\Delta p(x)$ и $\Delta n = \Delta p$ описываются либо уравнением биполярной диффузии (если полупроводник близок к собственному), либо (в случае электронного полупроводника) уравнением (1.38a). Что касается выражения для поля E , то оно в данном случае будет отличаться от (1.43), поскольку результирующий ток не равен нулю. Подставляя (1.33) и (1.34) в (1.32), легко получить:

$$E = 1/\sigma [j + q(D_p - D_n) \frac{dp}{dx}]. \quad (1.45)$$

Сравнивая с (1.43), замечаем, что поле при монополярной диффузии складывается из постоянной и демберовской составляющих. Строгое решение задачи о распределении носителей при монополярной диффузии встречает определенные трудности. Поэтому решим эту задачу для малых возмущений и в диффузионном приближении.

Величину возмущения принято характеризовать уровнем инжекции

$$\delta = \frac{\Delta p}{n_0} \approx \frac{\Delta n}{n_0},$$

т.е. отношением концентрации избыточных носителей к равновесной концентрации основных носителей. Малому возмущению соответствует неравенство $\delta \ll 1$, т.е.

$$\Delta p \approx \Delta n \ll n_0. \quad (1.45, a)$$

Это неравенство называют условием низкого уровня инжекции. Диффузионное приближение характеризуется тем, что в уравнении непрерывности пренебрегают составляющими, связанными с напряженностью поля. Стационарный и нестационарный режимы можно записать в виде:

$$\frac{\partial^2 \Delta p}{\Delta x^2} - \frac{\Delta p}{L^2} = \frac{\tau_p}{L^2} \frac{\partial \Delta p}{\partial t}; \quad (1.46)$$

$$\frac{\partial^2 \Delta p}{\Delta x^2} - \frac{\Delta p}{L^2} = 0. \quad (1.47)$$

Величина L определяется выражением

$$L = \sqrt{D\tau}.$$

L имеет размерность длины и носит название средней диффузионной длины или средней длины диффузии. Остановимся сначала на стационарном решении уравнения (1.46). Как известно, его решение является суммой двух экспонент:

$$\Delta p(x) = A_1 e^{-x/L_p} + A_2 e^{x/L_p}, \quad (1.48)$$

где коэффициенты A_1 и A_2 определяются из граничных условий $x=0$, $x=w$ и принципа физической реализуемости. На бесконечно большем расстоянии от границы инжекции величина избы-

точной концентрации должна стремиться к нулю в результате комбинации инжектированных носителей. Первый член уравнения (1.48) при $x = \infty$ и конечном значении коэффициента A_1 всегда равен нулю. Второй член при $x = \infty$ будет равен нулю, если $A_2 = 0$. Учтя эти условия, можно записать $A_2 = 0$ и $A_1 = \Delta p(0)$.

Тогда:

$$\Delta p(x) = \Delta p e^{-x/L_p}, \quad (1.49)$$

т.е. избыточные концентрации спадают по экспоненте. Это один из характерных случаев в теории полупроводниковых приборов. **Из выражения (1.49) следует, что диффузионная длина есть то расстояние, на котором концентрация диффундирующих носителей (при их экспоненциальном распределении) уменьшается в e раз. На участке длиной $(3 - 4)L$ концентрация уменьшается в **20—50 раз**, т.е. становится пренебрежимо малой по сравнению с исходным значением. В заключение следует отметить, что материал первой главы является фундаментом, на котором базируются последующие разделы.**

Таблица 1.1 — Основные параметры некоторых полупроводников

Параметр	Полупроводник			
	Кремний	Германий	Сплав <i>GaAs</i>	Сплав <i>InSb</i>
Заряд ядра	14	32	—	—
Атомный вес	28,1	72,6	—	—
Диэлектрическая проницаемость (отн.ед.)	12	16	11	16
Температура плавления, С	1420	940	1280	520
Коэффициент теплопроводности λ , Вт/(см*С)	1,2	0,55	—	—
Удельная теплоёмкость c , Дж/(г*С)	0,75	0,41	—	—
Эффективная масса электронов m_n (отн. ед.)	0,33	0,22	0,07	0,013

Окончание табл. 1.1

Параметр	Полупроводник			
	Кремний	Германий	Сплав <i>GaAs</i>	Сплав <i>InSb</i>
Эффективная масса дырок m_p (отн. ед.)	0,55	0,39	50,	0,6
Ширина запрещенной зоны φ_3 , эВ	1,11	0,67	1,4	0,18
Эффективная плотность состояний N_C , см ⁻³	$2,8 \cdot 10^{19}$	10^{19}	–	–
То же N_V , см ⁻³	10^{19}	$0,61 \cdot 10^{19}$	–	–
Подвижность электронов μ_n , см ² / (В*с)	500	1800	11000	65000
Подвижность дырок μ_p , см ² / (В*с)	500	1800	450	700
Собственное удельное сопротивление ρ_i , Ом*см	$2 \cdot 10^5$	60	$4 \cdot 10^8$	–
Собственная концентрация n_i , см ⁻³	$2 \cdot 10^{10}$	$2,5 \cdot 10^{13}$	$1,5 \cdot 10^8$	–
Коэффициент диффузии электронов D_n , см ² /с	36	100	290	1750
Коэффициент диффузии дырок D_p , см ² /с	13	45	12	17
Критическая напряженность поля E_{Kpn} , В/см	2500	900	–	–
То же E_{Kpp} , В/см	7500	1400	–	–
Максимальная скорость, см/с	10^7	$6 \cdot 10^6$	–	–

Таблица 1.2 — Основные физические константы, используемые в теории полупроводников

Основные физические константы	Приближенные значения
Элементарный заряд	$q = 1,6 * 10^{-19}$ Кл
Масса свободного электрона	$m = 9 * 10^{-31}$ кг
Постоянная Планка	$h = 6,6 * 10^{-34}$ Дж/с
Постоянная Больцмана	$k = 1,37 * 10^{-23}$ Дж/град.
Электрическая постоянная (диэлектрическая проницаемость вакуума)	$\epsilon_0 = 19 * 10^{-14}$ Ф/см
Магнитная постоянная (магнитная проницаемость вакуума)	$\mu_0 = 1,25 * 10^{-8}$ Г/см
Число Авогадро (количество атомов в 1 грамм-атоме вещества)	$N_0 = 6 * 10^{23}$

Вопросы для самопроверки

1. Какие типы полупроводников Вы знаете?
2. Для каких целей вводятся примеси?
3. Почему примеси не образуют зон?
4. Объясните физический смысл потенциала Ферми и уровни его залегания в различных типах полупроводниковых материалов.
5. Объясните физический смысл температурного потенциала ϕ_T .
6. Что такое эффективная плотность состояний в зонах проводимости и валентной?
7. От каких параметров зависят концентрации p и n в собственном полупроводнике?
8. Что такое электростатический потенциал полупроводника?
9. Дайте определение понятию «химический потенциал» и что означает условие $grad\phi_F = const$?
10. Какие полупроводники относятся к невырожденным?
11. Запишите потенциалы Ферми для дырочного и электронного полупроводников, если полупроводники невырожденные.

12. Где залегают уровни Ферми в дырочном, электронном и собственном полупроводниках?

13. Запишите уравнение нейтральности. В чем его физический смысл?

14. В типичном электронном полупроводнике уровень Ферми можно записать в виде $\varphi_F = \varphi_T + \varphi_T \ln \frac{N_D}{n_i}$. Если $N_D \rightarrow 0$, потенциал Ферми $\varphi_F \rightarrow -\infty$, но это противоречит физическому смыслу. В чем ошибка?

15. Что характеризует подвижность носителей заряда μ и почему подвижность с повышением температуры падает?

16. Запишите формулы удельной проводимости для собственного, электронного и дырочного полупроводников. Нарисуйте графики зависимости удельной проводимости от температуры.

17. Что означает понятие «время жизни носителей» и почему с увеличением концентраций доноров или акцепторов время жизни падает?

18. Почему время жизни с повышением температуры растёт.

19. От каких параметров зависят диффузионные и дрейфовые токи в полупроводнике?

20. Объясните, что такое диэлектрическая релаксация в полупроводниках и связана ли она с уравнением нейтральности?

21. Принципиальное отличие неоднородных полупроводников от однородных?

22. Объясните основные процессы при биполярной диффузии.

23. Объясните физическую сущность демберовского поля и напряжённости.

24. В чем сходство и принципиальное отличие между монополярной и биполярной диффузиями?

25. Выведите формулу для избыточной концентрации носителей при монополярной диффузии.

26. Что означает диффузионная длина L ?

2. ДИОДЫ

2.1 Введение

Полупроводниковый диод представляет собой комбинацию двух полупроводниковых слоев с различными типами проводимости (рис. 2.1). Такая комбинация обладает способностью гораздо лучше пропускать ток в одном направлении (от слоя p к слою n) и гораздо хуже в другом. Полярность напряжения, соответствующая большим токам, называется прямой, а соответствующая меньшим токам — обратной. Обычно пользуются терминами «прямое и обратное напряжение» (смещение), «прямой и обратный токи».

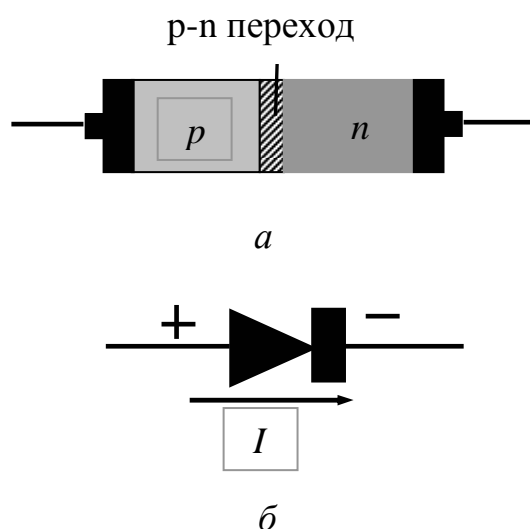


Рис. 2.1 — Плоскостной диод:
a — упрощенная структура; *б* — условное
 графическое обозначение

На рис. 2.1, *б* показаны символическое изображение диода, направление прямого тока и полярность прямого напряжения. Вентильное свойство диода является следствием ярко выраженной внутренней неоднородности структуры. Ступенчатая неоднородность даже в полупроводниках с одним типом проводимости сопровождается нарушением закона Ома в связи с образованием объемных зарядов и потенциального барьера. В данном случае, когда слои разнотипные, нелинейность, естественно, оказывается еще сильнее. Несмотря на кажущуюся простоту, диод яв-

ляется сложным прибором, так как имеет много областей, обладающих различными физическими свойствами.

1. Области, прилегающие к границам металл-полупроводник. Металл необходим для создания выводов.

2. Области, прилегающие к границе (металлургическая граница), образованной двумя полупроводниковыми материалами p и n .

3. Области полупроводников вдали от металлургической границы.

Поверхность, по которой контактируют слои p и n , называется металлургической границей, а прилегающие к ней области объемных (или пространственных) зарядов — электронно-дырочным переходом или p - n переходом. Два других внешних контакта в диоде должны быть не выпрямляющие: поэтому их называют омическими контактами.

Принципы получения омических контактов рассмотрим позже.

Плоскостными диодами и соответственно плоскостными переходами называют такие диоды и переходы, у которых граница между слоями плоская, а площади обоих слоев одинаковы. Эти условия не всегда соблюдаются на практике, но соблюдение этих условий облегчает анализ и в то же время позволяет получить правильное представление о процессах в реальном диоде и его характеристиках. Диоды — самостоятельный, весьма обширный класс полупроводниковых приборов. В то же время диод как простейший прибор с одним p - n переходом является основой многопереходных приборов — транзисторов и других. Мы изучим диоды именно с этой точки зрения.

В общих чертах процессы в полупроводниковом диоде можно охарактеризовать следующим образом. В отсутствие внешнего напряжения имеет место больцмановское равновесие: вблизи металлургической границы, где сконцентрированы объемные заряды и связанное с ними поле, диффузионные и дрейфовые потоки носителей уравновешены и результирующего тока нет. Полярность прямого напряжения способствует «выталкиванию» дырок из p -слоя в n -слой и электронов в обратном направлении. В результате повышается концентрация дырок в n -слое и концентрация электронов в p -слое, т.е. имеет место инжекция неосновных носителей в обоих слоях диода. Инжектированные носители диффундируют в глубь слоев, и эта монополярная диффу-

зия сопровождается протеканием прямого тока. Полярность обратного напряжения способствует «выталкиванию» дырок из n -слоя и электронов из p -слоя в область перехода, т.е. имеет место экстракция неосновных носителей и протекает соответствующий обратный ток.

Диффузионный ток определяется граничной величиной избыточной концентрации неосновных носителей. Поскольку при экстракции модуль избыточной концентрации не может превышать весьма малого значения равновесной концентрации, а при инжекции такого ограничения нет, прямой ток оказывается на много больше обратного, что и является основой вентильных свойств диода.

2.2 Электронно-дырочный переход

Плоскостной диод состоит из электронно-дырочного перехода, двух нейтральных (или квазинейтральных) слоев и омических контактов. Поскольку процессы в нейтральных полупроводниках были детально изучены в первой главе, следует, прежде всего, рассмотреть процессы в p - n переходе и учесть возможное влияние нейтральных областей на характеристики и параметры диодов.

Классификация p - n переходов

Прежде всего, заметим, что p - n переход нельзя осуществить путем простого соприкосновения двух разнородных полупроводниковых пластинок, так как при этом неизбежен промежуточный (хотя бы и очень тонкий) слой воздуха или поверхностных пленок. Настоящий переход получается в единой пластинке полупроводника, в которой тем или иным способом получена достаточно резкая граница между слоями p и n . Резкость границы играет принципиальную роль для образования перехода, так как нерезкая граница приводит к образованию плавного перехода, а как показывает теория, такой переход не обладает теми вентильными свойствами, которые требуются для работы полупроводниковых диодов и транзисторов.

Понятие резкости формулируется следующим образом: граница между слоями является резкой, если градиент концентрации примеси (считающийся постоянным в пределах перехода) удовлетворяет неравенству

$$\left| \frac{dN}{dx} \right| l_{Di} \gg n_i, \quad (2.1)$$

где N — эффективная концентрация примеси, l_{di} — дебаевская длина в собственном полупроводнике.

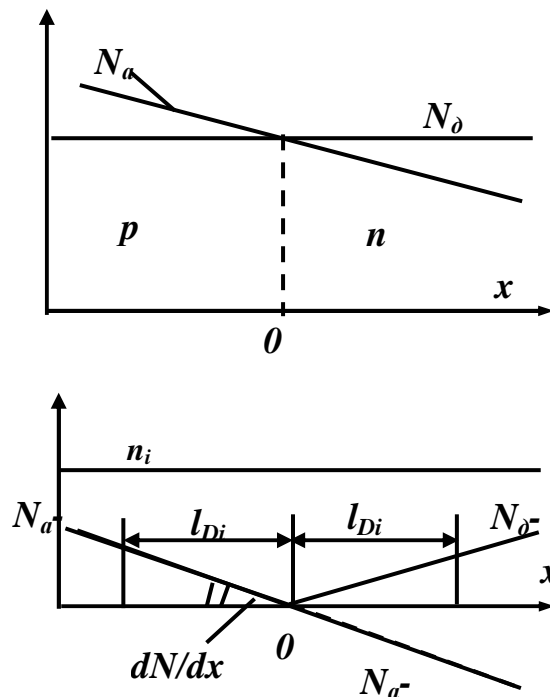


Рис. 2.2 — Распределение полных и эффективных концентраций примеси вблизи металлургической границы плавного перехода

Переходы, в которых имеется скачкообразное изменение концентрации на границе слоев ($dN/dx = \infty$), будем называть ступенчатыми. Они представляют собой предельный случай класса резких переходов, в которых градиент концентрации примесей конечен, но удовлетворяет неравенству (2.1). На практике ступенчатые переходы являются, конечно, известным приближением. Однако они хорошо отражают свойства многих реальных p - n

структур и, кроме того, оказываются проще для анализа. Поэтому ниже им будет уделено главное внимание.

Контакты, в которых условие (2.1) не соблюдается, не называют переходами, а относят к неоднородным полупроводникам. По соотношению концентраций основных носителей в слоях p и n переходы делятся на симметричные и несимметричные. В симметричных переходах имеет место соотношение $p_p \approx n_n$, т.е. концентрации основных носителей в обоих слоях почти одинаковы.

Такие переходы используются сравнительно редко и не являются типичными. Гораздо большее распространение имеют несимметричные переходы, в которых выполняется неравенство

$$p_p \gg n_n \text{ или } n_n \gg p_p \quad (2.2)$$

и концентрации различаются в несколько раз. Именно такие переходы будут анализироваться в дальнейшем, причем для определенности считаем, что слой p более низкоомный, чем слой n , т.е. $p > n$.

Полученные выводы легко использовать при обратном соотношении концентраций. В случае резкой асимметрии, когда концентрации основных носителей различаются более чем на порядок, переходы называют односторонними и обычно обозначают символами $(n-p^+$ или $n^+-p)$. Иногда, чтобы отличить несимметричные переходы от односторонних, используют для первых обозначения p^+-n (или n^+-p), а для вторых $p^{++}-n$ (или $n^{++}-p$).

Структура p - n перехода. Концентрации примесей и свободных носителей в каждом из слоев диода показаны на рис. 2.3, а, причем для наглядности разница в концентрациях p_p и n_n принята гораздо меньшей, чем это имеется в действительности.

Поскольку концентрация дырок в слое p значительно больше, чем в слое n , после соединения полупроводников часть дырок диффундирует из слоя p в слой n . При этом в слое n вблизи металлургической границы окажутся избыточные дырки. Вследствие диэлектрической релаксации избыточный заряд дырок будет компенсирован электронами полупроводника n -типа. Наличие разности концентраций дырок в полупроводнике n приведет к образованию разности химических потенциалов и к возникновению диффузии.

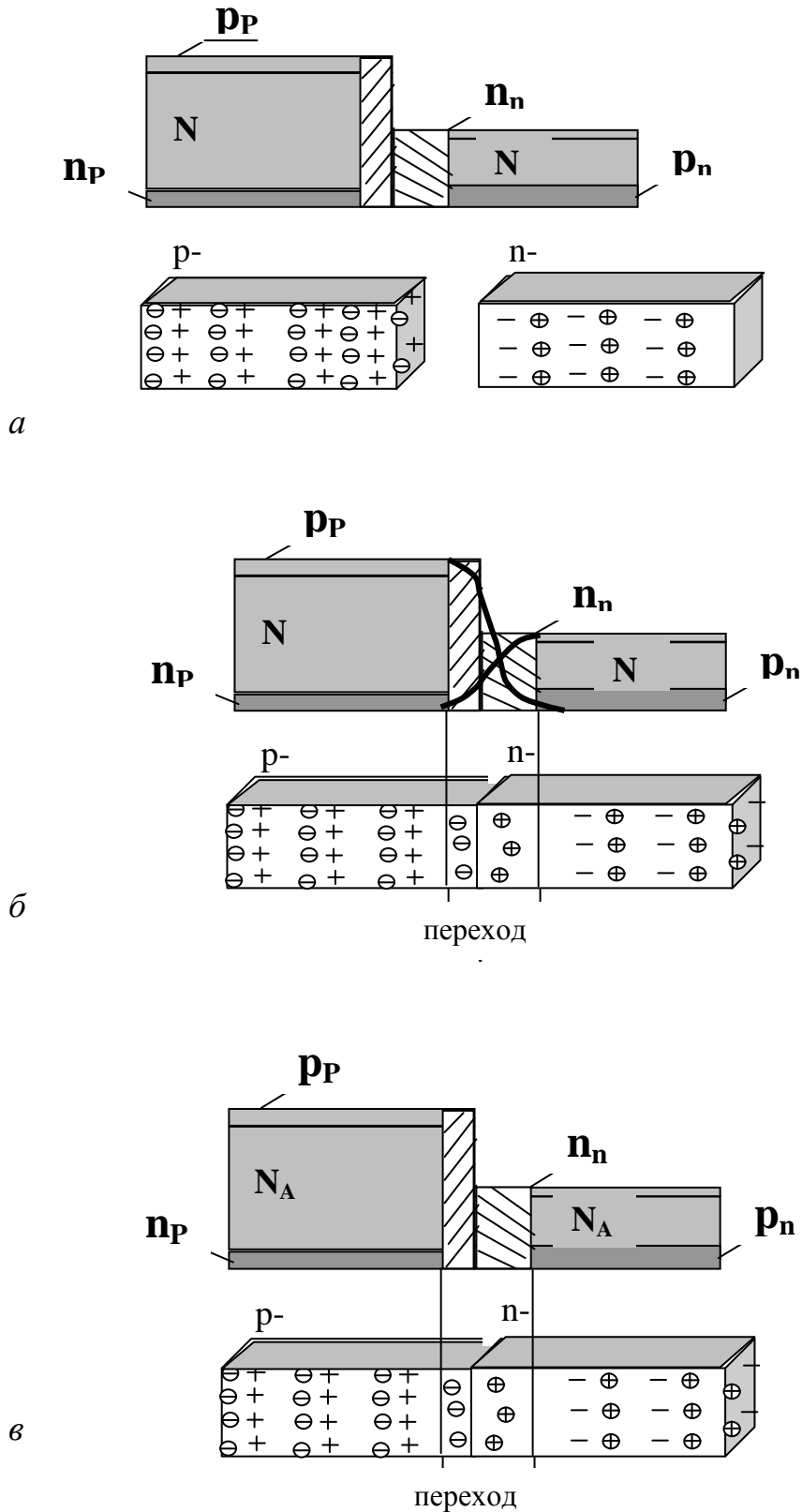


Рис. 2.3 — Структура p - n перехода: *a* — начальное состояние слоев; *б* — пространственные заряды в реальном переходе; *в* — пространственные заряды в идеализированном переходе; + — дырка, \ominus — ионизированный акцептор, — — электрон, \oplus — ионизированный донор

Дырки из области с повышенной концентрацией начнут двигаться в глубь полупроводника n , где концентрация ниже. По мере движения дырки будут рекомбинировать с электронами до тех пор, пока их концентрация не уменьшится до равновесной. Соответственно в области, прилегающей к металлургической границе, уменьшится концентрация свободных электронов и «обнажатся» некомпенсированные положительные ионы донорных атомов. Слева от границы «обнажатся» некомпенсированные отрицательные ионы акцепторных атомов, поскольку часть дырок перешла отсюда в слой n (рис. 2.3, б). Аналогичные рассуждения действительны для электронов слоя n , которые частично диффундируют в слой p . Однако в несимметричном переходе, в котором $n_n < p_p$, диффузия электронов в слой p малосущественна, поскольку разность концентраций $n_n - n_p$ существенно меньше разницы $p_p - p_n$, а именно этими разностями определяются градиенты концентраций и диффузионные токи. Область образовавшихся пространственных зарядов и есть область p - n перехода. Часто эту область называют обедненным или истощенным слоем, имея в виду резко пониженную концентрацию подвижных носителей в обеих ее частях.

Однако, строго говоря, переход и обедненный слой — не одно и то же: область перехода несколько шире, потому что объемные заряды и связанное с ними поле зарождаются уже при очень небольшом (несколько процентов) уменьшении концентрации носителей по сравнению с равновесной, тогда как понятию обедненного слоя соответствует спад концентрации носителей, по крайней мере на порядок. Промежуточные участки между «границами» обедненного слоя и перехода являются участками экранирования основных p - и n -слоев диода от поля, создаваемого зарядами обедненного слоя. При перепаде концентраций носителей на три порядка и более протяженность этих участков обычно не превышает 0,1 мкм, тогда как ширина собственно обедненного слоя, как увидим ниже, в несколько раз больше.

Поэтому есть основания идеализировать переход так, как показано на рис. 2.3, б, т.е. пренебречь наличием свободных носителей в переходе и считать его границы совпадающими с гра-

ницами обедненного слоя. Такая идеализация существенно упрощает решение многих задач, за исключением, конечно, тех, которые непосредственно связаны с анализом потоков носителей. Переход в целом, разумеется, нейтрален, т.е. отрицательный заряд в левой части и положительный заряд в правой части одинаковы. При этом условии различие в концентрациях акцепторной и донорной примесей неизбежно связано с различием в протяженности обоих зарядов: в слое с меньшей концентрацией примеси (в нашем случае в n -слое) область объемного заряда должна быть шире. Иначе говоря, несимметричный переход сосредоточен в высокоомном слое.

Особое внимание следует обратить на тот факт, что внутри p - n перехода имеется участок с собственной (т.е. сильно пониженной по отношению к основным слоям) проводимостью, так как в области перехода отсутствуют подвижные заряды, обусловленные примесями.

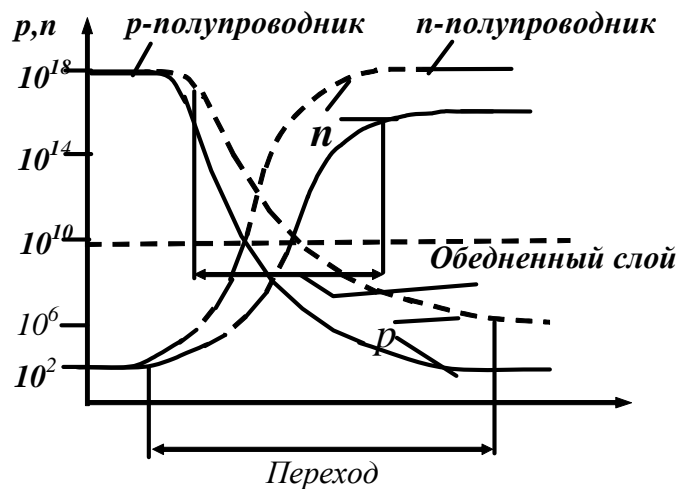


Рис. 2.4 — Распределение носителей в несимметричном переходе (полулогарифмический и линейный масштаб).

Пунктирными линиями показано распределение в симметричном переходе

Область перехода является наиболее высокоомной частью диодной структуры. Участок с собственной проводимостью в общем случае не совпадает с металлургической границей, а сдвинут в сторону того слоя, где сосредоточен переход. Пространст-

венные заряды в переходе образуют электрическое поле, которое направлено так, что оно ограничивает диффузию носителей.

В равновесном состоянии диффузионные потоки носителей, обусловленные градиентами концентрации, в любой точке равны дрейфовым потокам тех же носителей, обусловленным градиентом потенциала и направленным навстречу диффузионным потокам.

Строгий количественный и качественный анализ образования перехода можно провести только с позиций зонной теории.

Рассмотрим p - n переход с точки зрения зонной теории. В отсутствие контакта p - и n -слои характеризуются диаграммой на рис. 2.5, *a*, из которой видно, что потенциалы Ферми имеют разные значения.

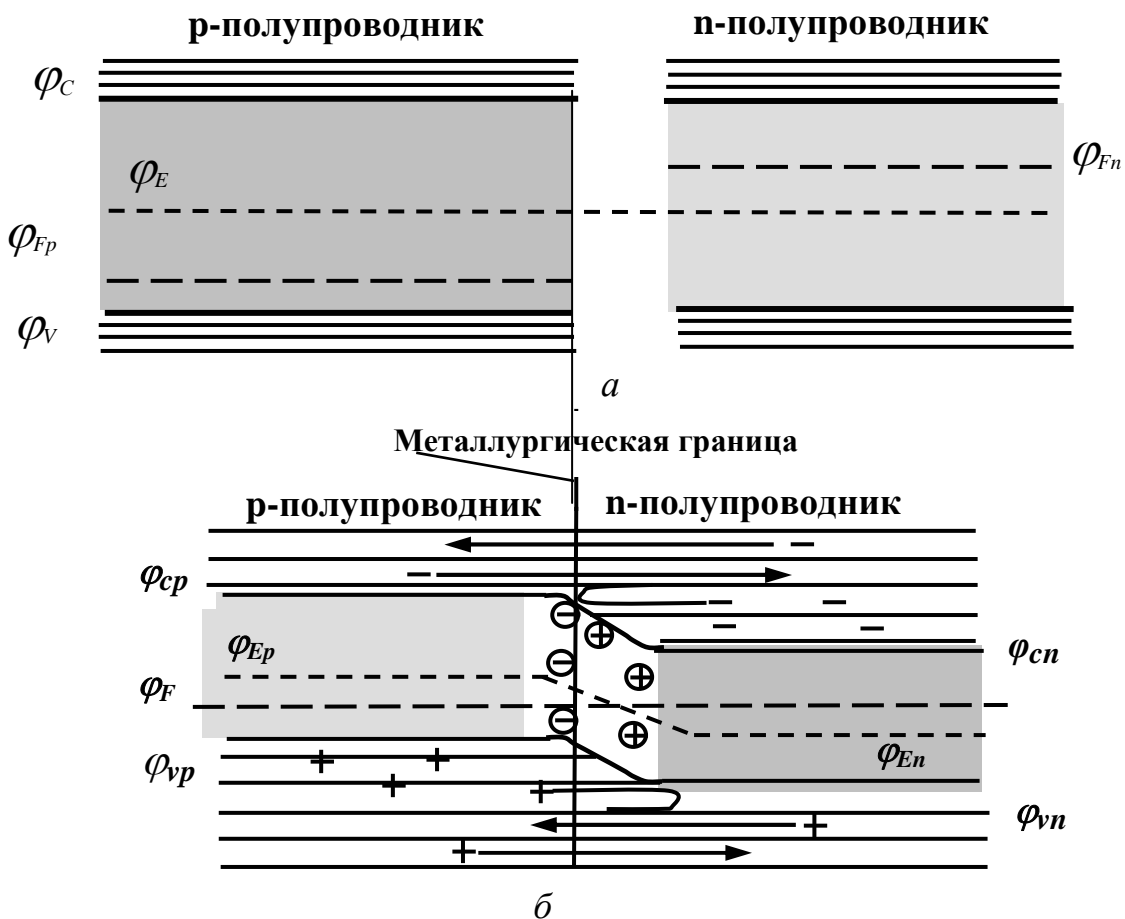


Рис. 2.5 — Зонная диаграмма слоев (а) и p - n перехода в равновесном состоянии (б)

При наличии контакта уровень Ферми в системе, находящейся в термодинамическом равновесии, должен быть единым, данное условие было обосновано в первой главе, а это должно приводить к неизбежному искривлению зон, расщеплению электростатических потенциалов и образованию потенциального барьера. Напомним, что такой барьер образуется также при ступенчатой неоднородности в полупроводнике с неизменным типом проводимости (рис. 1.29, б). Если воспользоваться образной интерпретацией движения носителей в зонах, то равновесное состояние перехода можно охарактеризовать следующим образом. Основная масса дырок p -слоя диффундирует слева направо в область перехода, но не может преодолеть потенциальный барьер и, проникнув в переход на некоторую глубину, «отражается» и возвращается в p -слой (рис. 2.6).

Дырки n -слоя независимо от энергии беспрепятственно «всплывают» в p -слой и образуют поток справа налево. Этот поток уравнивается встречным потоком дырок p -слоя с энергией, достаточной, чтобы преодолеть потенциальный барьер. Аналогичная ситуация имеет место по отношению и к электронам.

Глубина проникания носителей в переход тем больше, чем выше их энергия. В области перехода на рис. 2.5, б показаны ионизированные атомы акцепторов (слева) и доноров (справа). Как известно, уровни этих ионов расположены вдоль всего соответствующего слоя, но мы показываем их только в пределах перехода, чтобы подчеркнуть нескомпенсированность зарядов ионов на этих участках и тем самым подчеркнуть, что в переходе нет подвижных зарядов. В самом деле, расстояние между дном зоны проводимости и уровнем Ферми увеличивается от точки **а** влево, а значит, на участке — **б** быстро убывает вероятность заполнения этой зоны электронами. Поэтому если справа от точки **а** электроны компенсируют положительный заряд донорных ионов и слой n нейтрален, то слева от точки **а** концентрация электронов резко падает и такой компенсации нет. Аналогичные условия создаются на участке справа от точки **в** по отношению к акцепторным ионам. Очевидно, что ионы, показанные на рис. 2.5, б, создают пространственный заряд.

Анализ перехода в равновесном состоянии

Строгий анализ p - n переходов, в общем виде, весьма сложен. Наиболее просто анализируется ступенчатый переход, рассмотренный в предыдущем разделе. Но даже в этом случае необходимы упрощающие допущения. В самом деле, если решать задачу строго, т.е. исходить из структуры перехода, показанной на рис. 2.3, б, то нужно учитывать распределение подвижных носителей заряда и в переходе.

Задача анализа упрощается, если считать, что переход имеет структуру, показанную на рис. 2.3, в, где концентрации подвижных носителей внутри перехода равны нулю и размеры перехода малы, т.е. можно пренебречь процессами генерации и рекомбинации в переходе. Такое допущение вполне приемлемо. Во-первых, в этой области доноры и акцепторы все ионизированы, а во вторых, вследствие возникшей напряженности электрического поля на переходе свободные заряды уносятся в нейтральные слои соответствующих полупроводников.

Высоту равновесного потенциального барьера можно получить непосредственно из рис. 2.5, а:

$$\Delta\varphi_0 = \varphi_{F_n} - \varphi_{F_p}, \quad (2.3)$$

где величины в правой части — потенциалы Ферми в полупроводниках до их объединения.

Из данного выражения следует, что чем ближе уровни Ферми находятся к разрешенным зонам, тем больше диффузионный потенциал. В первой главе мы показали, чем выше концентрация примесей, тем ближе потенциал Ферми к потенциалам разрешенных зон. Следовательно, изменяя концентрацию примесей, можно управлять величиной диффузионного потенциала.

Максимальное значение потенциала Ферми для невырожденных полупроводников равно границам разрешенных зон полупроводника. Максимальная высота потенциального барьера перехода в невырожденных полупроводниках равна ширине запрещенной зоны полупроводникового материала.

Воспользуемся формулой (1.18а) и запишем эти потенциалы через концентрации свободных электронов в слоях:

$$\varphi_{F_p} = \varphi_T \ln \frac{n_{p0}}{n_i} + \varphi_E;$$

$$\varphi_{F_n} = \varphi_T \ln \frac{n_{n0}}{n_i} + \varphi_E,$$

где индекс 0 соответствует равновесному состоянию концентраций. Электрические потенциалы в обоих полупроводниках равны. Подставляя значения потенциалов Ферми в (2.3), приходим к выражению

$$\Delta\varphi_0 = \varphi_T \ln \frac{n_{n0}}{n_{p0}}. \quad (2.4a)$$

Если воспользоваться формулой (1.18б) или в выражении (2.4a) заменить концентрации электронов концентрациями дырок с помощью соотношения (1.16), то высота потенциального барьера запишется следующим образом:

$$\Delta\varphi_0 = \varphi_T \ln \frac{p_{p0}}{p_{n0}}. \quad (2.4б)$$

Величину $\Delta\varphi_0$ иногда называют диффузионным потенциалом, поскольку эта разность потенциалов, во-первых, образуется в результате диффузии носителей через переход и, во-вторых, противодействует диффузионным потокам носителей. Еще одно название для величины $\Delta\varphi_0$ — контактная разность потенциалов.

Выражая в любой из формул (2.4) концентрацию неосновных носителей (n_p или p_n) через концентрацию основных носителей в том же слое с помощью соотношения (1.16) и используя выражения (2.4б), легко получить зависимость высоты равновесного барьера от удельных сопротивлений слоев:

$$\Delta\varphi_0 = \varphi_T \ln \frac{\rho_i^2 (b+1)}{\rho_p \rho_n b}, \quad (2.5)$$

где $b = \frac{\mu_n}{\mu_p}$.

Для вычисления равновесной ширины потенциального барьера l_0 воспользуемся распределениями примесей и зарядов, которые показаны на рис. 2.6.

Примем, что поле в переходе направлено вдоль оси x и отсутствует во всех других направлениях. При выполнении этого

допущения действительны уравнения Пуассона, в которых плотности заряда в обеих частях ступенчатого перехода постоянны и определяются концентрациями соответствующих ионизированных примесей (рис. 2.6, б):

$$\lambda_p = -qN_a; \quad \lambda_n = -qN_D.$$

Из уравнений Пуассона следует, что плотность заряда в переходе прямо пропорциональна концентрации примесей. Учитывая, что для ступенчатого перехода выполняется условие $N_a \gg N_D$, можно заключить $|\lambda_p| \gg \lambda_n$. Переход должен быть нейтральным, т.е.

$$Q^+ = Q^- = \lambda_p V_p = \lambda_n V_n,$$

где V_p, V_n — объемы пространственного заряда в переходе, расположенные в полупроводниках p и n типа. Учитывая условие равенства зарядов в переходе, можно записать $V_p \ll V_n$, т.е. переход сосредоточен в основном в высокоомном полупроводнике.

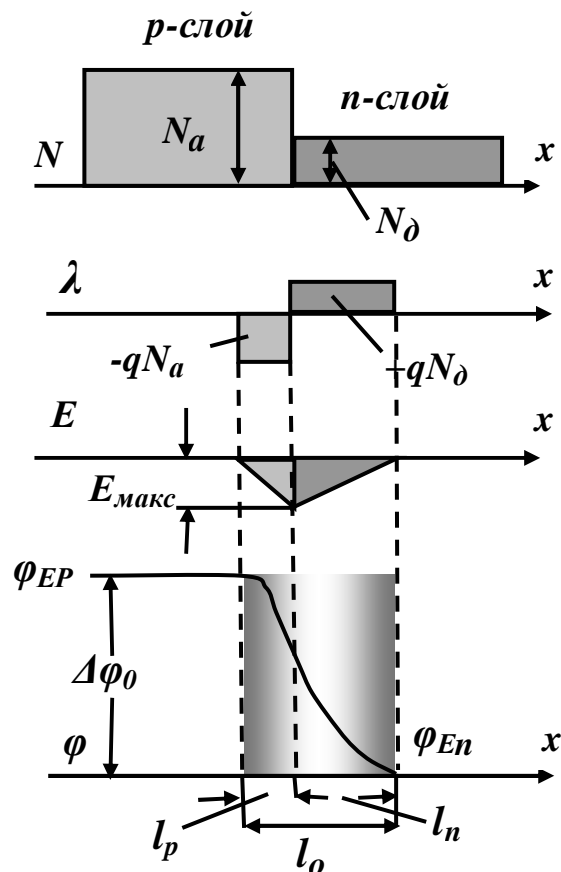


Рис. 2.6 — Распределение концентрации примесей (а), плотности заряда (б), напряженности поля (в) и потенциала (г) в ступенчатом p - n переходе

Напряженность электрического поля равна нулю вне перехода, если считать, что сопротивление перехода много больше сопротивления нейтральных слоев полупроводников и используя уравнения Пуассона, то напряженность в переходе изменяется по закону, показанному на рис. 2.6, в, а выражения для них запишутся в виде:

$$E_p = \frac{-qN_a}{\xi_0\xi}(x + l_p); x \leq 0; \quad (2.6a)$$

$$E_n = \frac{-qN_d}{\xi_0\xi}(l_n - x); x \geq 0. \quad (2.6b)$$

При этом потенциал в пределах перехода будет изменяться по квадратичному закону с точкой перегиба в месте излома кривой $E(x)$ (рис. 2.6, з).

$$\varphi_p - \varphi_{Ep} = -\frac{qN_a}{2\varepsilon_0\varepsilon}(x + l_p)^2; x \leq 0; \quad (2.7a)$$

$$\varphi_n - \varphi_{En} = -\frac{qN_d}{2\varepsilon_0\varepsilon}(x - l_n)^2; x \geq 0, \quad (2.7b)$$

где φ_{Ep} и φ_{En} — электростатические потенциалы соответствующих слоев вне перехода. Приравнивая значения напряженностей на металлургической границе $x = 0$ в выражениях (2.6a и 2.6b) E_{p0} и E_{n0} , получаем соотношение между шириной перехода в слоях и концентрациями p и n .

$$\frac{l_p}{l_n} = \frac{N_d}{N_a}. \quad (2.8)$$

Для несимметричного перехода $N_A \gg N_D$, следовательно, $l_p \ll l_n$, а значит, $l_0 \approx l_n$, т.е. можно считать, что переход полностью сосредоточен в высокоомном полупроводнике n .

Приравнивая $\varphi_p(0)$ и $\varphi_n(0)$ в выражениях (2.7a и 2.7b), используя соотношения $l_0 = l_n + l_p$ (2.8) и (2.3), можно получить зависимость между высотой барьера $\Delta\varphi_0$ и шириной перехода l_0 в следующем общем виде:

$$l_0 = \sqrt{\frac{2\xi_0\xi\Delta\varphi_0}{q}\left(\frac{1}{N_D} + \frac{1}{N_a}\right)}. \quad (2.9a)$$

Для несимметричного перехода при $N_a \gg N_D$ получаем:

$$l_0 = \sqrt{\frac{2\xi_0\xi\Delta\varphi_0}{qN_D}}. \quad (2.9б)$$

Используя формулу (1.20в), выражение для ширины перехода можно записать в виде:

$$l_0 = \sqrt{\frac{2\varepsilon_0\varepsilon\varphi_T \ln \frac{N_D N_a}{n_i^2}}{qN_D}}. \quad (2.9в)$$

Из выражения (2.9в) следует, чем выше концентрация доноров, тем меньше ширина перехода, т.е. влияние высокоомного полупроводника более существенно по сравнению с низкоомным. Увеличение концентрации доноров приводит к увеличению плотности заряда в этом слое, не изменяя плотности заряда в низкоомном полупроводнике. Учитывая, что заряды в переходе слева и справа от металлургической границы должны быть равны по абсолютной величине, ширина перехода в полупроводнике n уменьшится.

Анализ перехода в неравновесном состоянии

Подключим источник напряжения. между p - и n -слоями. Напряжение нарушает равновесие в системе и вызывает изменение величины тока. При этом высота потенциального барьера должна измениться, так как при равновесном значении $\Delta\varphi_0$ потоки носителей через переход уравновешены и ток отсутствует. Выше было показано, что удельное сопротивление обедненного слоя на несколько порядков выше, чем удельное сопротивление основных p - и n -слоев диода. Поэтому внешнее напряжение почти полностью падает на переходе, т.е. изменение высоты потенциального барьера должно быть равно величине приложенного напряжения. Когда внешнее напряжение приложено плюсом к p -слою, высота барьера уменьшается (рис. 2.7, а) и становится равной:

$$\Delta\varphi = \Delta\varphi_0 - U. \quad (2.10)$$

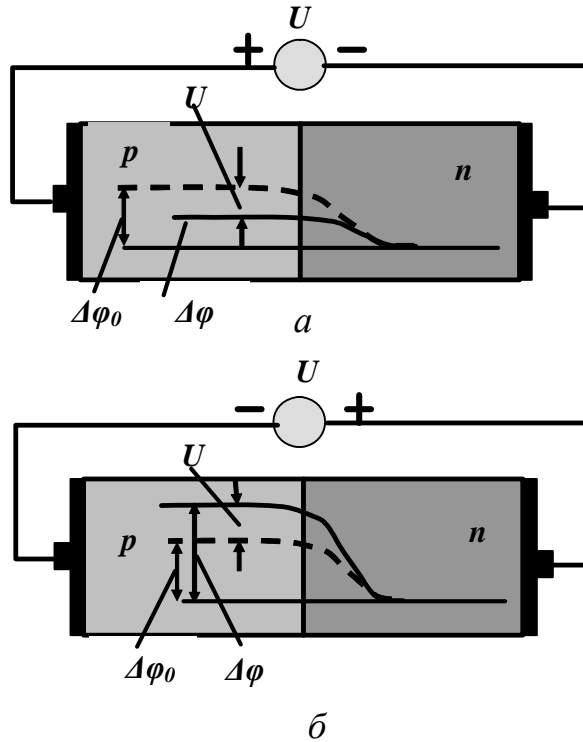


Рис. 2.7 — Смещение перехода в прямом (а) и обратном (б) направлениях

Такое включение перехода называется прямым. При отрицательном потенциале на p -слое (рис. 2.7, б) высота барьера увеличивается, и в формуле (2.10) нужно изменить знак перед U . Такое включение называется обратным. Изменение высоты барьера с помощью внешнего напряжения приводит к изменениям ширины перехода и граничных концентраций носителей. Рассмотрим эти изменения применительно к несимметричному переходу с низкоомным p -слоем. Подставляя значение $\Delta\varphi$ из (2.10) в (2.96), получаем:

$$l = \sqrt{\frac{2\xi\xi_0(\Delta\varphi_0 - U)}{qN_D}} = l_0 \sqrt{\frac{\Delta\varphi_0 - U}{\Delta\varphi_0}}, \quad (2.11)$$

где l_0 — ширина потенциального барьера при термодинамическом равновесии перехода. Положительное приращение потенциала соответствует уменьшению потенциальной энергии электронов, т.е. сдвигу энергетических уровней зоны проводимости полупроводника p вниз, уровней валентной зоны полупроводника n вверх.

Как видим, переход сужается при прямом напряжении ($U > 0$) и расширяется при обратном напряжении ($U < 0$). При прямом смещении вследствие идеализации перехода выражение (2.11) имеет существенную погрешность и отражает скорее качественную картину изменения ширины перехода. В то время как при обратном напряжении, удовлетворяющем неравенству $U \gg \varphi_T$, выражение (2.11) оказывается достаточно точным и широко используется на практике. В практических случаях часто выполняется условие $|U| \gg \Delta\varphi_0$. В этом случае ширину перехода можно записать в виде:

$$l = \sqrt{\frac{2\xi\xi_0}{qN_D} |U|} = l_0 \sqrt{\frac{|U|}{\Delta\varphi_0}}. \quad (2.12)$$

Из выражения (2.12) следует, что изменения внешнего напряжения приводит к изменениям ширины перехода, данное свойство широко используется на практике в специализированных диодах — **варикапах**.

Если в формулы (2.4) подставить высоту потенциального барьера (2.10) вместо равновесного значения $\Delta\varphi_0$ и считать концентрации основных носителей n_{n0} и p_{p0} неизменными, для граничных концентраций неосновных носителей получим выражения:

$$p_n = p_{p0} e^{-\Delta\varphi/\varphi_T} = \left(p_{p0} e^{-\Delta\varphi_0/\varphi_T} \right) e^{U/\varphi_T};$$

$$n_p = n_{n0} e^{-\Delta\varphi/\varphi_T} = \left(n_{n0} e^{-\Delta\varphi_0/\varphi_T} \right) e^{U/\varphi_T}.$$

Учитывая, что в скобках стоят равновесные граничные концентрации, определяемые формулами (2.4), запишем полученные выражения в следующей форме

$$p_n = p_{n0} e^{U/\varphi_T}; \quad (2.13a)$$

$$n_p = n_{p0} e^{U/\varphi_T}. \quad (2.13b)$$

Если напряжение U приложено в прямом направлении, то концентрации p_n и n_p на границах перехода согласно (2.13) воз-

растают по сравнению с равновесными значениями p_{n0} и n_{p0} . При $U = 3\phi_T$ концентрации неосновных носителей увеличиваются более чем в 25 раз. Иначе говоря, в каждом из слоев появляются избыточные неосновные носители, т.е. имеет место инжекция (рис. 2.8, а). Если напряжение U приложено в обратном направлении, то граничные концентрации p_n и n_p уменьшаются по сравнению с равновесными значениями, т.е. имеет место экстракция (рис. 2.8, б).

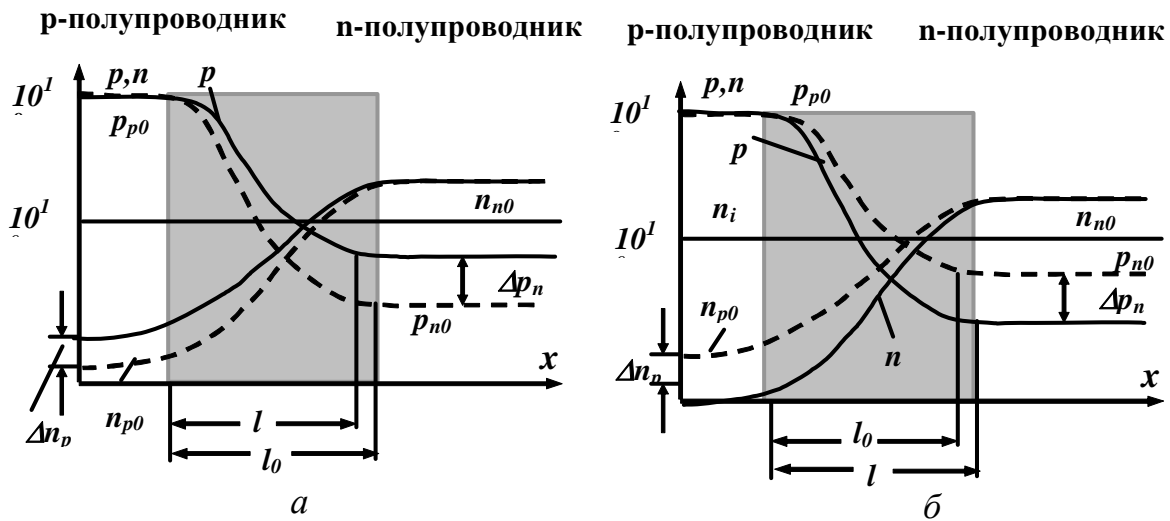


Рис. 2.8 — Распределение носителей в переходе при прямом (а) и обратном (б) напряжениях. Пунктиром показаны распределения в равновесном состоянии

Избыточные концентрации на границах перехода можно найти, вычитая из p_n и n_p соответственно величины равновесных концентраций p_{n0} и n_{p0} . Выражения для величин избыточных концентраций можно записать в виде:

$$\Delta p_n = p_{n0} (e^{U/\phi_T} - 1); \quad (2.14a)$$

$$\Delta n_p = n_{p0} (e^{U/\phi_T} - 1). \quad (2.14б)$$

Сравним избыточные граничные концентрации в слоях p и n , разделив (2.14а) на (2.14б) и заменив в правой части концентрации p_{n0} и n_{p0} на p_{p0} и n_{n0} по формуле (1.16). Тогда:

$$\frac{\Delta p_n}{\Delta n_p} = \frac{p_{p0}}{n_{n0}}. \quad (2.15)$$

Для несимметричных переходов концентрации p_{p0} и n_{n0} сильно различаются, поэтому концентрация инжектированных неосновных носителей будет гораздо больше в высокоомном слое, чем в низкоомном. Таким образом, в несимметричных переходах можно считать, что инжекция носит односторонний характер: неосновные носители инжектируются в основном из низкоомного слоя в высокоомный. **Инжектирующий слой с относительно малым удельным сопротивлением называют эмиттером, а слой с относительно большим удельным сопротивлением, в который инжектируются неосновные для него носители, — базой.**

Формулы (2.13) подтверждают, что в режиме экстракции граничные концентрации неосновных носителей могут быть сколь угодно малы, но всегда положительны.

2.3 Контакты металл-полупроводник

Контакты полупроводника с металлом играют важную роль в полупроводниковых приборах. Структура и свойства этих контактов зависят в первую очередь от взаимного расположения уровней Ферми в исходных слоях. На рис. 2.9 вверху показаны зонные диаграммы разделенных слоев, а внизу — зонные диаграммы соответствующих контактов. Поскольку уровень Ферми в металле всегда расположен в зоне проводимости и для равновесной системы должен быть единым, нетрудно в каждом случае оценить характер результирующей диаграммы. Так, на рис. 2.9, а имеем $\varphi_{Fm} > \varphi_{Fp}$. Это значит, что энергетические уровни, соответствующие зоне проводимости полупроводника, заполнены в металле больше, чем в полупроводнике. Следовательно, после соприкосновения слоев часть электронов перейдет из металла в полупроводник и концентрация электронов в металле в области, прилегающей к металлургической границе, уменьшится. Наличие дополнительных электронов приводит к уменьшению расстояния между уровнем Ферми и дном зоны проводимости в этой облас-

ти, поэтому энергетические уровни полупроводника искривляются вниз. Знаки плюс и минус говорят о том, что в области металлургической границы образуется равновесный поверхностный потенциал φ_{so} . На рис. 2.9, б показан обратный случай, когда после соприкосновения слоев электроны переходят из полупроводника в металл и соответственно уровни искривляются вверх. Область искривления зон (т.е. область пространственных зарядов) имеет протяженность порядка дебаевской длины.

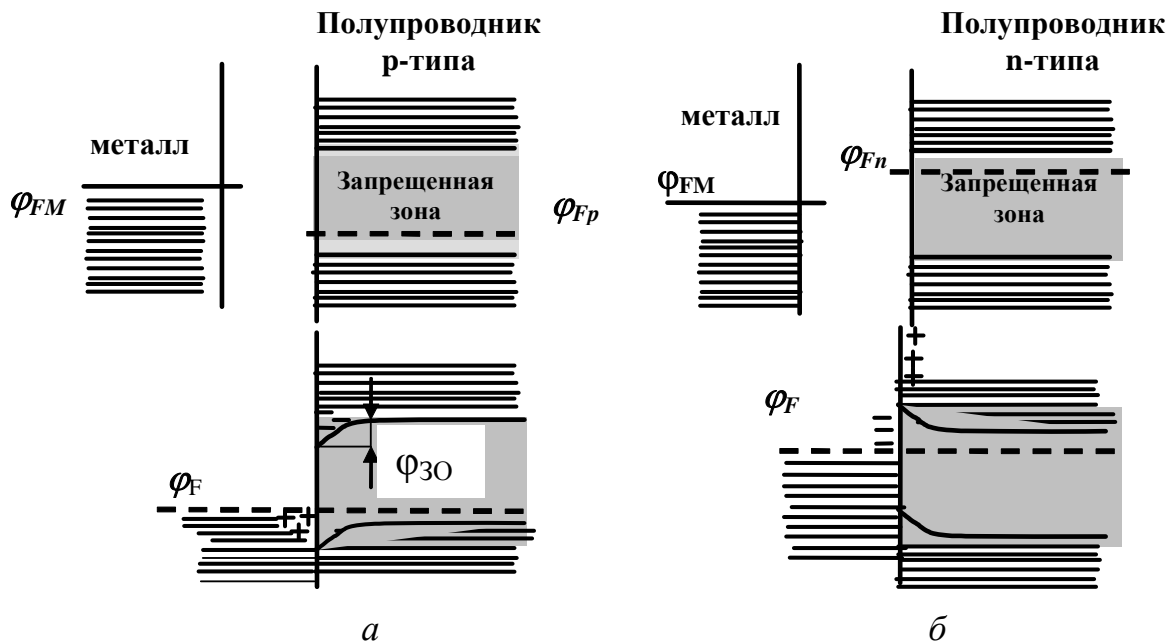


Рис. 2.9 — Зонные диаграммы выпрямляющих контактов металла с полупроводником: а — контакт с полупроводником p -типа; б — контакт с полупроводником n -типа

Оба контакта, показанные на рис. 2.9, а и б, характерны тем, что концентрация основных носителей на границе полупроводника понижена по сравнению с концентрацией вдали от контакта. Следовательно, граничный слой обладает повышенным удельным сопротивлением и поэтому определяет сопротивление всей системы. В зависимости от полярности приложенного напряжения меняется высота приповерхностного потенциального барьера.

Так, если внешнее напряжение приложено плюсом к металлу и минусом к полупроводнику, то потенциальный барьер на рис. 2.9, а повышается, а на рис. 2.9, б понижается. При этом граничный слой на рис. 2.9, а еще больше обедняется дырками и бу-

дет иметь повышенное сопротивление, а граничный слой на рис. 2.9, *б* обогащается электронами и будет иметь пониженное сопротивление по сравнению с равновесным состоянием. Значит, такая полярность будет обратной для контактов на рис. 2.9, *а* и прямой для контактов на рис. 2.9, *б*.

Если изменить полярность приложенного напряжения, то по аналогичным причинам сопротивление граничного слоя на рис. 2.9, *а* понизится и контакты будут работать в прямом направлении, а сопротивление граничного слоя на рис. 2.9, *б* повысится, т.е. система будет работать в обратном направлении.

Таким образом, контакты на рис. 2.9, *а* и *б* хотя и не обеспечивают инжекции, но обладают вентильными свойствами и лежат в основе диодов Шоттки. Принципиальное отличие таких контактов от системы с образованием *p-n* перехода заключается в том, что ток обусловлен основными носителями, подвижность которых значительно выше неосновных. Кроме того, отсутствие режима инжекции не приводит к накоплению и рассасыванию избыточных носителей заряда. Два этих фактора способствуют существенному повышению быстродействия в диодах Шоттки.

Иная картина получается тогда, когда для контакта металла с полупроводником *p*-типа имеет место соотношение $\varphi_{Fm} < \varphi_{Fp}$, а для контакта металла с полупроводником *n*-типа — соотношение $\varphi_{Fm} > \varphi_{Fn}$. В этих случаях искривление зон в полупроводниках получается обратным по сравнению с тем, какое показано на рис. 2.9, *а* и *б*, т.е. граничные слои оказываются не обедненными, а обогащенными основными носителями (рис. 2.10). Соответственно, удельные сопротивления граничных слоев оказываются ниже, чем у основных слоев полупроводника вдали от границы, так что наличие граничного слоя в системе малосущественно с точки зрения ее суммарного сопротивления. Следовательно, результирующее сопротивление системы близко к сопротивлению основного слоя полупроводника и почти не зависит от полярности и величины внешнего напряжения. Такие контакты называют омическими, или невыпрямляющими. Невыпрямляющие контакты являются основой омических контактов.

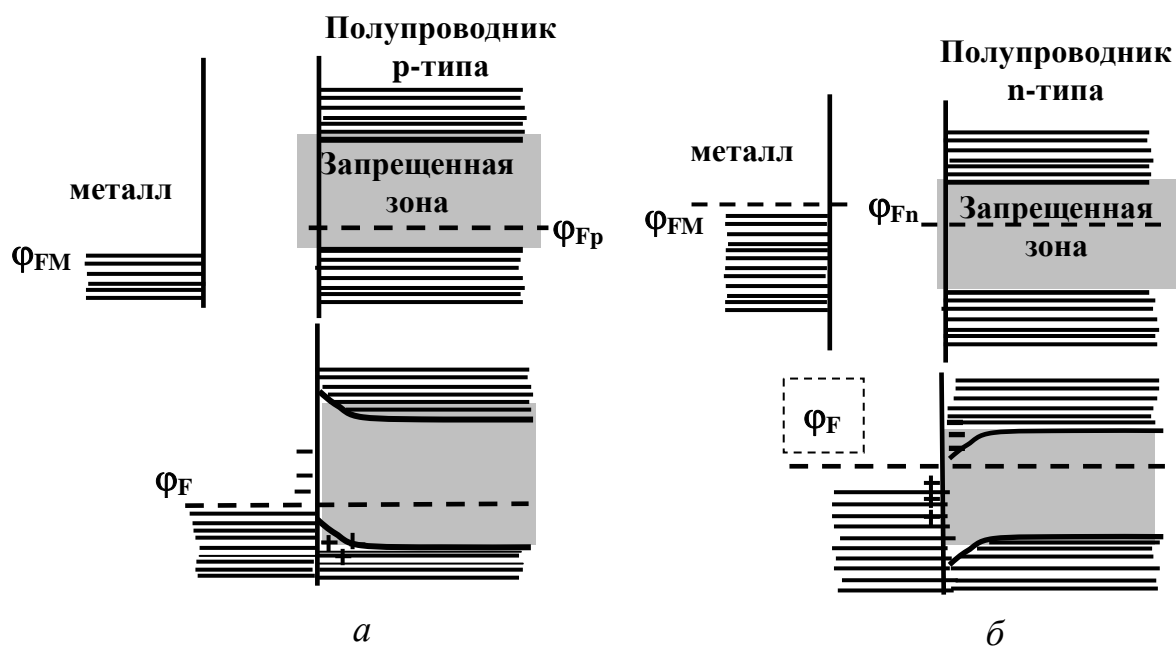


Рис. 2.10 — Зонные диаграммы невыпрямляющих контактов металла с полупроводником: *а* — контакт с полупроводником *p*-типа; *б* — контакт с полупроводником *n*-типа

2.4 Анализ идеализированного диода

Несмотря на то, что диод представляет собой один из простейших полупроводниковых приборов, процессы, происходящие в нем, достаточно сложны. Для того чтобы выяснить основные параметры и характеристики диода, проведем сначала упрощенный анализ, а затем в последующих параграфах уточним полученные результаты.

Исходные предпосылки. Будем, как и раньше, считать *p-n* переход несимметричным и *p*-слой значительно больше легированным, чем *n*-слой. При этом, как известно, инжекция и экстракция носят односторонний характер, а значит, можно сосредоточить внимание на анализе процессов в базе и результаты анализа распространить на аналогичные, но менее существенные процессы в эмиттере.

Анализ существенно упрощается, если принять следующие допущения, а в дальнейшем ввести уточнения.

1. Слой базы является ярко выраженным электронным полупроводником. Это значит, что вместо «объединенного» элек-

тронно-дырочного уравнения можно пользоваться уравнением (1.33а) и положить $\frac{\partial E}{\partial x} = 0$.

2. Концентрация дырок, инжектируемых в базу, невелика, т.е. выполняется условие низкого уровня инжекции. При этом полная концентрация дырок в базе $p = p_0 = \Delta p$ остается значительно меньше концентрации электронов $n = n_0$. Следовательно, можно пренебречь дрейфовой составляющей дырочного тока в базе. По аналогичным причинам можно пренебречь дрейфовой составляющей электронного тока в эмиттере. Соответственно вместо уравнения непрерывности можно использовать уравнения диффузии.

3. Падение напряжения в базе (а тем более в низкоомном эмиттерном слое) значительно меньше внешнего напряжения, так что последнее можно считать приложенным непосредственно к переходу.

4. Ширина перехода настолько мала, что процессами генерации и рекомбинации в области перехода можно пренебречь. Это дает право считать электронные токи на обеих границах перехода одинаковыми; то же самое относится к дырочным токам. Сразу же оговоримся, что это допущение очень часто не реализуется, особенно для кремния и арсенида галлия, т.е. для полупроводниковых материалов с относительно большой шириной запрещенной зоны.

5. Обратные напряжения значительно меньше напряжения пробоя, так что можно пренебречь предпробойными явлениями в переходе.

6. Отсутствуют всякого рода поверхностные утечки, шунтирующие переход, а следовательно, и токи утечки, которые добавляются к токам, обусловленным инжекцией и экстракцией.

Учитывая принятые допущения, работу диода можно описать следующим образом. При прямом смещении перехода концентрация дырок на его базовой границе повышается, и эти избыточные дырки диффундируют в глубь базы. По мере удаления от перехода концентрация дырок убывает и в установившемся режиме получается некоторое распределение избыточных дырок $\Delta p(x)$ (рис. 2.11, а). Инжекция дырок в базу нарушает ее ней-

тральность и вызывает приток избыточных электронов из внешней цепи. Эти электроны распределяются таким образом, чтобы компенсировать поле дырок, т.е. накапливаются в той же области, что и дырки. Поэтому кривые $\Delta p(x)$ и $\Delta n(x)$ оказываются почти одинаковыми. Небольшая разница между этими кривыми обусловлена различием подвижностей электронов и дырок (эффект Дембера). В установившемся режиме в базе протекает диффузионный дырочный ток, пропорциональный градиенту концентрации в каждой точке кривой $\Delta p(x)$.

Так как полный ток диода должен быть одинаковым в любом сечении, то уменьшение диффузионного дырочного тока сопровождается ростом электронной составляющей.

Структура полного тока рассмотрена нами раньше. Однако величину полного тока удастся вычислить без учета этой структуры, если воспользоваться сделанными выше допущениями. Действительно, учитывая допущения 2 и 4, можем для базовой границы перехода записать:

$$j(0) = j_p(0) + j_n(0) = j_{p\text{диф}}(0) + j_{n\text{диф}}(-l),$$

где координата $x = -l$ соответствует эмиттерной границе. В одномерном случае плотность тока $j(0)$ сохраняется в любом сечении.

Реальные структуры полупроводниковых диодов и транзисторов неоднородные, однако анализ проводится применительно к одномерной модели (в данном случае применительно к модели, показанной на рис. 2.8), после чего в случае необходимости делаются те или иные поправки.

Таким образом, чтобы рассчитать ток диода, нужно, зная величину приложенного напряжения, найти распределения дырок в базе и электронов в эмиттере, определить градиенты этих распределений соответственно в точках $x = 0$ и $x = -l$ и затем с помощью формул, известных нам из первой главы, получить компоненты полного тока $j_{p\text{диф}}(0)$ и $j_{n\text{диф}}(-l)$.

Решение диффузионного уравнения. Чтобы получить статическую вольт-амперную характеристику диода, нужно найти стационарное распределение дырок в базе. Для этого в уравнении

диффузии следует положить $\frac{\partial \Delta p}{\partial t} = 0$, после чего оно приводится к виду

$$\frac{d^2(\Delta p)}{dx^2} - \frac{\Delta p}{L^2} = 0, \quad (2.16)$$

где для простоты опущен индекс p при параметре L .

Поскольку в базе $n \gg p$, кривые $n(x)$ должны были бы лежать намного выше, чем кривые $p(x)$. Чтобы этого избежать, на рис. 2.11 сделан разрыв на оси ординат.

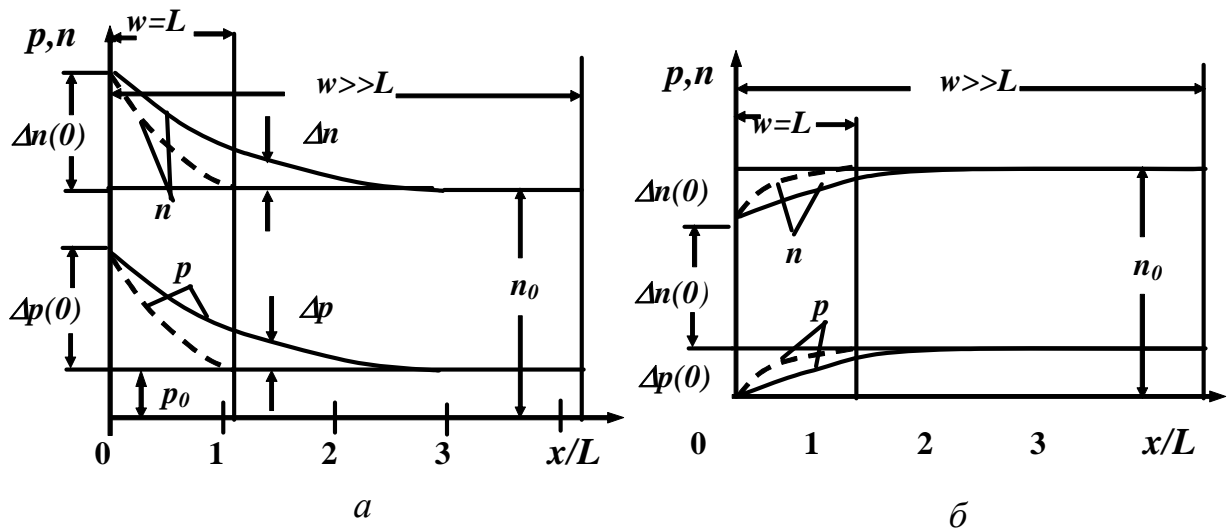


Рис. 2.11 — Распределение носителей в диодах с толстой и тонкой базой:
а — при прямом смещении; б — при обратном смещении

Как известно, решением (2.16) является сумма двух экспонент:

$$\Delta p(x) = A_1 e^{x/L} + A_2 e^{-x/L}. \quad (2.17)$$

Для того чтобы в решении (2.17) определить коэффициенты A_1 и A_2 , нужно знать граничные условия. В разделе 1.13 (разд. «Монополярная диффузия») значение $\Delta p(0)$ использовалось без расшифровки, а толщина базы принималась бесконечно большой, и соответственно полагалось $\Delta p(\infty) = 0$. Теперь, учитывая конечные размеры базы, выразим граничную концентрацию $\Delta p(0)$ через приложенное напряжение с помощью (2.14а):

$$\Delta p(0) = p_0 \left(e^{U/\phi_T} - 1 \right). \quad (2.18a)$$

Вторую граничную концентрацию запишем в виде

$$\Delta p(w) = 0, \quad (2.18b)$$

поскольку концентрации носителей на омическом контакте сохраняют равновесное значение независимо от распределения концентраций в базе. При выполнении граничных условий (2.18) коэффициенты A_1 и A_2 имеют значения:

$$A_1 = \frac{\Delta p(0)}{2sh\left(\frac{w}{L}\right)} e^{-w/L}, \quad A_2 = \frac{\Delta p(0)}{2sh\left(\frac{w}{L}\right)} e^{w/L}, \quad (2.19)$$

а распределение $\Delta p(x)$ принимает вид:

$$\Delta p(x) = p_0 \left(e^{U/\phi_T} - 1 \right) \frac{sh\left(\frac{w-x}{L}\right)}{sh\left(\frac{w}{L}\right)}. \quad (2.20)$$

В случае толстой базы $|w > (2-3)L|$ можно положить $w \rightarrow \infty$ Тогда коэффициенты A_1 и A_2 упрощаются:

$$A_1 = 0; \quad A_2 = \Delta p(0),$$

а распределение дырок оказывается экспоненциальным

$$\Delta p(0) = p_0 \left(e^{U/\phi_T} - 1 \right) e^{-x/L}. \quad (2.21a)$$

В случае тонкой базы $|w < 0.5L|$ можно считать $shz \approx z$. Тогда из (2.20) получаем практически линейное распределение примесей в базе

$$\Delta p(x) = p_0 \left(e^{U/\phi_T} - 1 \right) \left(1 - \frac{x}{w} \right), \quad (2.21b)$$

которое более характерно для реальных диодов. На рис. 2.11, а распределения (2.21а) и (2.21б) показаны соответственно сплошными и пунктирными линиями. Для режима экстракции ($U < 0$) все выведенные формулы остаются в силе, а соответствующие распределения дырок и электронов показаны на рис. 2.11, б.

Вольт-амперная характеристика. Для получения вольт-амперной характеристики продифференцируем выражение (2.20) по координате x и, подставив в (1.20а), получим распределение плотности дырочного тока в базе

$$j_p(x) = \frac{qD_{p\bar{\delta}}P_{0\bar{\delta}}}{L_{p\bar{\delta}}} \left(e^{U/\Phi_T} - 1 \right) \frac{ch \frac{w_{\bar{\delta}} - x}{L_{p\bar{\delta}}}}{sh \frac{w_{\bar{\delta}}}{L_{p\bar{\delta}}}}. \quad (2.22a)$$

Здесь для определенности введены индексы для базового слоя и для «рабочих» носителей — дырок. По аналогии можно записать для плотности электронного тока в эмиттерном слое:

$$j_n(x) = \frac{qD_{n\bar{\delta}}n_{0\bar{\delta}}}{L_{n\bar{\delta}}} \left(e^{U/\Phi_T} - 1 \right) \frac{ch \frac{w_{\bar{\delta}} - x}{L_{n\bar{\delta}}}}{sh \frac{w_{\bar{\delta}}}{L_{n\bar{\delta}}}}, \quad (2.22b)$$

где $L_{n\bar{\delta}}$ — диффузионная длина для электронов в эмиттере, а координата x отсчитывается от перехода в глубь эмиттера. Полагая $x = 0$ в формулах (2.22), умножая обе части на площадь s и складывая токи $I_p(0)$ и $I_n(0)$, получаем искомую вольт-амперную характеристику идеализированного диода. Необходимо помнить, что уменьшение диффузионной составляющей плотности тока сопровождается ростом дрейфовой составляющей. Обычно статическую вольт-амперную характеристику записывают в следующей форме:

$$I = I_0 \left(e^{U/\Phi_T} - 1 \right), \quad (2.23)$$

где

$$I_0 = \frac{qD_{p\bar{\delta}}s}{L_{p\bar{\delta}}th\left(\frac{w_{\bar{\delta}}}{L_{p\bar{\delta}}}\right)} P_{0\bar{\delta}} + \frac{qD_{n\bar{\delta}}s}{L_{n\bar{\delta}}th\left(\frac{w_{\bar{\delta}}}{L_{n\bar{\delta}}}\right)}. \quad (2.24)$$

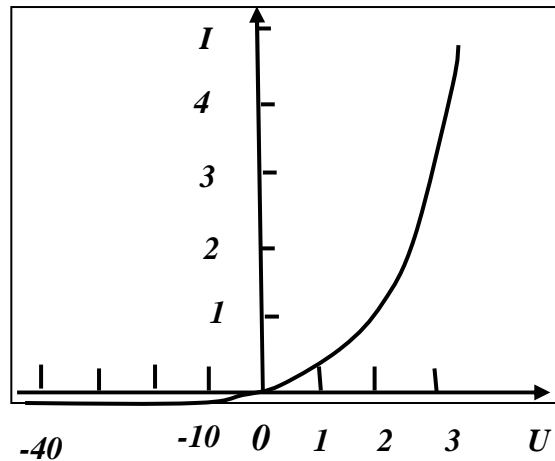


Рис. 2.12 — Статическая вольт-амперная характеристика идеализированного плоскостного диода

Формула (2.23) — одна из важнейших в транзисторной технике — представлена на рис. 2.12. Ток I_0 , определяющий «масштаб» характеристики, называется **тепловым током**. Термин «тепловой» отражает сильную температурную зависимость тока I_0 , а также тот факт, что он равен нулю при абсолютном нуле температуры. Другим распространенным термином является «обратный ток насыщения», происхождение которого связано с тем, что при отрицательном напряжении $|U| \gg \phi_T$ обратный ток идеализированного диода равен — I_0 и не зависит от напряжения.

Вторым слагаемым в (2.24) обычно пренебрегают, т.к. инжекция электронов из базы в эмиттер очень мала. Тогда тепловой ток равен

$$I_0 \approx q \frac{DS}{L \operatorname{th}(w/L)} p_{0б}. \quad (2.25a)$$

Если выполняется условие $w \ll L$ $\operatorname{th}\left(\frac{w}{L}\right) \approx w/L$,

$$I_0 \approx q \frac{DS}{w} p_0. \quad (2.25b)$$

При $w \gg L$ $\operatorname{th}\left(\frac{w}{L}\right) \approx 1$

$$I_0 \approx q \frac{Ds}{L} p_0. \quad (2.26)$$

Свойства теплового тока будут подробно рассмотрены позже. Сейчас только заметим, что тепловой ток диода пропорционален концентрации неосновных носителей электронного полупроводника $p_{0б}$, которая жестко связана с собственной концентрацией n_i . Собственная концентрация n_i у кремния гораздо меньше, чем у германия, следовательно, и тепловой ток у кремниевых диодов несравненно меньше, чем у германиевых.

Одной из важных особенностей характеристики (2.23) является очень крутая (экспоненциальная) прямая ветвь. Поэтому весьма большие прямые токи порядка нескольких ампер и выше получаются у полупроводниковых диодов при напряжении не более 1 В, т.е. намного меньшем, чем в случае вакуумных и газонаполненных диодов. В связи с большой крутизной прямой ветви обычно удобнее задавать в качестве аргумента ток, а напряжение считать его функцией, т.к. задать на практике напряжение с высокой точностью бывает трудно. Соответственно, формулу (2.23) целесообразно преобразовать к следующему виду

$$U = \varphi_T \ln \left(\frac{I}{I_0} + 1 \right). \quad (2.27)$$

Вентильные свойства любого диода выражены тем ярче, чем меньше обратный ток при заданном обратном напряжении и чем меньше прямое напряжение при заданном прямом токе. К сожалению, эти два требования в данном случае противоречивы. В самом деле, из формулы (2.26) видно, что изменение теплового тока, какими бы причинами оно ни вызывалось, сопровождается изменением прямого напряжения в противоположном направлении. Это хорошо видно из рис. 2.13, а, где различие токов I_0 обусловлено разницей в площадях переходов при прочих равных условиях. Важным следствием этой общей зависимости является то, что прямые напряжения у кремниевых диодов заметно больше, чем у германиевых, поскольку тепловой ток у первых на несколько порядков меньше. Различие в прямых напряжениях германиевых и кремниевых диодов составляет обычно **0,3–0,4 В** и сохраняется вплоть до малых токов.

Характеристические сопротивления. Нелинейность характеристики удобно оценивать, сопоставляя сопротивления диода при смещениях в прямом и обратном направлениях.

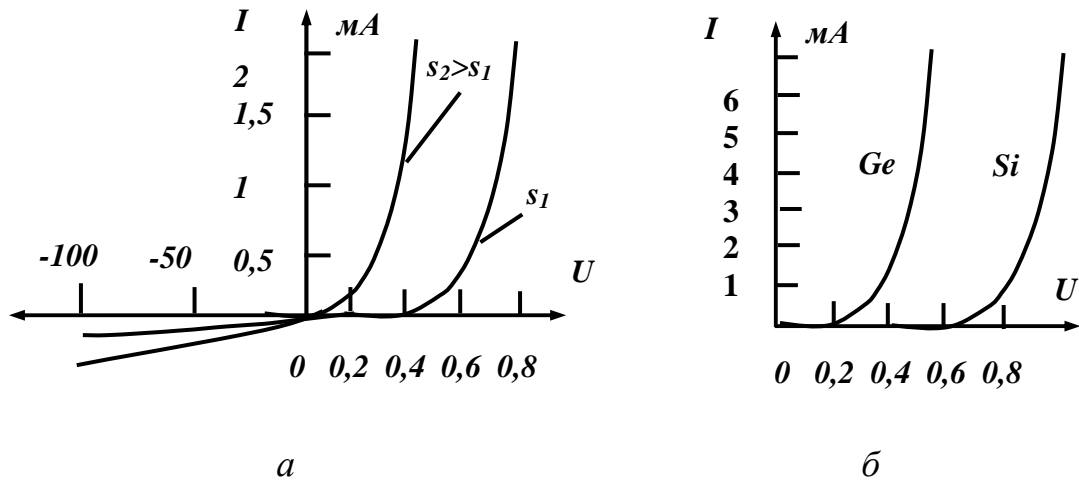


Рис. 2.13 — Характеристики идеализированных диодов с разными площадями переходов (а) и разными тепловыми токами — германиевого и кремниевого (б)

Как и в ламповой технике, различают дифференциальные сопротивления и сопротивления постоянному току.

Найдя производную от функции (2.26), легко представить дифференциальное сопротивление как функцию тока (рис. 2.14, а)

$$r_D = \frac{dU}{dI} = \frac{\Phi_T}{I + I_0} \approx \frac{\Phi_T}{I}. \quad (2.28)$$

Приближенное выражение, разумеется, действительно только для прямой ветви при условии $I \gg I_0$. На обратной ветви величина r_D резко возрастает, а на прямой ветви величина r_D , наоборот, быстро уменьшается. При таком сопротивлении изменение прямого напряжения даже на **5–10 мВ** приводит к значительным изменениям тока. Поэтому задать прямое напряжение с целью получить нужный ток весьма трудно, и для полупроводникового диода, работающего в прямом направлении, более характерен режим генератора тока. На это указывалось при выводе формулы (2.27). Часто при расчете диодных схем пользуются сопротивлением постоянному току — статическое сопротивление. Из формулы (2.23) получаем зависимость такого сопротивления от тока (рис. 2.14, а)

$$R_D = \frac{U}{I} = \frac{\Phi_T}{I} \ln \left(\frac{I}{I_0} + 1 \right). \quad (2.29a)$$

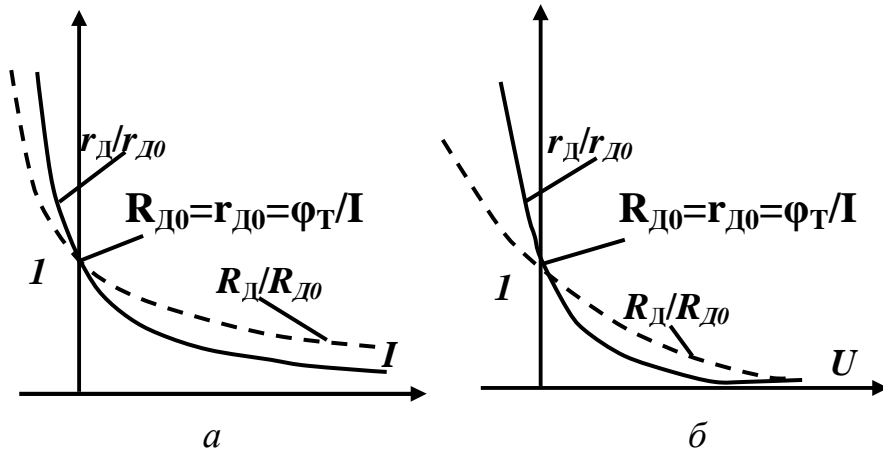


Рис. 2.14 — Зависимость сопротивлений идеализированного диода — дифференциального сопротивления r_D и R_D — от прямого тока (а) и прямого напряжения (б)

Из выражений 2.28 и 2.29 следует, что о дифференциальном и статическом сопротивлениях можно говорить только для конкретной точки вольт-амперной характеристики.

Из формулы (2.29а) получаем зависимость статического сопротивления от напряжения (рис. 2.14, б):

$$R_D = \frac{U}{I_0 \left(e^{U/\varphi_T} - 1 \right)}. \quad (2.296)$$

Отсюда следует, что на обратной ветви характеристики, когда $U \gg \varphi_T$, сопротивление R_D прямо пропорционально напряжению:

$$R_D = \frac{U}{I_0}.$$

В нулевой точке $U = 0$ и $I = 0$ сопротивления r_D и R_D , как нетрудно убедиться, имеют одно и то же значение

$$r_{D0} = R_{D0} = \frac{\varphi_T}{I_0}. \quad (2.29в)$$

Необходимо подчеркнуть, что формулы (2.27)–(2.29) выведены на основе выражения (2.23), которое не полностью отражает все свойства реального диода.

Поэтому на практике сопротивления могут значительно отличаться от значений, полученных по этим формулам. Правиль-

нее, видимо, говорить о характеристических сопротивлениях перехода в диоде и при необходимости в зависимости от значений сопротивлений нейтральных слоев и выводов вносить поправки.

2.5 Обратная и прямая характеристики реального диода

Обратная характеристика реального диода

Опыт показывает, что обратный ток диода не остается постоянным и равным I_0 , как следует из формулы (2.23). Он всегда более или менее сильно растет с увеличением обратного напряжения. Но и при малых напряжениях обратный ток больше теплового, особенно у кремниевых диодов, у которых это превышение достигает 2 — 3 порядков. Главные причины отклонения реальной обратной характеристики от идеализированной заключаются: в термогенерации носителей в области перехода, поверхностных утечках, а также в явлениях, которые при достаточно большом обратном напряжении приводят к пробоем перехода.

Таким образом, обратный ток диода состоит из нескольких компонентов, из которых при выводе характеристики (2.23) мы познакомились лишь с одним — тепловым током. Однако и этот ток следует рассмотреть подробнее.

Тепловой ток. Используя в формуле (2.26) соотношения (1.16), (1.21) и (1.28 — 1.30), нетрудно представить тепловой ток в следующем виде:

$$I_0 = \frac{b}{(b+1)^2} \frac{\varphi_T s \rho_{\bar{o}}}{L \rho_i^2},$$

где $b = \frac{\mu_a}{\mu_p}$. Например, для германиевых и кремниевых диодов при $\rho_{\bar{o}} = 5$ Ом см, $\tau = 5$ мксек и $s = 0,05$ см² получим $(I_0)_{Ge} = 30$ мкА, $(I_0)_{Si} = (3 \cdot 10) 10^{-7}$ мкА, т.е. различие в токах I_0 составляет около семи порядков. Маленький тепловой ток кремниевых диодов является одним из их важнейших преимуществ, хотя полный обратный ток реальных кремниевых диодов, как будет видно ниже,

значительно больше. В мощных германиевых диодах с большой площадью ток I_0 может достигать до нескольких миллиампер.

Запишем формулу (2.24) в несколько иной форме. Умножим и разделим слагаемые в правой части на соответствующие диффузионные длины, положим $th(w/L)=1$. Тогда

$$I_0 = q(SL_{p\bar{o}}) \frac{p_{0\bar{o}}}{\tau_{\bar{o}}} + q(SL_{n\bar{e}}) \frac{n_{0\bar{e}}}{\tau_{\bar{e}}}. \quad (2.30)$$

Поскольку множители $p_{0\bar{o}}/\tau_{\bar{o}}$ и $n_{0\bar{e}}/\tau_{\bar{e}}$ есть скорости генерации дырок в базе и электронов в эмиттере, выражение (2.30) можно трактовать следующим образом: тепловой ток обусловлен генерацией неосновных носителей в объемах SL , прилегающих к переходу, откуда эти носители диффундируют в область потенциального барьера и уносятся полем в другой слой (рис. 2.15). В равновесном состоянии эти потоки компенсируются встречными потоками аналогичных носителей, имеющих достаточную энергию, чтобы преодолеть потенциальный барьер.

Особый интерес представляет температурная зависимость теплового тока. Если пренебречь электронной составляющей, то согласно выражению (2.24) тепловой ток зависит от температуры через параметры: коэффициент диффузии, время жизни носителей, концентрацию. Из них главную роль играет равновесная концентрация неосновных носителей p_0 .

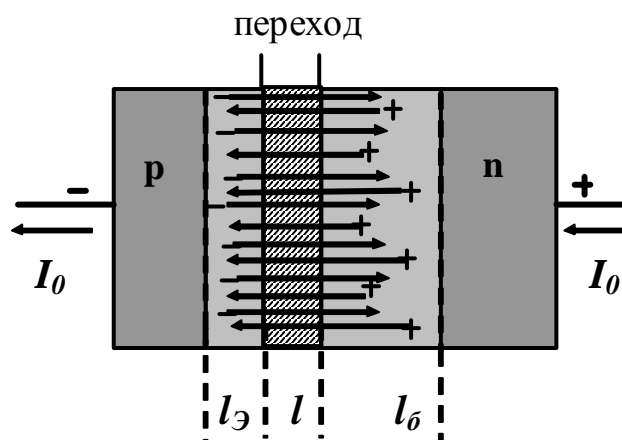


Рис. 2.15 — Происхождение электронного и дырочного компонентов теплового тока

Принимая во внимание выражение (1.21б), можно сделать заключение, что тепловой ток пропорционален квадрату собственной концентрации полупроводника. Подставляя значение n_i из (1.15), выразим тепловой ток в виде:

$$I_0(T) = I_{00} e^{-\varphi_3 / \varphi_T}. \quad (2.31a)$$

Здесь ток I_{00} содержит величины, мало зависящие от температуры. С ростом температуры примесный полупроводник постепенно превращается в собственный. В этом случае с учетом (1.15) выражение для теплового тока примет вид:

$$I_0(T) = I_{00} e^{-\varphi_3 / 2\varphi_T}. \quad (2.31б)$$

Использовать на практике сравнительно простую зависимость (2.31б) невозможно из-за сложности определения I_{00} .

На практике всегда известен тепловой ток при некоторой (обычно «комнатной») температуре T_0 , и требуется определить его значение при другой температуре. Из формулы (2.31б) легко получить:

$$\frac{I_0(T)}{I_0(T_0)} = e^{\frac{\varphi_3}{2} \left(\frac{1}{\varphi_{T_0}} - \frac{1}{\varphi_T} \right)}. \quad (2.32)$$

Учитывая, что абсолютная температура в рабочем диапазоне меняется не очень сильно (т.е. $TT_0 \approx T_0^2$), получаем вместо (2.32) простое приближенное соотношение

$$I_0(T) \approx I_0(T_0) e^{a\Delta T}, \quad (2.33)$$

где $a = 0,13\varphi_3$;

$$a_{Ge} = 0,09 \text{ град}^{-1};$$

$$a_{Si} = 0,13 \text{ град}^{-1}.$$

Этими значениями коэффициента a можно пользоваться вплоть до температур **70 — 80° C** для германия и **120 — 150° C** для кремния.

Несмотря на простоту, формула (2.33) не всегда удобна для быстрых «прикидок». Поэтому целесообразно заменить степень числа e степенью числа 2. Заменив основание степени по извест-

ным правилам, формулу (2.33) можно представить в следующем виде

$$I_0(t^0C) \approx I_{0(t=20^0C)} * 2^{\frac{t^0C-20}{10}}. \quad (2.34)$$

На практике весьма распространено правило: «тепловой ток удваивается на каждые $10^\circ C$ приращения температуры», что соответствует значению $a = 0,07$. Из изложенного ясно, что это правило не универсально и обычно занижает фактические изменения теплового тока в несколько раз. Главная трудность вычисления теплового тока по формуле 2.34 заключается в том, что значение теплового тока неизвестно для конкретного экземпляра. Поэтому выражением 2.34 можно пользоваться для качественной оценки изменения тока диода в диапазоне температур.

Выше отмечалось, что тепловой ток это только одна из составляющих обратного тока.

Ток термогенерации. В идеализированном диоде мы считали переход бесконечно узким и, следовательно, могли пренебречь генерацией и рекомбинацией носителей в этой области. Реальный переход имеет конечную ширину, поэтому эти процессы, имеющие место так же, как в любом другом слое полупроводника, играют немаловажную роль. Электрическое поле, которое всегда есть в переходе, быстро уносит генерируемые носители в соответствующий слой диода. Это вызывает протекание некоторого тока, который получил название тока термогенерации I_G . В равновесном состоянии диода этот ток компенсируется равным ему встречным током — током рекомбинации I_R . Ток рекомбинации обусловлен теми носителями, которые непрерывно проникают в переход из эмиттера и базы, но не имеют достаточной энергии, чтобы перейти в смежный слой. Вблизи точки «отражения» такие носители имеют малую скорость и успевают рекомбинировать. В неравновесном состоянии диода взаимная компенсация токов нарушается. В случае обратного включения диода превалирует ток термогенерации, так как высота потенциального барьера увеличивается и проникание носителей в область перехода с последующей рекомбинацией затрудняется. При величине смещения $|U \gg \phi_T|$ практически остается только ток термогене-

рации, величина которого пропорциональна объему «генерирующего» обедненного слоя, т.е. ширине перехода (рис. 2.16). Ток I_G накладывается на тепловой ток I_0 , и результирующий обратный ток оказывается больше, чем это следует из формулы (2.24). Более того, ток растет с расширением перехода, т.е. с ростом напряжения, что приводит к конечному наклону обратной ветви характеристики. Соответственно, ток термогенерации по аналогии с выражением (2.30) запишется в следующем виде:

$$I_G = q(Sl) \frac{n_i}{\tau_\infty}, \quad (2.35)$$

где l — ширина перехода. Пусть, например, $S = 0,05 \text{ см}^2$, $l = 1,25 \text{ мкм}$, $\tau_\infty = 4 \text{ мксек}$; тогда из формулы (2.35) для кремниевого диода получаем $I_G = 0,005 \text{ мкА}$. Для германиевого диода ток I_{Ge} будет в 1 000 раз больше, т.е. около 5 мкА. Сравним токи I_G и I_0 . Разделив (2.35) на первый член (2.30), выразив концентрации через удельные сопротивления и полагая для простоты $\tau_\infty = \tau_\delta$ и $\mu_m = \mu_z$, получим:

$$\frac{I_G}{I_0} \approx \rho_i / \rho_\delta \frac{l}{L}. \quad (2.36)$$

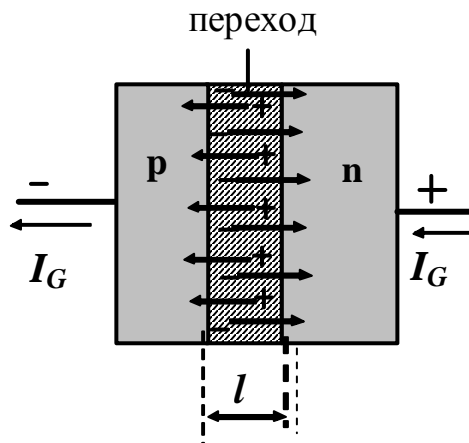


Рис. 2.16 — Происхождение тока термогенерации в переходе

Для германия при $\rho_\delta \text{ Ом см}^{-1}$, $l = 1 \text{ мкм}$, $L = 150 \text{ мкм}$ и комнатной температуре отношение токов составит около 0,1. Для

кремния при прочих равных условиях отношение токов получится около 1000. Таким образом, при комнатной температуре ток термогенерации в германии много меньше теплового.

Значения токов термогенерации при любой температуре можно рассчитать по формуле (2.34). Заметим, что у кремниевых диодов ток термогенерации является главным компонентом обратного тока при комнатной температуре. С повышением температуры тепловой ток растет быстрее. Обычно это происходит при температуре $+100\text{ }^\circ\text{C}$ и выше. У германиевых диодов при комнатной температуре доминирует тепловой ток, а ток термогенерации начинает играть существенную роль лишь при отрицательной температуре.

Однако в этом диапазоне величина обратного тока делается вообще малозначительной.

Ток утечки. Поверхностные утечки представляют собой нередко главный фактор, влияющий на обратную характеристику. Ток утечки не всегда является результатом «загрязнения» поверхности. Он обусловлен в первую очередь поверхностными энергетическими уровнями, которые способствуют активной генерации — рекомбинации, а также молекулярными или ионными пленками, шунтирующими переход (это могут быть молекулы окислов основного материала, молекулы газов, воды, ионы водорода и т.п.). При повышении напряжения ток утечки растет сначала почти линейно, а затем более сильно (рис. 2.17). Почти линейный начальный участок характеристики можно охарактеризовать эквивалентным сопротивлением утечки R_y .

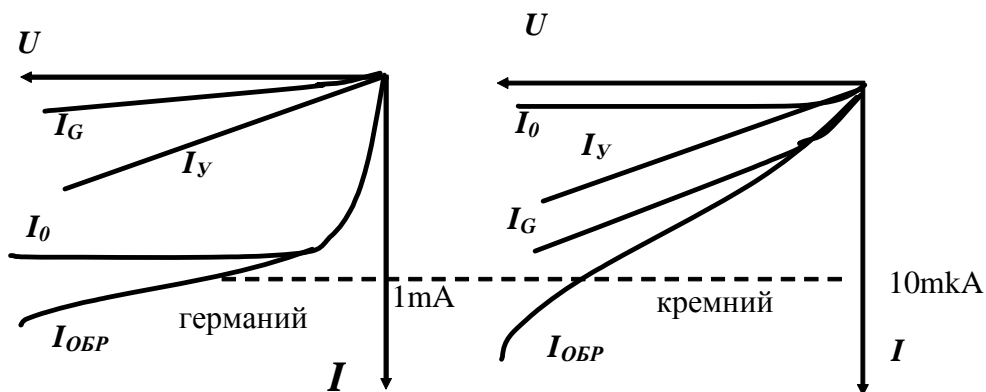


Рис. 2.17 — Обратные характеристики реальных диодов — германиевого (а) и кремниевого (б)

Характерная черта тока утечки заключается в его временной нестабильности, которую часто называют «ползучестью». Ток утечки зависит от температуры сравнительно слабо. Поэтому будем считать его постоянным при изменениях температуры.

Соответствующая формула для такой идеализированной характеристики имеет вид

$$I = I_{обр} + \frac{U}{r_{обр}}. \quad (2.37)$$

Несмотря на приближенность формулы (2.37), она, как и эквивалентная схема на рис. 2.18, б, позволяет производить полезные количественные оценки в широком диапазоне напряжений. Параметры эквивалентной схемы определяются по данным справочников или путем измерений.

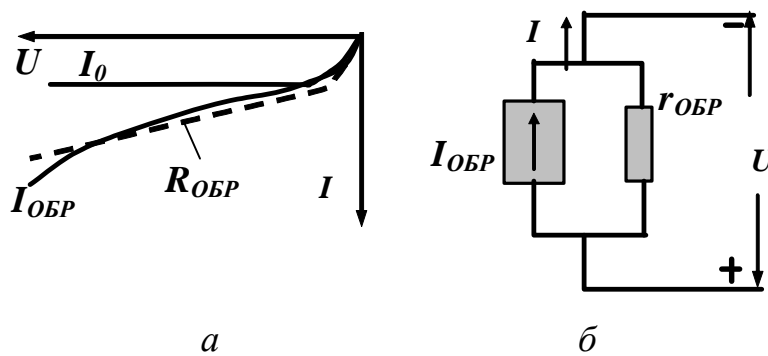


Рис. 2.18 — Обратная характеристика реального диода, ее идеализация (а) и эквивалентная схема диода при обратном включении (б)

Прямая характеристика реального диода

При напряжении $U > \phi_T$ прямая ветвь характеристики согласно (2.23) должна быть экспоненциальной функцией. Между тем, анализ показывает, что реальные характеристики состоят из нескольких участков с разными наклонами, так что формула (2.23) представляет собой лишь некоторое приближение. Рассмотрим главные причины, по которым реальная характеристика особенно при прямом смещении существенно отличается от идеализированной.

Ток рекомбинации. Известно, что в равновесном состоянии токи термогенерации и рекомбинации в переходе взаимно компенсируются. При прямом смещении перехода крутизна потен-

циального барьера уменьшается и носители, не способные преодолеть барьер проникают в переход гораздо глубже. Соответственно увеличивается вероятность их рекомбинации в переходе, что приводит к появлению тока рекомбинации. Рассмотрим участок перехода, в котором концентрации электронов и дырок одинаковы (такой участок всегда имеется внутри перехода). Полагая $p = n$ и $U > 2\phi_T$, получаем для этого участка:

$$p = n = n_i e^{U/2\phi_T} \gg n_i = p_i;$$

$$pn = n_i^2 e^{U/\phi_T} \gg n_0 p_0.$$

Тогда скорость рекомбинации равна

$$V \approx -\frac{n_i}{\tau_\infty} e^{U/\phi_T}. \quad (2.38)$$

Умножив модуль $|V|$ на заряд электрона q объем перехода Sl , мы получим приближенное значение тока рекомбинации. Приближение обусловлено тем, что скорость рекомбинации внутри перехода непостоянная.

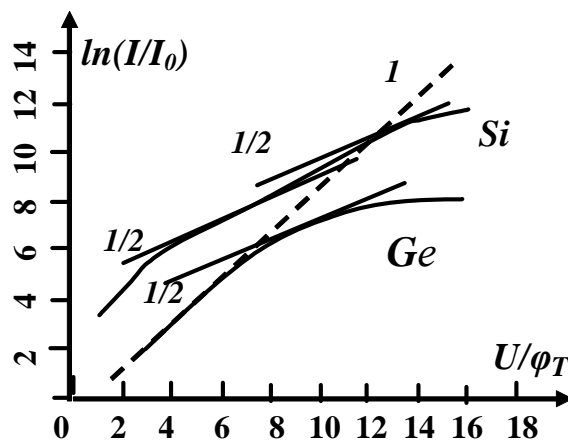


Рис. 2.19 — Характеристики реальных диодов при прямом смещении — германиевого и кремниевых — в полулогарифмическом масштабе

Цифры характеризуют наклон кривых.

Для получения точного выражения тока рекомбинации необходимо взять интеграл от скорости рекомбинации.

Ток рекомбинации можно записать в следующем виде

$$I_R = q(Sl) \frac{\varphi_T}{\Delta\varphi_0 - U} \frac{n_i}{\tau} e^{U/2\varphi_T}. \quad (2.39)$$

Как видим, ток I_R подобно току I_0 пропорционален собственной концентрации, а потому его величина и доля его в общем, прямом токе диода существенно зависят от материала. Ток рекомбинации, так же как и ток термогенерации, играет главную роль в кремниевых диодах. В германиевых диодах его роль может стать заметной при пониженной температуре, когда тепловой ток I_0 сильно уменьшается. Соответственно наклон такой характеристики имеет другую величину. Очевидно, что диффузионный ток сильнее зависит от напряжения, чем ток рекомбинации. Поэтому даже тогда, когда ток I_R играет главную роль при малых напряжениях, с ростом напряжения он неизбежно уступает эту роль диффузионному току. В кремниевых диодах это имеет место при напряжении $U = 0,2 - 0,3$ В.

Сопротивление базы. Будем считать, что в отсутствие инжекции или при малом ее уровне, сопротивление базы определяется обычной формулой:

$$r_{\bar{\sigma}} = \rho_{\bar{\sigma}} \frac{w}{S}, \quad (2.40)$$

где w — толщина базы; S — площадь поперечного сечения. Обычно $r_{\bar{\sigma}}$ лежит в пределах от **1 — 2** до **20 — 30 Ом**. Падение напряжения на базе составляет:

$$U_{\bar{\sigma}} = I r_{\bar{\sigma}} = j \rho_{\bar{\sigma}} w. \quad (2.41)$$

Это напряжение является той поправкой, которую, вообще говоря, следует ввести в формуле (2.23), чтобы учесть различие между падением напряжения на переходе и величиной приложенного напряжения.

$$I = I_0 \left(e^{\frac{U_{\bar{\sigma}} - I r_{\bar{\sigma}}}{\varphi_T}} - 1 \right); \quad (2.42a)$$

$$U_{\bar{\sigma}} = \varphi_T \ln \left(\frac{I}{I_0} + 1 \right) + I r_{\bar{\sigma}}. \quad (2.42b)$$

С увеличением тока напряжение U_B растет линейно, а напряжение U — логарифмически, т.е. более слабо. Поэтому при

достаточно большом токе всегда превалирует напряжение U_B и экспоненциальная характеристика диода вырождается.

Следует указать, что дифференциальное и статическое сопротивления диодов могут существенно отличаться от аналогичных сопротивлений, вычисленных для перехода, из-за влияния сопротивления базы.

2.6 Переходные характеристики плоскостного диода

Наличие реактивных компонентов (паразитные емкости и индуктивности) приводит к появлению переходных процессов в устройствах, что ограничивает их быстродействие и уменьшает полосу пропускания.

Полупроводниковый диод также является инерционным элементом по отношению к быстрым изменениям тока или напряжения, поскольку новое распределение носителей устанавливается не сразу. Индуктивностями выводов пренебрегаем. Как известно, внешнее напряжение приводит к изменению ширины перехода, а значит, и величины пространственных зарядов в переходе. Кроме того, при инжекции (или экстракции) меняются заряды в области базы. Следовательно, наряду с проводимостью, которая в первом приближении характеризуется выражением (2.23), диод обладает емкостью, которую можно считать подключенной параллельно $p-n$ переходу. Эту емкость принято разделять на две составляющие: барьерную емкость, отражающую перераспределение зарядов в переходе, и диффузионную емкость, отражающую перераспределение зарядов в базе. Понятие диффузионная емкость является вообще-то условным, т.к. физически этой емкости нет. Введение понятия «диффузионная емкость» позволяет проводить аналитические расчеты, не прибегая к расчету распределения носителей в базе. Такое разделение, в общем, весьма условно, но удобно на практике, тем более что соотношение обеих емкостей различно при разных полярностях смещения.

При прямом смещении главную роль играют заряды в базе и соответственно диффузионная емкость. При обратном смещении (режим экстракции) заряды в базе меняются мало, и главную

роль играет барьерная емкость. Характеристика инерционных свойств диода с помощью емкостей особенно удобна в случае малых переменных сигналов, действующих на фоне больших постоянных смещений. При этом емкости оказываются почти линейными и весьма наглядно дополняют эквивалентную схему диода. В случае больших сигналов использование емкостей, особенно диффузионной, становится нецелесообразным, так как их нелинейность проявляется в такой мере, что эквивалентная схема теряет свою наглядность, а анализ делается отнюдь не более простым, чем при использовании уравнений непрерывности. После этих предварительных замечаний мы рассмотрим сначала свойства барьерной и диффузионной емкостей, а затем исследуем переходный процесс при большом сигнале.

Барьерная емкость (емкость перехода). Определим величину барьерной емкости, считая, что переход несимметричен и сосредоточен в базе типа. Заряды на пластинах воображаемого конденсатора — перехода — одинаковы ($-Q_p = Q_n$). Найдем один из них, например Q_n , для случая ступенчатого перехода

$$Q_n = qN_d (Sl),$$

где S — площадь перехода; l — ширина перехода. Подставляя сюда выражение (2.12) и дифференцируя (Q_n) по $|U|$, получаем дифференциальную барьерную емкость ступенчатого перехода при обратном смещении

$$C_{II} = \frac{S\sqrt{0,5\xi\xi_0qN_d}}{\sqrt{|U|}} = \frac{\xi\xi_0S}{l_0} \sqrt{\frac{\Delta\phi_0}{|U|}}. \quad (2.43a)$$

Вторая форма выражения удобна тем, что ее первый множитель является емкостью обычного плоского конденсатора с расстоянием между пластинами, равным l_0 — равновесной ширине перехода, определяемой по формуле (2.9б). Как видим, емкость перехода зависит от приложенного напряжения. Величина барьерной емкости современных диодов лежит в пределах (2–200) пф. В общем случае, используя (2.11), для барьерной емкости будем иметь:

$$C_{II} = \frac{\xi\xi_0 S}{l_0} \sqrt{\frac{\Delta\varphi_0}{\Delta\varphi_0 - U}}. \quad (2.436)$$

Эта формула может давать большую погрешность при прямых напряжениях (**более 0,1 — 0,2 В**), так как в исходных выражениях (2.9) не учтен заряд подвижных носителей в переходе, существенный при прямых смещениях. **Необходимо заметить, что изменение ширины перехода (т.е. перезаряд емкости) требует притока или отвода основных (для каждого слоя) носителей. Отсюда следует, что на высоких частотах должен уменьшаться коэффициент инжекции неосновных носителей.**

$$C_{II} = \frac{\xi\xi_0 S}{l_0} \sqrt[3]{\frac{\Delta\varphi_0}{|U|}}. \quad (2.43в)$$

Зависимость барьерных емкостей от обратного напряжения показана на рис. 2.20. Емкость перехода при прочих равных условиях зависит от концентрации примесей, или от удельного сопротивления материала. Чем больше удельное сопротивление, тем меньше емкость

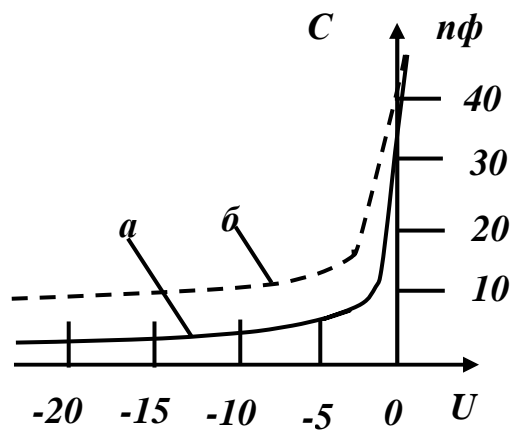


Рис 2.20 — Зависимость емкости ступенчатого (а) и плавного (б) переходов от обратного напряжения

Диффузионная емкость. При прямом смещении диода переход, как известно, сужается и соответственно растет барьерная емкость. Однако эта емкость оказывается менее существенной, чем емкость, обусловленная возрастающим зарядом носителей в базовом слое, которую называют диффузионной, так как этот за-

ряд лежит в основе диффузии носителей в базе. Диффузионная емкость C_D «заряжается» как инжектированными дырками, так и электронами, компенсирующими заряд инжектированных дырок, поскольку избыточные заряды электронов и дырок одинаковы. Найдем один из них, а именно заряд дырок, исходя из распределения (2.20):

$$\Delta Q = qS \int_0^w \Delta p(x) dx = \frac{qSL\Delta p(0)}{\operatorname{th}\left(\frac{w}{L}\right)} \left(1 - \operatorname{sech} w/L\right).$$

Подставляя сюда $\Delta p(0)$ из (2.20), ток I_0 из (2.26а), получаем:

$$\Delta Q = I\tau \left(1 - \operatorname{sech} w/L\right). \quad (2.44)$$

Если поделить этот заряд на напряжение U , то интегральная диффузионная емкость запишется в следующем виде:

$$C_D = \frac{\Delta Q}{U} = \frac{\tau}{R_D} \left(1 - \operatorname{sech} w/L\right),$$

где $R_D = U/I$ — сопротивление диода постоянному току. Дифференциальная диффузионная емкость будет иметь аналогичную форму:

$$C_D = \frac{d(\Delta Q)}{dU} = \frac{\tau}{r_D} \left(1 - \operatorname{sech} w/L\right), \quad (2.45)$$

где $r_D = \partial U / \partial I$ — сопротивление диода переменному току. Как видим, диффузионная емкость (2.45) является функцией прямого тока, подобно тому, как барьерная емкость (2.43а) является функцией обратного напряжения. Кроме того, диффузионная емкость находится в прямой зависимости от толщины базы, уменьшаясь с уменьшением отношения w/L . В случае толстой базы получаем:

$$\Delta Q \approx I\tau; \quad (2.46а)$$

$$C_D \approx \frac{\tau}{r_D} = \frac{I\tau}{\Phi_T}. \quad (2.46б)$$

Например, если $\tau = 5$ мксек, $I = 10$ мА, то $C_D = 2$ мкФ. Такие значения на несколько порядков превосходят величину барьерной емкости.

В случае тонкой базы C_D приводятся к виду:

$$\Delta Q \approx It_D; \quad (2.47a)$$

$$C_D \approx \frac{t_D}{r_D}, \quad (2.47b)$$

где
$$t_D = \frac{w^2}{2D} \quad (2.48)$$

есть среднее время диффузии или среднее время пролета носителей через тонкую базу при чисто диффузионном механизме движения.

При воздействии импульсных сигналов протекают переходные процессы, которые возникают из-за действия зарядной ёмкости перехода и процессов накопления или рассасывания носителей заряда в базовой области.

Удобно рассматривать переходные процессы в диоде в отдельности при отпирании и запирании перехода.

Отпирание p - n перехода. Импульсное сопротивление.

Время установления прямого сопротивления

Рассмотрим переходный процесс в p - n переходе при его отпирании скачком тока. Такой режим можно обеспечить, выбрав сопротивление R в схеме (рис. 2.21) на порядок и более превышающим дифференциальное сопротивление перехода, т.е. обеспечить режим генератора тока.

$$I = \frac{U_{\Gamma} - U_D}{R} \approx \frac{U_{\Gamma}}{R}.$$

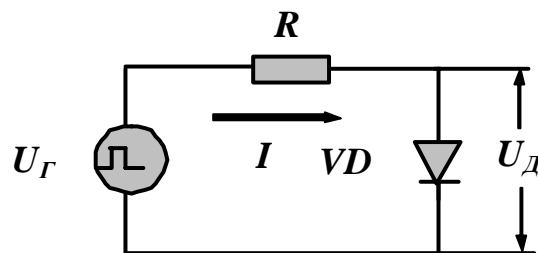


Рис. 2.21 — Схема включения источника импульсного напряжения U_{Γ} при отпирании p - n перехода

Эпюры тока и напряжений показаны на рис. 2.22. Как известно, напряжение на электронно-дырочном переходе U_D складывается из напряжения на объёмном сопротивлении базы U_{OB} и напряжения на переходном слое U_{II} , т.е.

$$U_D = U_{OB} + U_{II}.$$

При подаче ступеньки тока напряжение U_{OB} , скачком увеличиваясь, достигает величины $U_{OB} \approx Ir_B$.

Напряжение же на переходе

$$U_{II} = \varphi_T \ln \frac{Q}{Q_0}$$

возрастает по мере накопления неосновных носителей в базе.

Время установления $\tau_{уст}$ определяется в основном средним временем пролета носителей через область базы (тонкая база).

Запирание p - n перехода. Рассасывание носителей заряда. Время восстановления обратного сопротивления

Рассмотрим переходные процессы, которые возникают при запирании электронно-дырочного перехода.

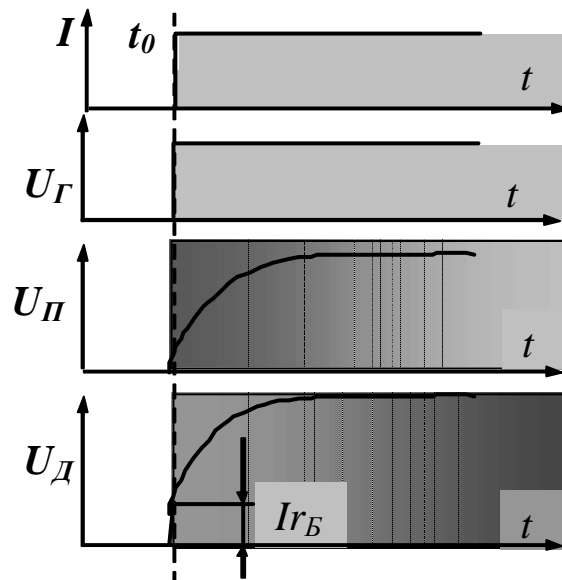


Рис. 2.22 — Эпюры тока и напряжений при отпирании p - n перехода

В исходном состоянии переход открыт напряжением E_{CM} (рис. 2.23), и через него протекает прямой ток $I_D = I_{PP}$.

В момент времени t_0 включается перепад напряжения U_G (рис. 2.23). Амплитуда этого напряжения выбрана так, чтобы полностью запереть переход. При этом $U_D \approx -(U_G - E_{CM})$ и $I_D \approx -I_{OBR}$.

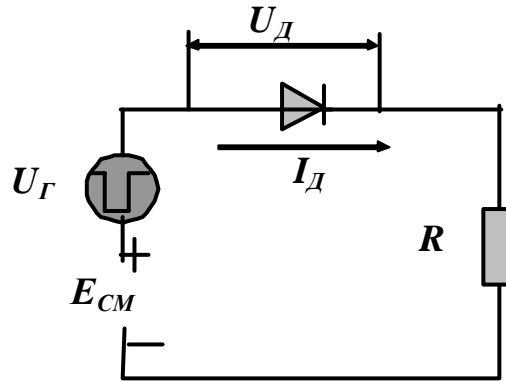


Рис. 2.23 — Схема включения источника импульсного напряжения U_G при запираии $p-n$ перехода

Если исходить из вольт-амперной характеристики, то можно было ожидать следующее: в момент времени t_0 с включением запирающего перепада ток перехода должен был уменьшиться до величины I_{OBR} и далее оставаться постоянным.

На рис. 2.24 эта эпюра показана штриховыми линиями. В момент времени t_0 ток скачком уменьшается до величины $I_{OBR.ИМП}$, в десятки и сотни раз превышающей величину тока I_{OBR} .

В течение некоторого времени этот всплеск тока практически не изменяется (рис. 2.24). Затем начинается спад обратного тока до своего установившегося значения I_{OBR} .

Время, в течение которого происходит рассасывание избыточного заряда, пока переход включен в прямом направлении, называется временем рассасывания $t_{рас}$.

Время, в течение которого ток через диод спадает от $I_{OBR.ИМП}$ до уровня $1,1I_{OBR}$, называется длительностью среза t_{CP} . Время восстановления обратного сопротивления можно считать равным сумме времен среза и рассасывания.

С ростом прямого тока высота потенциального барьера

уменьшается, стремясь в пределе к нулю. Следовательно, при больших токах наличие перехода делается все менее существенным, и диод постепенно превращается в двухслойную полупроводниковую пластинку, в которой главную роль играет слой базы. Оценим величину тока, при котором наступает вырождение экспоненциальной характеристики.

Для этого проще всего положить дифференциальное сопротивление r_{δ} равным сопротивлению r_{σ} . Используя (2.28), приходим к соотношению:

$$I_B = \Phi_T / r_{\sigma}. \quad (2.49)$$

Омический участок может составлять значительную, а иногда и основную рабочую часть характеристики рис. 2.24.

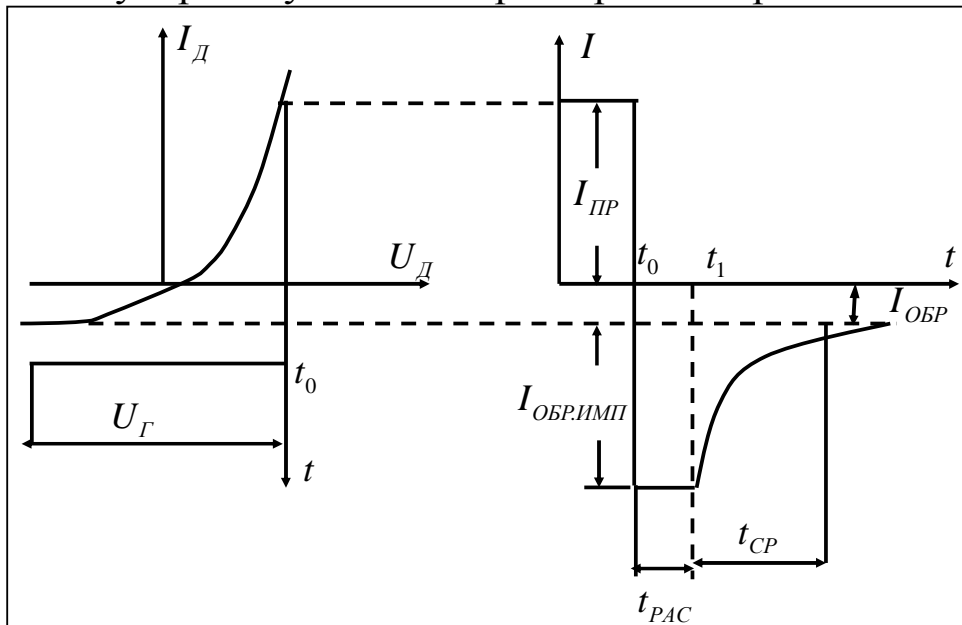


Рис. 2.24 — Вольт-амперная характеристика и временные диаграммы тока, протекающего через p - n переход при его запирации

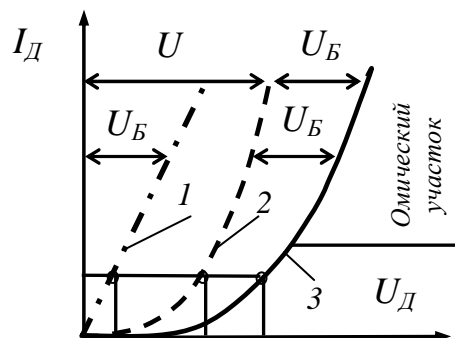


Рис. 2.25 — Влияние объемного сопротивления базы на прямую характеристику реального диода

Зависимость прямого напряжения от температуры. Рассмотрим сначала идеализированный диод, пользуясь формулой (2.27). Пренебрегая единицей в формуле (2.27), подставляя I_0 из выражения (2.31а) и логарифмируя, получаем:

$$U(T) = \varphi_T \ln \frac{I}{I_{00}} + \varphi_3.$$

Учитывая, что I_{00} и φ_3 слабо зависят от температуры, функция $U(T)$ близка к линейной. Полагая, что $I_{00} = \text{const}$ и $\ln(I/I_{00}) = \text{const}$ и дифференцируя U по T , получим температурную чувствительность напряжения перехода

$$\zeta_{II} = \frac{dU}{dT} = \frac{U - \varphi_3}{T} < 0.$$

Отрицательный знак температурной чувствительности обусловлен очевидным неравенством $U < \Delta\varphi_0 < \varphi_3$. Это неравенство вытекает из условия, что для изготовления диодов используются невырожденные полупроводники. В первой главе мы доказали, что в невырожденных полупроводниках должно выполняться условие $\Delta\varphi_0 < \varphi_3$. С увеличением прямого напряжения, а следовательно, и прямого тока модуль температурной чувствительности уменьшается. Для реальных диодов температурная чувствительность с увеличением тока уменьшается значительно быстрее, чем следует из формулы, и при некотором токе температурная чувствительность становится равной нулю, а затем меняет знак. Это связано с тем, что при больших токах падение напряжения на объёмном сопротивлении базы r_b растёт и с увеличением температуры возрастает сопротивление базы. Использовать точку с нулевой чувствительностью для обеспечения высокой температурной стабильности на практике не представляется возможным из-за большой величины тока через диод. При разработке практических схем с применением диодов, как правило, задают ток в цепи диода, запитывая его через большое сопротивление. В этом случае ток через диод изменяться не будет, т.е. обеспечивается режим стабилизации тока. Напряжение на диоде будет изменяться с изменением температуры.

Вопросы для самопроверки

1. Какая поверхность называется металлургической границей?
2. Что называется областями объёмного заряда?
3. Запишите условие образования резкого перехода?
4. Нарисуйте диаграмму распределения носителей в несимметричном переходе.
5. Нарисуйте зонную диаграмму p - n перехода в равновесном состоянии.
6. Объясните, почему дырки n -слоя, а электроны p -слоя могут свободно переходить соответственно в p -слой и n -слой?
7. Объясните понятие диффузионного потенциала и принцип его образования?
8. Объясните, почему при равновесии ток через переход равен нулю?
9. Почему в равновесной системе $\varphi_{Fn} = \varphi_{Fp}$, в то время как до объединения полупроводников $\varphi_{Ep} \neq \varphi_{En}$.
10. От каких параметров зависит величина диффузионного потенциала и почему?
11. Почему переход в основном сосредоточен в полупроводнике с более низкой концентрацией примесей?
12. Почему ширина перехода с увеличением концентрации носителей уменьшается?
13. Как изменяется ширина перехода в зависимости от полярности приложенного внешнего напряжения?
14. Изменяется ли распределение носителей в переходе при прямом и обратном смещении на переходе?
15. Какие типы переходов, кроме резкого перехода, Вы знаете?
16. Принципы образования выпрямляющих контактов.
17. Принципы образования невыпрямляющих контактов.
18. Образуется ли переход в выпрямляющих контактах.
19. Почему при анализе перехода можно пользоваться уравнениями диффузии, а не уравнениями непрерывности?
20. Выведите выражение для распределения дырок в базе от величины внешнего напряжения.
21. Выведите формулу, описывающую статическую вольт-

амперную характеристику диода.

22. Объясните физическую природу теплового (масштабного, насыщения) тока.

23. Что такое характеристические сопротивления диода и чему они равны?

24. Почему при реальном использовании диодов, как правило задают прямой ток, а не прямое напряжение?

25. Запишите выражение, определяющее температурную зависимость теплового тока?

26. Объясните, почему с повышением температуры тепловой ток увеличивается, используя только физические основы полупроводников.

27. В чем причина тока термогенерации и его зависимость от температуры?

28. Причина возникновения токов утечки в переходе.

29. Зависит ли статическая вольт-амперная характеристика диода от температуры при прямом смещении на диоде?

30. Нарисуйте эквивалентную схему диода при прямом смещении.

31. Объясните причину барьерной емкости и её зависимость от величины приложенного напряжения.

32. Физическая причина появления диффузионной ёмкости и её зависимость от величины тока, протекающего через переход.

33. Переходные процессы в диоде при резких изменениях входных напряжений.

3. РАЗНОВИДНОСТИ ДИОДОВ

Теория и свойства плоскостных полупроводниковых диодов, изложенные в предыдущей главе, лежат в основе всех других типов диодов, количество которых в настоящее время довольно велико. Специфика каждого из этих специальных диодов требует особого анализа, но мы ограничимся их качественной характеристикой в той мере, в какой это полезно при разработке схем.

3.1 Точечные диоды

Точечные диоды (рис. 3.1) появились намного раньше плоскостных диодов. Однако процессы, происходящие в них, сложнее и до сих пор еще полностью не изучены. Эффект выпрямления на границе между металлической иглой и пластинкой полупроводника имеет место даже при простом их соприкосновении. Однако в настоящее время точечный контакт получают, как правило, путем так называемой формовки. Процесс формовки заключается в пропускании через собранный точечный диод сравнительно мощных, но коротких импульсов тока в прямом или обратном направлении. Количество, величина и длительность формирующих импульсов выбирают на основании опытных данных.

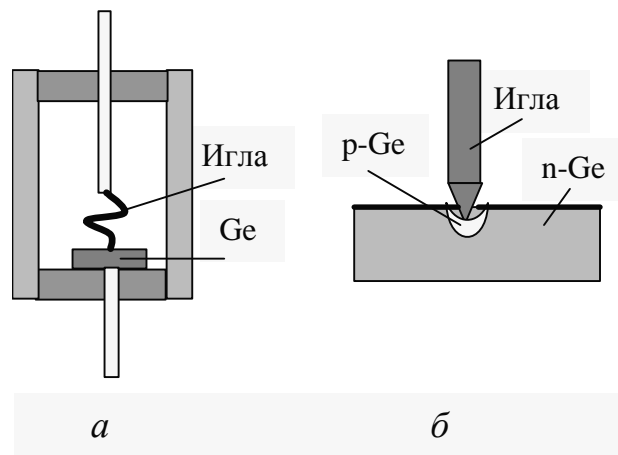


Рис. 3.1 — Конструкция точечного диода (а) и структура его перехода (б)

Общей целью формовки является сильный местный нагрев контакта, при котором происходит своеобразное сплавление кончика иглы с полупроводником. Сплавление обеспечивает ста-

бильность и механическую прочность контакта, что и было первоначальной целью формовки. Однако, как выяснилось позднее, при сплавлении происходит также изменение типа проводимости в тонком слое полупроводника, прилегающем к игле. Это превращение объясняется диффузией примесей из иглы в полупроводник при сильном разогреве и частичном расплавлении обоих элементов в месте контакта.

Например, в случае когда исходная пластинка германия имеет электронную проводимость, а материалом иглы является бериллиевая бронза, может происходить диффузия бериллия в германий. Бериллий, будучи акцептором, по отношению к германию, обуславливает наличие тонкого *p*-слоя в германии вблизи иглы (рис. 3.1). Как видим, в данном случае получается *p-n* переход, правда, своеобразной (не плоской) конфигурации и малой площади. Обычно при анализе форму перехода в точечном диоде принято считать полусферической, что, несомненно, близко к действительности. Не проводя самого анализа, отметим лишь характерные особенности точечных переходов и диодов. Очевидно, что малая площадь перехода обуславливает не только малую емкость перехода, но и малую допустимую мощность. Величина допустимых прямых токов у точечных диодов значительно меньше, чем у плоскостных, а эффективное сопротивление базы больше из-за малой площади эмиттера. Сопротивления базы у точечных диодов составляют десятки и сотни Ом, а прямые токи обычно не больше 10 — 20 мА. Существенное превышение допустимого тока (даже в течение короткого интервала времени) приводит у точечных диодов к «переформовке» контакта и ухудшению или утрате выпрямляющих свойств. Для прямой ветви вольт-амперной характеристики точечного диода выражение (2.23) недействительно, так как даже при очень небольших токах уровень инжекции оказывается весьма высоким из-за малой площади перехода. При больших токах уровень инжекции получается настолько высоким, что экспоненциальная характеристика вырождается и приближается к параболической. Особенно сильно различаются статические характеристики точечных и плоскостных диодов в области обратных токов (рис. 3.2). Поскольку площадь перехода мала, мал и тепловой ток I_0 . Однако участок на-

сыщения небольшой и обычно плохо заметен, так как уже при сравнительно небольших напряжениях (несколько вольт) обратный ток существенно возрастает за счет утечек, а также за счет заметного повышения температуры перехода, теплоотвод затруднен из-за малой площади контакта, и большого теплового сопротивления.

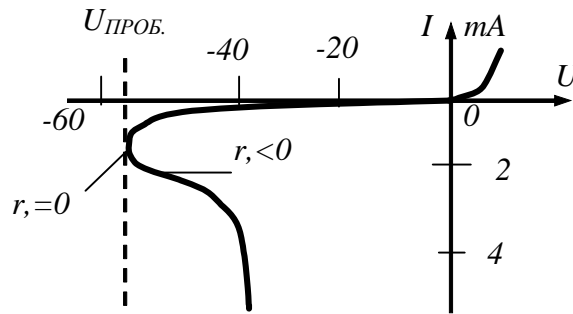


Рис. 3.2 — Статическая характеристика точечного диода

Важной особенностью обратной характеристики является участок с отрицательным дифференциальным сопротивлением, который обусловлен тепловым пробоем.

Несмотря на принципиальную ценность указанного участка, его практическое использование нецелесообразно вследствие большого разброса величины отрицательного сопротивления и координат начала и конца участка, плохой стабильности во времени и малого срока службы диода в таком перенапряженном режиме. Переходные процессы протекают в точечных диодах качественно так же, как и в плоскостных. Количественные различия связаны с меньшей площадью перехода и временем жизни. Время жизни у точечных диодов меньше, чем у плоскостных, так как роль поверхностной рекомбинации возрастает с уменьшением площади перехода. Точечные диоды широко используются в высокочастотных схемах и импульсной технике. Обширный и своеобразный класс точечных полупроводниковых диодов составляют германиевые и кремниевые детекторные диоды, применяющиеся главным образом в технике СВЧ, что обусловлено очень малым временем жизни носителей, а вместе с малой площадью контакта обеспечиваются хорошие частотные свойства, необходимые для работы в области СВЧ.

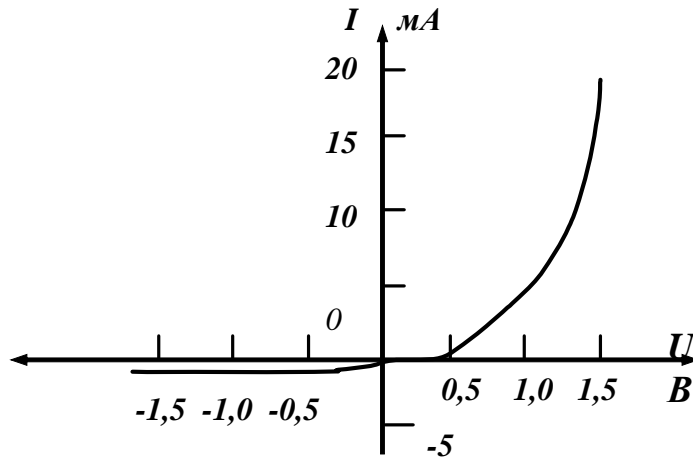


Рис. 3.3 — Статическая характеристика детекторного диода

Для таких диодов, помимо рабочей частоты, важными параметрами являются такие, как коэффициент шума, коэффициент преобразования и др. Эти параметры могут быть достаточно хорошими, несмотря на «плохую» (с точки зрения «обычных» диодов и электронных схем) вольт-амперную характеристику. Конструкция детекторных диодов обычно приспособлена к сочленению с элементами волноводов, измерительных головок и других деталей систем СВЧ (рис. 3.4). Применение такой конструкции позволяет существенно уменьшить паразитные емкости диодов и тем самым увеличить их быстродействие.

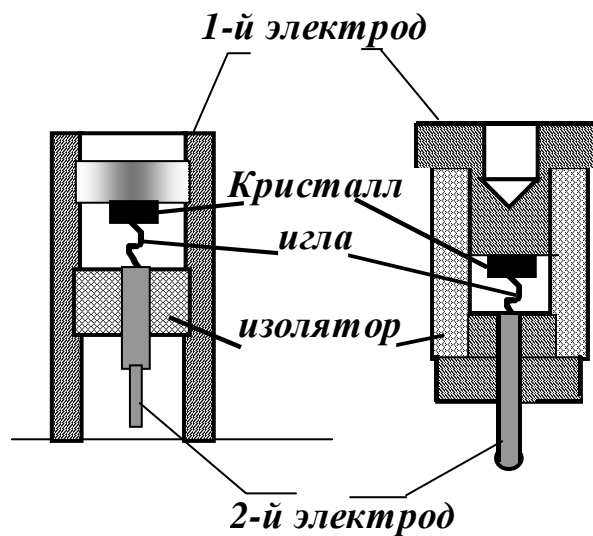


Рис. 3.4 — Конструктивные варианты детекторных диодов

В настоящее время более перспективными диодами для области СВЧ являются диоды Шоттки, которые имеют высокое быстродействие, работают при больших прямых токах и больших прямых напряжениях.

3.2 Полупроводниковые стабилитроны

Одним из основных условий нормального функционирования любой электронной системы особенно с применением интегральных микросхем является высокая стабильность питающих напряжений.

Основным элементом практически любого стабилизатора напряжения является стабилитрон — кремниевый плоскостной диод, работающий в области восстановления пробоя.

Такой «полупроводниковый стабилитрон» по большинству параметров превосходит газоразрядный, и, главное, он может иметь значительно меньшие рабочие напряжения стабилизации, что совершенно необходимо в транзисторных схемах, особенно в интегральных схемах, которые практически всегда низковольтные. Выбор кремния в качестве материала для полупроводниковых стабилитронов обусловлен главным образом малым обратным током, следовательно, невысокой мощностью рассеивания в предпробойной области. При этом саморазогрев диода в предпробойной области отсутствует, и переход в область пробоя получается достаточно резким (рис. 3.5). Кроме того, в самой области пробоя, даже при большом токе, нагрев диода не носит лавинообразного характера, т.к. в этом случае дифференциальное сопротивление диода резко уменьшается и рассеиваемая мощность не превышает допустимого значения. Из рисунка 3.5 видно, что статическая вольт-амперная характеристика стабилитрона близка к симметричной и отличается только величиной напряжения стабилизации $U_{СТ}$.

При токах $I_D > 5$ мА дифференциальные сопротивления также равны $r_{диф.} \approx r_B$. Поэтому при анализе работы стабилитрона как элемента стабилизатора не будем делать различий, для какого включения проводится анализ.

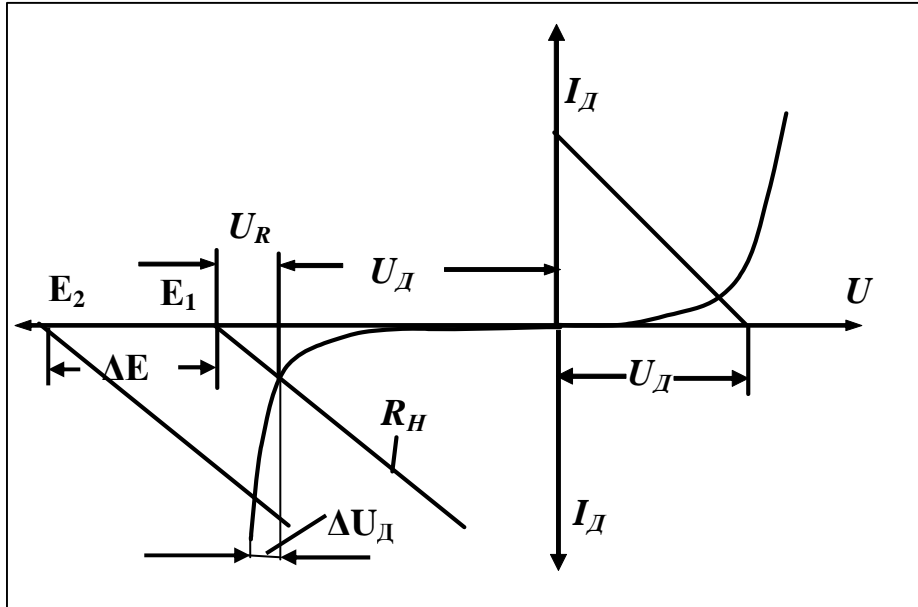


Рис. 3.5 — Статическая характеристика полупроводникового стабилизатора

На рис. 3.6 приведен температурный коэффициент нестабильности стабилизатора в режиме стабилизации при обратном включении. Для прямого включения эта зависимость рассмотрена раньше.

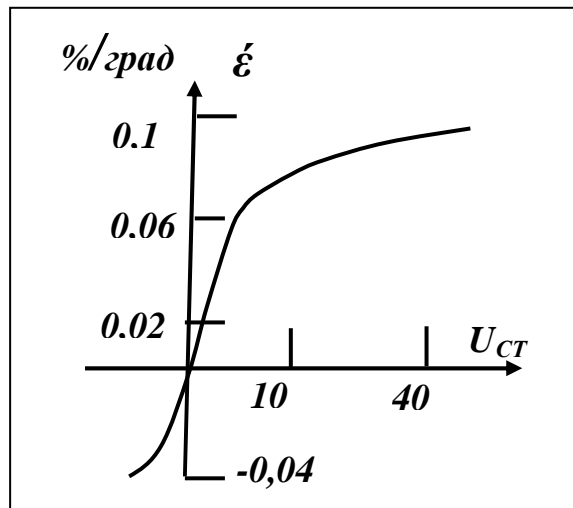


Рис. 3.6 — Зависимость ТКН на рабочем участке кремниевого стабилизатора от напряжения пробоя

У низковольтных стабилизаторов пробой носит туннельный эффект, т.к. база выполнена из низкоомного полупроводника, а в

этом случае напряжение пробоя прямо пропорционально ширине запрещенной зоны, которое с повышением температуры падает.

В высоковольтных стабилитронах пробой носит лавинный характер, а с повышением температуры подвижность носителей падает и растет напряжение пробоя.

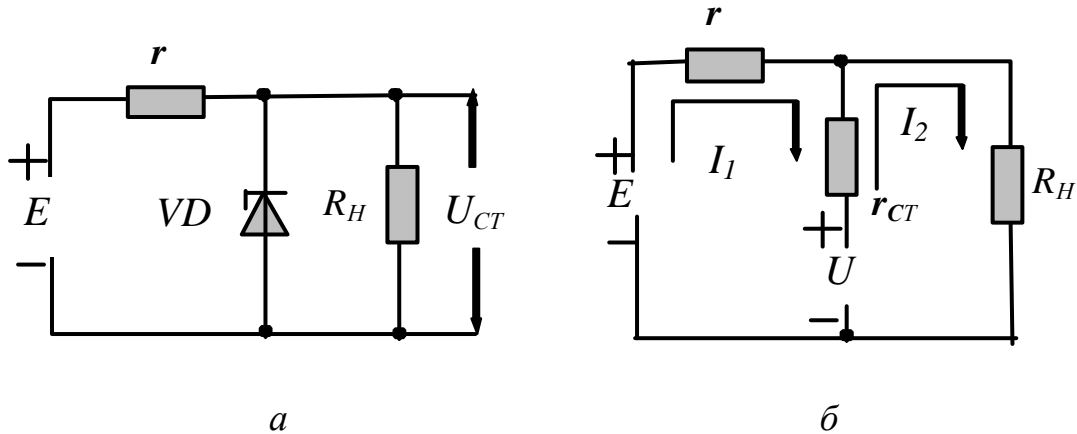


Рис. 3.7 — Принципиальная (а) и эквивалентная (б) схемы стабилизатора напряжения. Здесь U_{CT} — напряжение стабилизации, E — идеальный генератор напряжения, $r_{CT} = r_{\delta}$ — дифференциальное сопротивление стабилитрона

Анализируя эквивалентную схему рис. 3.7, б, уравнения контурных токов запишем в виде:

$$E - U = I_1(r + r_{CT}) + I_2 r_{CT}; \quad (3.1)$$

$$-U = I_1 r_{CT} + I_2(r_{CT} + R_H). \quad (3.2)$$

Решив систему уравнений (3.1) и (3.2), получим:

$$U_{CT} = I_2 R_H = U_{R_H} = \frac{R_H(Er_{\delta} + Ur)}{R_H(r + r_{\delta}) + rr_{\delta}}. \quad (3.3)$$

Для нахождения неустойчивости U_{CT} при изменении напряжения первичного источника питания E продифференцируем (3.3)

$$\frac{dU_{CT}}{dE} = \frac{R_H r_{\delta}}{R_H(r + r_{\delta}) + rr_{\delta}}. \quad (3.4)$$

Перейдя к конечным приращениям (3.4), запишем в виде:

$$\Delta U_{CT} = \Delta E \frac{R_{\text{ЭКВ}}}{r + R_{\text{ЭКВ}}}, \quad (3.5)$$

где $R_{ЭКВ} = \frac{R_H r_B}{R_H + r_B}$.

Для ослабления нестабильности $\Delta U_{СТ}$ при изменениях напряжения первичного источника питания необходимо выполнить условие:

$$r \gg R_{ЭКВ}. \quad (3.6)$$

Неравенство (3.6) говорит о том, что питание выходной цепи необходимо обеспечивать от генератора тока.

При работе радиоэлектронного устройства ток нагрузки изменяется, т.е. $R_H \neq const$, а следовательно, изменяется и $U_{СТ}$. Проанализируем нестабильность $U_{СТ}$ при изменениях величины нагрузки, при этом считаем $r_{СТ} = r_B = const$.

Продифференцируем (3.3) по R_H . Тогда:

$$\Delta U = E \frac{\Delta R_H r_B^2}{r (R_H + r_B)^2}. \quad (3.7)$$

Для уменьшения влияния изменений нагрузки необходимо выполнить условие

$$R_H \gg r_B; \quad (3.8)$$

$$\Delta U_{СТ} = E \delta_H \frac{r_B^2}{R_H r}, \quad (3.9)$$

где $\delta_H = \frac{\Delta R_H}{R_H}$ — относительная нестабильность нагрузки.

Анализ выражений (3.8) и (3.9) показывает, что для уменьшения влияния нагрузки необходимо, чтобы ток нагрузки был как можно меньше тока, протекающего через стабилитрон.

3.3 Туннельные диоды

Туннельные диоды, разработанные в 1958–1959 гг. японским физиком Есаки, интересны тем, что, будучи двухполюсниками, они могут усиливать сигналы. Это объясняется наличием участка с отрицательным сопротивлением на их вольт-амперной характеристике. В отличие от точечных диодов отрицательное сопротивление у туннельных диодов имеется не на обратной, а на

прямой ветви характеристики.

Рассмотрим происхождение такой формы характеристики, воспользовавшись энергетическими диаграммами p - n перехода. Отличительными особенностями туннельного диода являются очень малые удельные сопротивления p - и n -слоев и соответственно очень малая ширина перехода. В этом случае, как отмечалось в гл. 1, полупроводник вырождается, превращаясь в полуметалл. Уровни примесных атомов сливаются в зоны, а последние в свою очередь сливаются с соответствующими основными зонами слоев. В результате уровни Ферми, как и в металле, располагаются не в запрещенных зонах p - и n -слоев, а в разрешенных зонах: в валентной зоне p -слоя и в зоне проводимости n -слоя. При этом энергетическая диаграмма **симметричного перехода** в равновесном состоянии будет примерно такой, как показано на рис. 3.9, *а*. Как видим, нижняя часть зоны проводимости в слое n и верхняя часть валентной зоны в слое p оказались разделенными весьма узким запирающим слоем. Если ширина его не превышает 0,01 — 0,02 мк, то носители имеют возможность переходить в смежный слой «по горизонтали», т.е. не преодолевая потенциального барьера. Это явление обусловлено туннельным эффектом, откуда и происходит название диодов. Ниже мы рассмотрим работу диода, анализируя движение электронов. Поведение дырок совершенно аналогично, а относительная роль обоих типов носителей, как обычно, зависит от степени симметрии диода, т.е. от соотношения удельных сопротивлений слоев. Распределение электронов по энергиям отражено на рис. 3.9 разными расстояниями между кружками, обозначающими электроны. Стрелками снабжены те электроны, которые способны перейти в смежный слой тем или иным путем. Результирующий ток через переход оценивается на рис. 3.9 как разность электронных потоков, проходящих из одного слоя в другой. В равновесном состоянии потоки электронов в обоих направлениях уравниваются, и ток отсутствует (рис. 3.9, *а*). Приложим к диоду внешнее напряжение обратной полярности (т.е. плюсом к n -слою). Энергетическая диаграмма для этого случая показана на рис. 3.9, *б*. Так как количество электронов с энергией, превышающей уровень Ферми, невелико, то поток электронов из p -слоя в n -слой увеличится, а обратный поток останется почти неизменным. Следовательно, ре-

зультирующий ток будет протекать в направлении от n -слоя к p -слою. Этот ток очень быстро возрастает с увеличением обратного напряжения (см. рис. 3.8), поскольку плотность электронов в глубине валентной зоны огромна и малейшее приращение разности $\varphi_{Fp} - \varphi_{Fn}$ сопровождается существенным изменением потока электронов из p -слоя в n -слой. Теперь приложим к диоду небольшое прямое напряжение. Энергетическая диаграмма для этого случая показана на рис. 3.9, в. Легко заметить, что поток электронов из p -слоя в n -слой сильно убывает, а обратный поток меняется сравнительно слабо. Следовательно, результирующий ток протекает в направлении от p -слоя к n -слою и при небольших прямых напряжениях возрастает с увеличением напряжения (рис. 3.8). На рис. 3.9, г изображены зонные диаграммы перехода при напряжении между точками 1 — 2 статической вольтамперной характеристики (рис. 3.8, а). На рис. 3.9, д зонные диаграммы соответствуют напряжениям между точками 2 — 3.

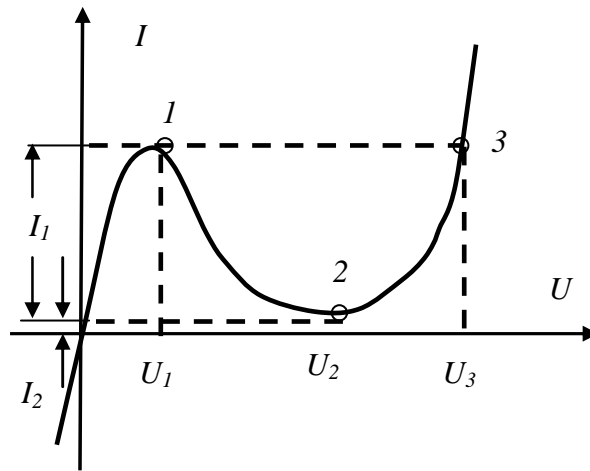


Рис. 3.8 — Статическая характеристика туннельного диода

Туннельный диод представляет собой универсальный прибор, способный выполнять все функции, свойственные активным элементам электронных схем. Поэтому вопросы применения туннельных диодов составляют отдельную область прикладной электроники. Заметим лишь, что диодные схемы настолько сильно отличаются от триодных, что требуют пересмотра многих привычных методов построения и расчета электронных схем.

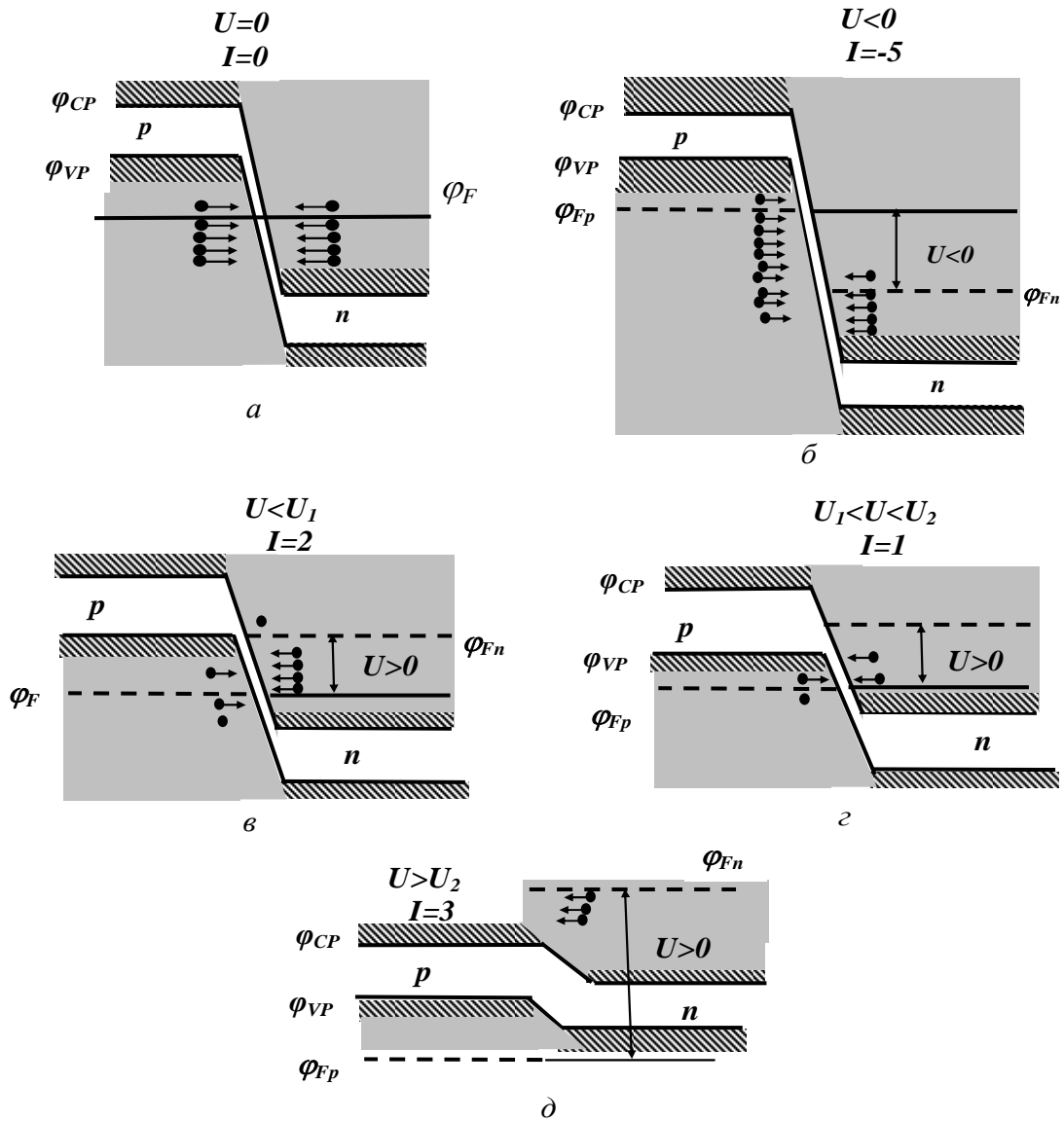


Рис. 3.9 — Энергетические диаграммы туннельного диода на разных участках характеристики:
a — равновесное состояние; *б* — обратное включение,
в — прямое включение; *г* — $U_1 < U < U_2$; *д* — $U > U_2$.

Весьма интересным вариантом туннельного диода является так называемый обращенный диод, характеристика которого показана на рис. 3.10.

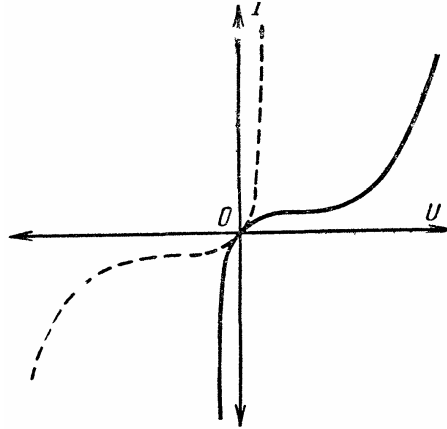


Рис. 3.10 — Статическая характеристика
обращенного диода

Как видим, особенность обращенного диода состоит в том, что отсутствует (или очень мал) максимум на прямой ветви. В этом случае логично повернуть характеристику на 180° (показано пунктиром) и считать прямую ветвь обратной, а обратную — прямой. Таким образом, обращенный диод имеет значительно меньшее прямое напряжение, чем обычные плоскостные диоды (у которых оно составляет десятки и сотни милливольт), что очень ценно для многих применений. Однако его обратное напряжение тоже весьма мало (0,3 — 0,5 В), и с этим нужно считаться при расчете схем. Технология туннельных и обращенных диодов почти одинакова, различие состоит главным образом в подборе исходных материалов (у обращенных диодов концентрация примесей меньше, чем у туннельных).

3.4 Диоды Шоттки

В основе диодов Шоттки лежит контакт между металлом и полупроводником. Такой контакт, как отмечалось, при определенных условиях может обладать выпрямительными свойствами. Для этого необходимо, чтобы приповерхностный слой полупроводника в равновесном состоянии был обеднен основными носителями и чтобы сопротивление обедненного слоя было много больше сопротивления остальной части полупроводниковой пластины. Несмотря на то, что теория контакта между металлом и полупроводником была развита более 60 лет назад, реализовать данный тип

диодов удалось лишь в 70-е годы. На пути реализации такого, казалось бы, простого прибора стояли следующие трудности: во-первых, нужно было осуществить контакт металл — полупроводник без каких бы то ни было промежуточных слоев, (прижимной контакт не обеспечивал этого условия); во-вторых, нужно было добиться малого сопротивления полупроводниковой пластины, не жертвуя при этом ее удельным сопротивлением. Основным преимуществом диодов Шоттки по сравнению с диодами с $p-n$ переходом является тот факт, что у них отсутствует явление инжекции неосновных носителей при прямом смещении, а значит, и явления накопления и рассасывания этих носителей.

Соответственно инерционность диодов Шоттки обусловлена только барьерной емкостью контакта и может быть сделана весьма малой путем уменьшения размеров структуры. Типичный диапазон рабочих частот составляет 3 — 15 ГГц, а времена переключения доходят до 0,1 нс. Еще одним преимуществом диодов Шоттки является то, что экспоненциальный характер статической вольт-амперной характеристика сохраняется для них в гораздо более широком диапазоне токов, чем для обычного $p-n$ перехода, поскольку отсутствует модуляция сопротивления базы неосновными носителями. Обратный ток диодов Шоттки выражается формулой:

$$I_0 = \frac{1}{4} q n_s v_T S.$$

Обратные токи могут составлять всего несколько пикоампер. Обратные напряжения лежат в пределах от 10 до 1000 В. Величины прямых токов зависят, конечно, от площади структуры и качества теплоотвода. В настоящее время разработаны диоды Шоттки на прямые токи порядка 50А и более. При этом прямое падение напряжения составляет всего 0,5 В, т.е. величину, примерно вдвое меньшую, чем у кремниевых диодов. Основным полупроводником — материалом, используемым в диодах Шоттки, является кремний. В качестве металлов используются молибден, нихром, золото, а также алюминий.

3.5 Фотоприёмники (приёмники оптического излучения)

В живой природе оптическая связь распространена весьма широко; ее функции чрезвычайно разнообразны. Человек почти девяносто процентов информации получает с помощью зрения.

Основным элементом устройств является преобразователь оптических сигналов в электрические (фотоприемники).

Из всех известных полупроводниковых фотоприёмников фоторезистор является наиболее простым, но и наиболее универсальным датчиком.

По отношению к сигнальной цепи фоторезистор представляет собой двухполюсник, переменным параметром которого является его импеданс Z_ϕ .

При этом $Z_\phi = f(B, \lambda, \omega_b, t, T)$, откуда следуют основные характеристики фоторезистора: световая, спектральная, частотная, переходная, температурная.

B — световой поток, λ — длина волны оптического излучения, ω_b — частота, t — время, T — температура.

На рис. 3.11 показаны две возможные конструкции фоторезисторов: поперечная (*а*), продольная (*б*).

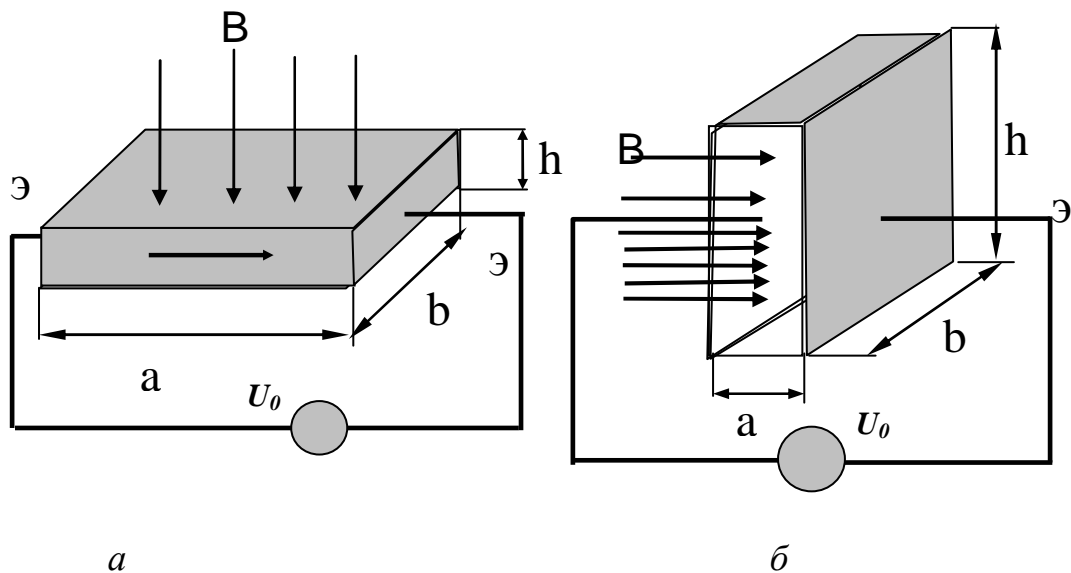


Рис. 3.11 — Принципиальные конструкции фоторезисторов: *а* — поперечная; *б* — продольная; Э — электроды

В первом случае прикладываемое к фоторезистору электрическое поле и возбуждающий свет действуют во взаимно перпендикулярных плоскостях, во втором — в одной плоскости.

Свет, поглощаясь в полупроводнике, возбуждает в нём свободные носители зарядов: электроны и дырки, которые изменяют его проводимость. Если к фоторезистору приложено электрическое поле, то с изменением освещённости будет изменяться и ток в цепи, в которую включен фоторезистор.

Для поперечного и продольного фоторезисторов токи можно записать:

$$I_{\Phi} = \frac{qa\tau\mu k_H}{d^2} BU_{\Phi}, \quad (3.10)$$

$$I_{\Phi} = \frac{qa\tau\mu k_H^1}{d} BU_{\Phi}, \quad (3.11)$$

где q — заряд электрона; a — квантовый выход; τ , μ — среднее время жизни и подвижность носителей тока в полупроводнике; d — расстояние между электродами фоторезистора; k_H^1 — коэффициент поглощения в полупроводнике; k_H — безразмерный коэффициент, показывающий долю поглощенного в образце излучения. Диапазон освещенностей фоторезисторов лежит в пределах $(10^{-3} - 10^2)$ лк.

3.6 Фотодиоды

Принцип работы фотодиода с $p-n$ перехода основан на поглощении света в области базы, или в переходе, в результате чего образуются дырки и электроны, что приводит к изменению равновесных концентраций в базе и эмиттере, а следовательно, к появлению фототока.

Основные соотношения, определяющие характеристики $p-n$ перехода как приёмника излучения, можно записать в виде:

$$J_{\Phi} = aqk_H B; \quad (3.12)$$

$$J = J_{\Phi} - J_{OB} = J_{\Phi} - J_0 \left(e^{U/\Phi T} - 1 \right), \quad (3.13)$$

где a — квантовый выход внутреннего фотоэффекта; j_ϕ , j_{OB} — плотности фототока и обратного тока p - n перехода, обусловленные неосновными носителями тока в полупроводнике.

Уравнение (3.12) отвечает семейству (по параметру B) вольт-амперных характеристик фотодиода. В фотодиодном режиме на p - n переход подаётся обратное смещение. При этом семейство вольт-амперных характеристик фотодиода будет выглядеть, как это показано на рис. 3.12, a .

Ордината участков насыщения прямо пропорциональна уровню возбуждения B . При обратном смещении $U < 0$ и $|U| \gg \phi_T$ (3.13) упрощается и принимает вид

$$J = J_\phi + J_0. \quad (3.14)$$

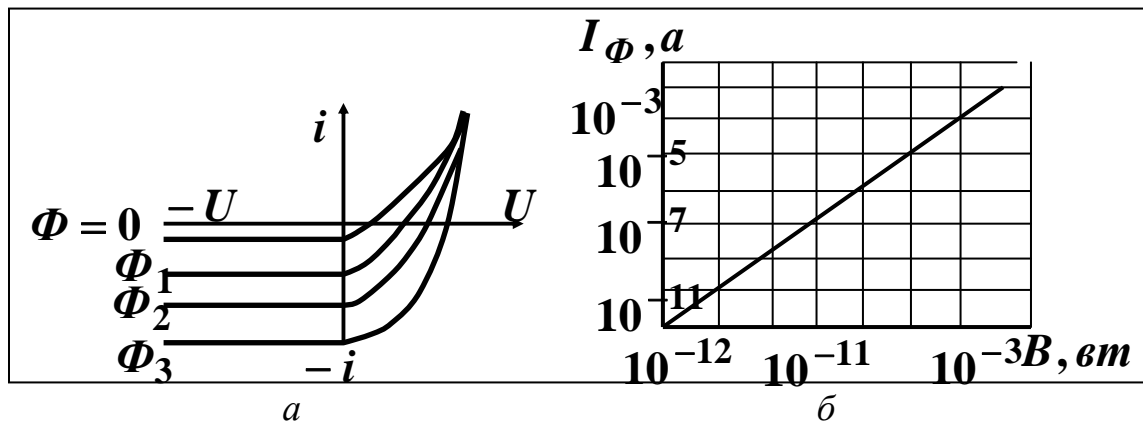


Рис. 3.12 — Характеристики фотодиода:
 a — вольт-амперная $\Phi_1 < \Phi_2 < \Phi_3$;
 b — световая (кремниевый фотодиод)

Фотодиод — быстродействующий прибор, инерционность которого практически не зависит от уровня возбуждения.

Применение фотодиодов на основе контактов Шоттки вообще решает проблему быстродействия (включая СВЧ-диапазон). Режим работы диода при обратном смещении называется фотодиодным, а при прямых смещениях — вентильным.

Вопросы для самопроверки

1. Технология изготовления точечных диодов.
2. Особенности статической вольт-амперной характеристики

ки точечного диода.

3. Объясните причину отрицательного участка на вольт-амперной характеристике точечного диода.

4. Почему точечные диоды применяются в основном в диапазоне СВЧ?

5. Полупроводниковые стабилитроны и их основное применение.

6. Почему для производства стабилитронов в основном используется кремний?

7. Какие физические принципы лежат в основе работы диодов Шоттки?

8. Основные достоинства диодов Шоттки?

9. Какие полупроводники используются при изготовлении туннельных диодов?

10. В чём особенность статической вольт-амперной характеристики туннельного диода при прямом смещении на диоде?

11. Чем объясняется высокое быстродействие туннельных диодов?

12. Физические принципы образования фототока в фоторезисторах.

13. Какие конструкции фоторезисторов Вы знаете?

14. Принцип образования фототока в фотодиоде?

15. Статические вольт-амперные характеристики фотодиодов.

16. От каких параметров зависит чувствительность фотодиодов и фоторезисторов?

4. БИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

4.1 Введение

Биполярный транзистор наиболее распространенный тип транзисторов. Он может с успехом выполнять как усилительные функции, так и функции переключателя, т.е. представляет собой универсальный элемент электронных схем. Транзистор представляет собой двухпереходный прибор (рис. 4.1). Переходы образуются на границах тех трех слоев, из которых состоит транзистор. Контакты с внешними электродами — омические. В зависимости от типа проводимости крайних слоев различают транзисторы $p-n-p$ и $n-p-n$ с взаимно противоположными рабочими полярностями, что не имеет аналогии в электровакуумной технике. Чтобы не дублировать всех рассуждений и выводов, мы будем в дальнейшем рассматривать только транзисторы $p-n-p$. Условные обозначения обоих типов транзисторов, рабочие полярности напряжений и направления токов показаны на рис. 4.2. Переход, работающий в прямом направлении, называется эмиттерным, а соответствующий крайний слой — коллекторным. Такое название, как и у диодов, отражает факт инжекции неосновных носителей через переход. Средний слой называется базой. Второй переход, нормально смещенный в обратном направлении, называется коллекторным, а соответствующий крайний слой — коллектором. Это название отражает функцию «собираения» инжектированных носителей, прошедших через слой базы. Для того чтобы такое «собираение» было возможно, база должна иметь достаточно малую толщину. В противном случае инжектированные носители успеют рекомбинировать в процессе перемещения через базу, что мы видели на примере диодов с толстой базой. Необходимо подчеркнуть, что транзистор представляет собой, вообще говоря, обратимый прибор, т.е. эмиттер и коллектор можно поменять местами, сохранив в той или иной мере работоспособность прибора. Такой вывод вытекает из однотипности крайних слоев. Однако в связи с несимметричностью реальной структуры (рис. 4.3), а также различием материалов эмиттера и коллектора в большинстве типов транзисторов нормальное и инверсное включения неравноценны, в чем мы убедимся позднее.

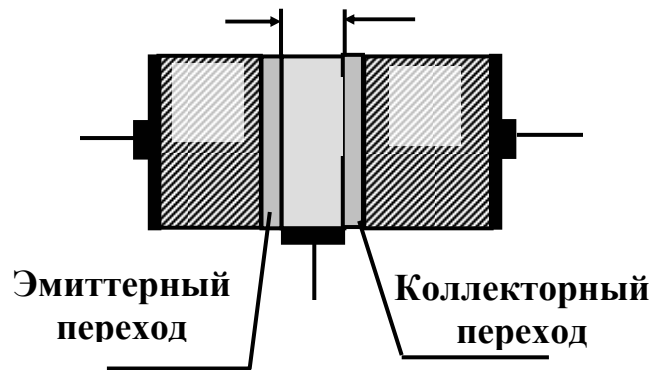
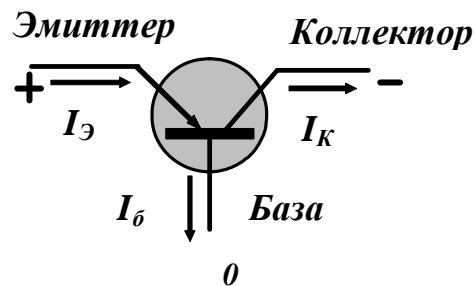
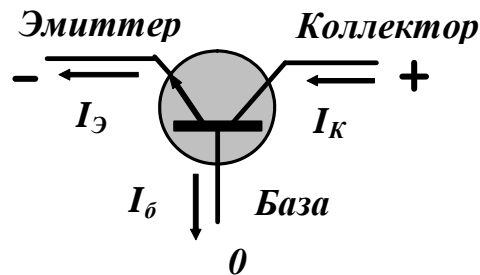


Рис. 4.1 — Упрощенная структура плоскостного транзистора



a



б

Рис. 4.2 — Условные обозначения транзисторов
a — транзистор *p-n-p*; *б* — транзистор *n-p-n*

В транзисторах типа *n-p-n* рабочими носителями являются электроны и полярности получаются такие же, как у электронных ламп. В транзисторах типа *p-n-p* рабочими носителями являются дырки и полярности соответствуют полярностям воображаемой позитронной лампы.

Транзистор иногда работает в режиме, когда оба перехода

смещены в прямом направлении. При этом, по существу, имеются двусторонняя инжекция и двустороннее «собираение» неосновных носителей. Если функция инжекции превалирует на обоих переходах, транзистор превращается в двойной диод. Однако чаще на одном из переходов (обычно коллекторном) превалирует функция «собираения», и тогда ток через него протекает в направлении, не соответствующем полярности смещения. Такой режим называется режимом насыщения.

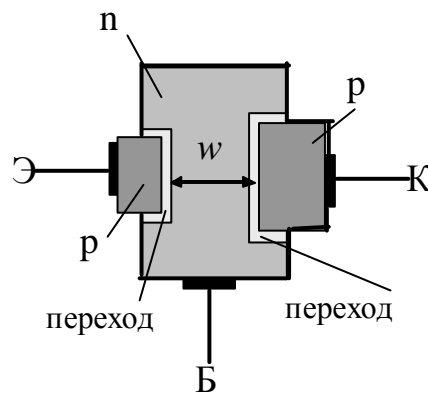


Рис. 4.3 — Реальная структура биполярного транзистора

Из всего сказанного следует, что плоскостной транзистор является системой двух взаимодействующих p - n переходов и что неизменным условием такого взаимодействия является достаточно малая толщина базы ($w \ll L$, где L — диффузионная длина неосновных носителей).

Основные свойства транзистора определяются процессами в базе, и им в дальнейшем будет уделено главное внимание. Характер движения инжектированных носителей в базе в общем случае заключается в сочетании диффузии и дрейфа. Электрическое поле, в котором происходит дрейф, может быть результатом высокого уровня инжекции, а также результатом неоднородности слоя. Последний случай имеет особенно большое значение, так как собственное поле неоднородного полупроводника обуславливает дрейфовый механизм движения носителей независимо от уровня инжекции. Транзисторы без собственного поля базы называются диффузионными или бездрейфовыми, а с собственным полем — дрейфовыми. Оба названия отражают главный меха-

низм перемещения носителей, хотя, как правило, диффузия и дрейф сочетаются. Более простыми для анализа являются бездрейфовые транзисторы, которым ниже уделено основное внимание. Особенности дрейфовых транзисторов будут рассмотрены в отдельном параграфе. Выше подразумевалось, что оба напряжения (U_{ε} и U_B) отсчитываются от базы, принятой за основной электрод, который является общим для входного и выходного напряжений транзистора. Такое включение транзистора (рис. 4.4, а), позволяющее строго и наглядно изучить его физические свойства и параметры, называют включением «с общей базой». Это включение будет в дальнейшем обозначаться буквами ОБ. Однако схема ОБ является не только не единственно возможной, но даже не наиболее распространенной на практике. Это объясняется рядом обстоятельств (например, отсутствием усиления тока), которые будут ясны из последующих параграфов.

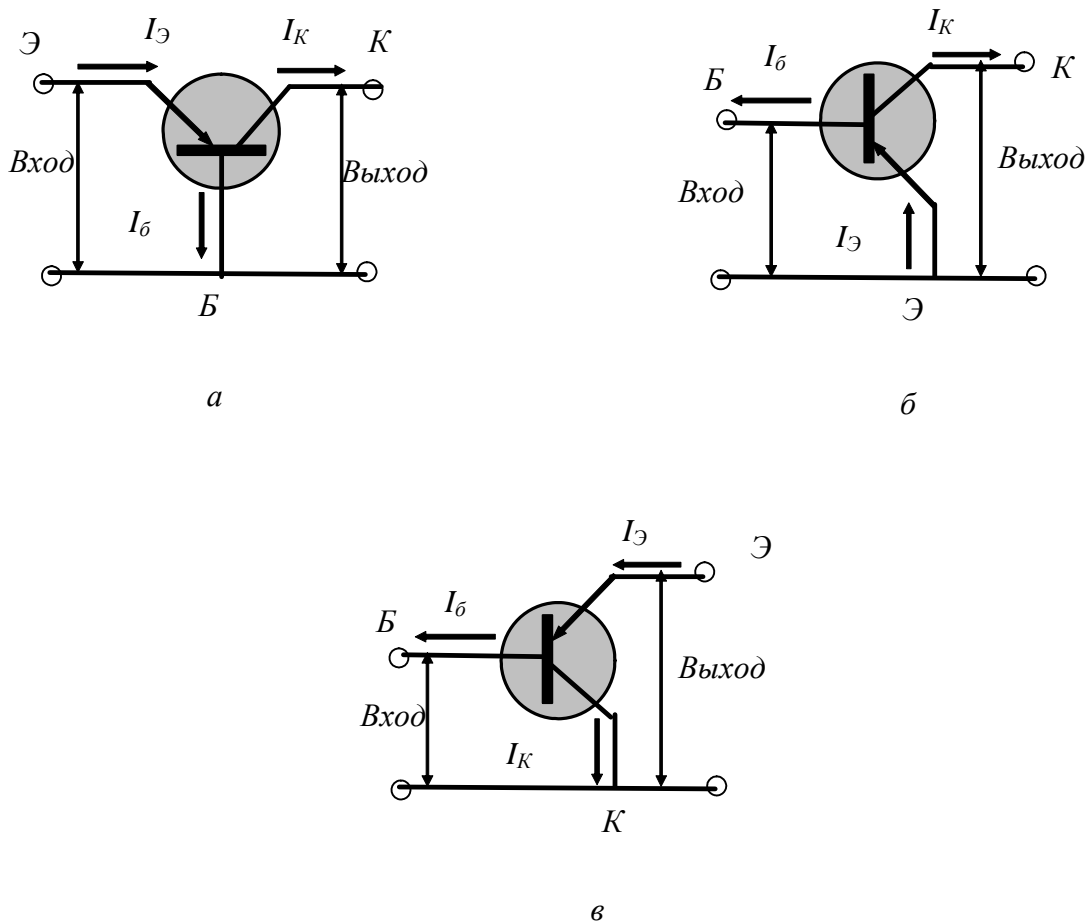


Рис. 4.4 — Схемы включения транзистора:
 а — с общей базой; б — с общим эмиттером;
 в — с общим коллектором

Основное применение в схемах находит другое включение транзистора, которое по вполне понятным причинам называют включением «с общим эмиттером» (рис. 4.4, б), и схема включения «с общим коллектором».

Мы будем обозначать их буквами ОЭ и ОК. Преимущества схемы ОЭ выяснятся чуть позже; однако можно сразу отметить, что она дает усиление по току, поскольку ток базы, являющийся для нее входным, гораздо меньше токов эмиттера и коллектора, это позволяет большее усиление по мощности в относительно узкой полосе частот. Третий вариант включения — схема «с общим коллектором» (ОК) — показан на рис. 4.4, в.

Несмотря на практические недостатки, которые в большей степени неверны, схема ОБ является основой при рассмотрении физических процессов в транзисторе. Поэтому она является основой при анализе основных процессов в биполярном транзисторе.

4.2 Основные процессы в биполярном транзисторе

На рис. 4.3 показан разрез бездрейфового транзистора со сплавными переходами, имеющего дисковую структуру. По сравнению с другими типами транзисторов сплавной транзистор наиболее прост и удобен для анализа. База этого транзистора однородна, поэтому механизм движения носителей — диффузионный. Удельные сопротивления слоев эмиттера и коллектора практически одинаковы, так что фактором, обуславливающим асимметрию транзистора, является различие площадей $S_{\text{Э}}$ и $S_{\text{К}}$. Асимметрия транзистора преследует ту цель, чтобы дырки, инжектируемые эмиттером и диффундирующие под некоторым углом к оси транзистора, по возможности полнее собирались коллектором. База сплавного транзистора отличается от базы идеализированной структуры (см. рис. 4.1) наличием трех участков, которые называют активной, промежуточной (или коллекторной) и пассивной областями базы. Активной областью базы является цилиндрический объем с высотой w и площадью, равной поверхности эмиттера $S_{\text{Э}}$. Промежуточной областью базы является кольцевой объем с площадью основания $S_{\text{К}} - S_{\text{Э}}$ и высотой, равной расстоянию от коллектора до противоположной поверхности

базовой пластинки. Наконец, пассивной областью базы является ее объем, расположенный вне коллектора.

Для первоначального ознакомления с транзистором можно пренебречь пассивной и промежуточной областями базы и считать транзистор симметричным, имеющим одинаковую площадь $S_{\mathcal{E}} = S_{\mathcal{K}}$ во всех сечениях (см. рис. 4.1). Обычно соблюдается соотношение $w \ll \sqrt{S}$ т.е. размеры транзистора в направлениях, перпендикулярных главной оси, много больше толщины базы. В таком транзисторе краевые эффекты не очень существенны, и его можно считать одномерным, т.е. можно предположить движение носителей только вдоль главной оси, без отклонения в стороны. Такая одномерная модель будет всегда подразумеваться, если не сделано специальных оговорок

Инжекция и собирание неосновных носителей. На рис. 4.5, *а* показана энергетическая диаграмма транзистора в равновесном состоянии. Диаграмма показывает, что эмиттер и коллектор представляют собой низкоомные слои (уровень Ферми лежит вблизи уровней акцепторов), а база — сравнительно высокоомный слой (уровень Ферми расположен вблизи середины запрещенной зоны). Легко видеть, что электроны базы и дырки эмиттера и коллектора находятся в «потенциальных ямах», из которых они могут перейти в смежный слой только благодаря достаточно большой тепловой энергии. Наоборот, дырки базы и электроны эмиттера и коллектора находятся на «потенциальных гребнях», с которых они могут свободно переходить в смежный слой. В равновесном состоянии на обоих переходах имеется динамическое равновесие между потоками дырок (а также между потоками электронов), протекающих в ту и другую стороны. Пусть на эмиттерном переходе задано нормальное для него положительное смещение, а коллекторный переход по-прежнему замкнут (рис. 4.5, *б*). Тогда потенциальный барьер эмиттера понизится и начнется инжекция дырок в базу и электронов в эмиттер. При большой разнице в удельных сопротивлениях слоев электронная составляющая тока, как известно, не играет большой роли, и ею пока можно пренебречь. Инжектированные дырки, пройдя базу, доходят до коллекторного перехода и свободно проходят в коллектор. Значит, в выходной цепи будет протекать ток,

близкий к току эмиттера, поскольку рекомбинация в тонкой базе невелика. Небольшая разность между эмиттерным и коллекторным токами связана с током базы, который обусловлен электронами, поступающим из внешней цепи. Электроны внешней цепи компенсируют убыль электронов в базе в результате их рекомбинации с дырками. Поскольку напряжение U_K равно нулю, полезная мощность не выделяется и усиление отсутствует. Если в выходную цепь включить резистор нагрузки (рис. 4.5, в), то падение напряжения на этом сопротивлении создаст положительное смещение коллектора. При этом, наряду с собиранием дырок, дошедших от эмиттера, будет происходить инжекция дырок самим коллектором. В результате коллекторный ток станет заметно меньше тока эмиттера и мощность в нагрузке будет очень невелика. В нормальном усилительном режиме на коллекторный переход задается достаточно большое отрицательное смещение, в результате чего потенциальный барьер у коллектора соответственно увеличивается (рис. 4.5, г). Теперь можно включать в выходную цепь резистор со значительным сопротивлением без опасения вызвать инжекцию через коллекторный переход. При этом «собирательные функции» коллекторного перехода никак не ухудшатся.

При этом можно получить значительную выходную мощность, а главное — усиление мощности, так как токи I_E и I_K почти одинаковы, а сопротивление нагрузки превышает сопротивление эмиттерного перехода.

Чтобы коллекторный переход не сместился в прямом направлении, должно выполняться условие.

$$-E_K + I_K R_K < 0.$$

Пусть теперь коллекторный переход смещен в обратном направлении, а эмиттер «оборван» (рис. 4.5, з).

Высокий потенциальный барьер коллектора практически исключает уход дырок из коллектора. Следовательно, ток через коллекторный переход будет обусловлен неуравновешенным потоком дырок из базы в коллектор. Токи I_{k0} и I_{B0} в этом случае невелики и, конечно, равны друг другу. Экстракция дырок из базы через коллекторный переход создает отрицательный градиент концентрации их вдоль всей базы. В результате тот поток дырок

из базы в эмиттер, который был в равновесном состоянии (рис. 4.5, а), уменьшается и остается неуравновешенный поток достаточно «энергичных» дырок из эмиттера в базу. Соответственно в эмиттерном слое образуется отрицательный заряд, а в базовом слое такой же положительный заряд; энергетические уровни электронов в эмиттере повышаются, затрудняя переход дырок слева направо. Этот процесс продолжается до тех пор, пока потоки дырок из эмиттера и в эмиттер снова не уравновесятся.

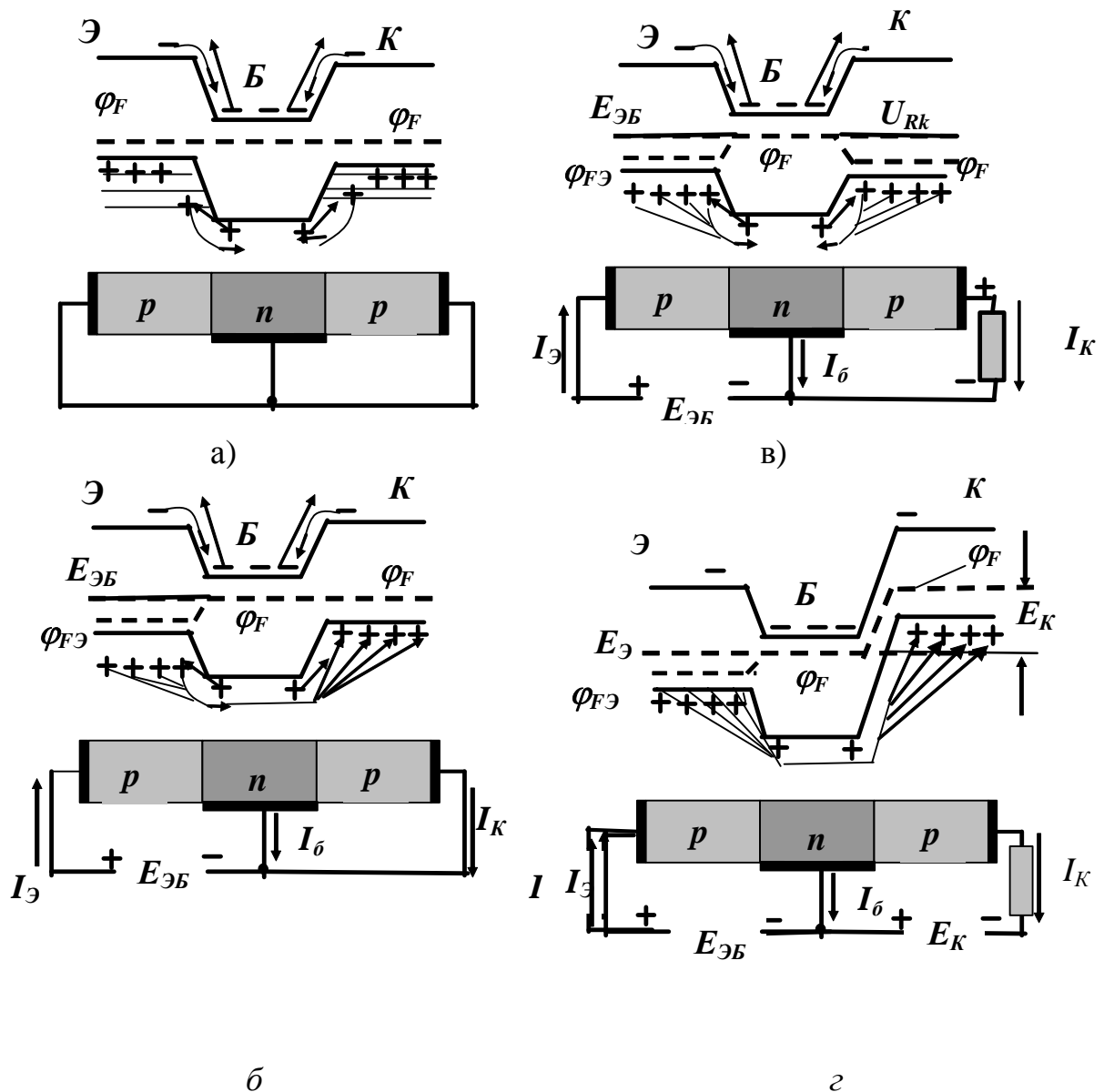


Рис. 4.5 — Зонные диаграммы для плоскостного транзистора при различных режимах его работы

Во всех случаях у транзистора $p-n-p$ главными рабочими носителями, образующими токи через переходы, являются дырки, тогда как ток базы всегда обусловлен электронами; последние компенсируют избыточный заряд дырок в базе и обеспечивают ее нейтральность как во время переходных процессов, так и в стационарном режиме, когда убыль дырок обусловлена только рекомбинацией.

Распределение носителей в базе. Дырки, инжектированные эмиттером, достигают коллектора не сразу, а с некоторой задержкой, обусловленной их перемещением вдоль базы. Кроме того, в связи с хаотичностью движения дырок коллекторный ток нарастает не скачком, а плавно.

Соответственно ток базы в первый момент равен току эмиттера, а затем постепенно уменьшается до стационарной величины. Примерная картина переходного процесса показана на рис. 4.6, где t_3 — время задержки, а t_ϕ — длительность фронта.

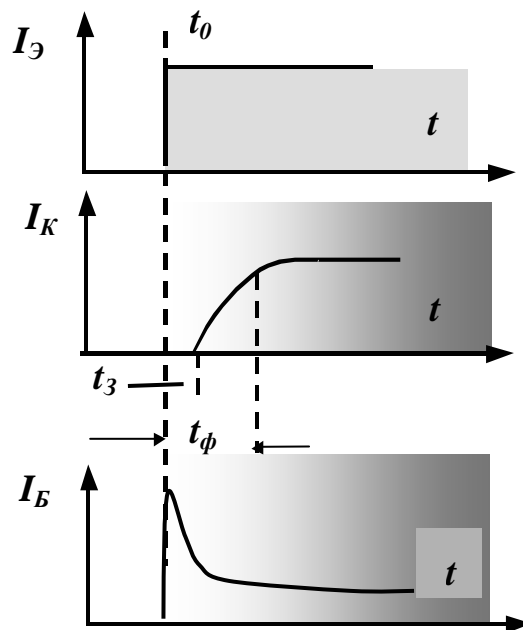


Рис. 4.6 — Переходные процессы при подаче ступеньки эмиттерного тока

Распределение дырок в базе почти линейно, как показано на рис. 4.7. На самом деле градиент концентрации около коллекторного перехода несколько меньше, чем около эмиттерного, поскольку ток коллектора (из-за рекомбинации) немного меньше эмиттерного тока. Эту разницу в градиентах следует иметь в виду,

но ее трудно отразить на графике. Линейному распределению дырок должно соответствовать почти линейное распределение компенсирующих (избыточных) электронов в базе (рис. 4.7). Заряды избыточных носителей пропорциональны площадям под кривыми их распределения. Поскольку база в целом нейтральна, можно считать эти площади одинаковыми. Для оценки заряда воспользуемся распределением дырок. Очевидно, что заряд дырок пропорционален толщине базы и току транзистора, определяющему наклон линии $p(x)$. Вопрос о том, какому из двух токов (I_E или I_K) пропорционален заряд, не очень существен, так как эти токи в стационарном режиме почти одинаковы. В большинстве случаев удобнее считать заряд пропорциональным току коллектора, так как в схемах этот ток обычно не претерпевает скачкообразных изменений. Эквивалентную емкость, обусловленную изменениями заряда в базе, называют, как и в диоде, диффузионной емкостью.

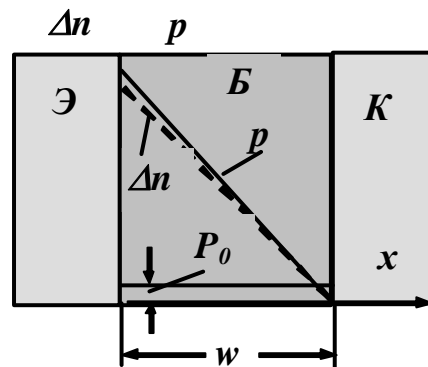


Рис. 4.7 — Распределение дырок и избыточных электронов в базе

Модуляция толщины базы. Как известно, ширина p - n перехода зависит от напряжения на нем. Поскольку эмиттерный переход смещен в прямом направлении, его ширина мала и изменения этой ширины при изменениях U_E не имеют существенного значения.

Коллекторный же переход, смещенный в обратном направлении, имеет сравнительно большую ширину, и изменения ее при изменениях напряжения U_K важны для работы транзистора. А именно, поскольку коллекторный переход сосредоточен в базе (как более высокоомном слое), приращение его ширины оказываются практически равными приращениям толщины базы. В результате получается зависимость $w(U_K)$, которую называют мо-

дуляцией толщины базы или эффектом Эрли. Проанализируем эффект модуляции базы.

Во-первых, изменение толщины базы влияет на ту долю инжектированных дырок, которая доходит до коллектора, избегая рекомбинации. Чем меньше толщина базы, тем эта доля больше. Значит, при неизменном токе эмиттера модуляция толщины базы приводит к изменениям тока коллектора. Соответственно коэффициент передачи эмиттерного тока оказывается функцией коллекторного напряжения, а коллекторный переход имеет конечное дифференциальное сопротивление.

Во-вторых, модуляция толщины базы сопровождается изменением заряда дырок в базе; иначе говоря, имеет место зависимость заряда от коллекторного напряжения, т.е. коллекторный переход обладает некоторой диффузионной емкостью дополнительно к обычной барьерной.

В-третьих, модуляция толщины базы меняет время диффузии дырок через базу; тем самым коллекторное напряжение влияет на частотные свойства транзистора.

В-четвертых, поскольку тепловой ток эмиттерного перехода $I_{Э0}$ при тонкой базе обратно пропорционален ее толщине, напряжение U_K , модулируя толщину базы, модулирует также ток $I_{Э0}$, а вместе с ним, согласно (2.33), всю вольт-амперную характеристику эмиттерного перехода. Следовательно, если одна из входных величин ($I_{Э}$ или $U_{Э}$) задана, то вторая оказывается функцией коллекторного напряжения (рис. 4.8). Такое влияние разумно назвать внутренней обратной связью по напряжению.

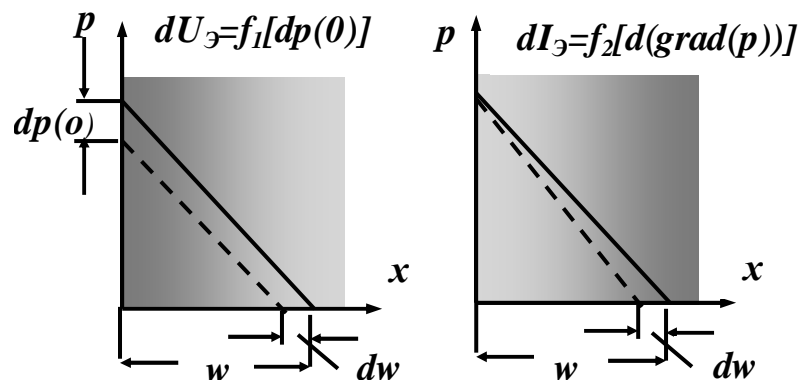


Рис. 4.8 — Влияние модуляции толщины базы на входные величины: а — $I_{Э} = const$; б — $U_{Э} = const$

4.3 Статические характеристики транзистора

Выше была рассмотрена идеализированная модель транзистора. Идеализация заключалась не только в том, что модель считалась одномерной, но и в том, что не учитывались объемные сопротивления слоев. Сопротивления слоев эмиттера и коллектора существенны только в некоторых ключевых режимах. Сопротивление же базы существенно почти во всех случаях, но чтобы не усложнять предварительный анализ, мы учтем его позднее.

Формулы Мола-Эберса. Для транзистора можно принять эквивалентную схему, которая показана на рис. 4.9. Здесь каждый из переходов изображен в виде диода, а взаимодействие их отражено генераторами токов. Если эмиттерный переход открыт и через него протекает ток I_1 , то в цепи коллектора, как известно, будет протекать несколько меньший ток, поскольку часть инжектированных носителей рекомбинирует.

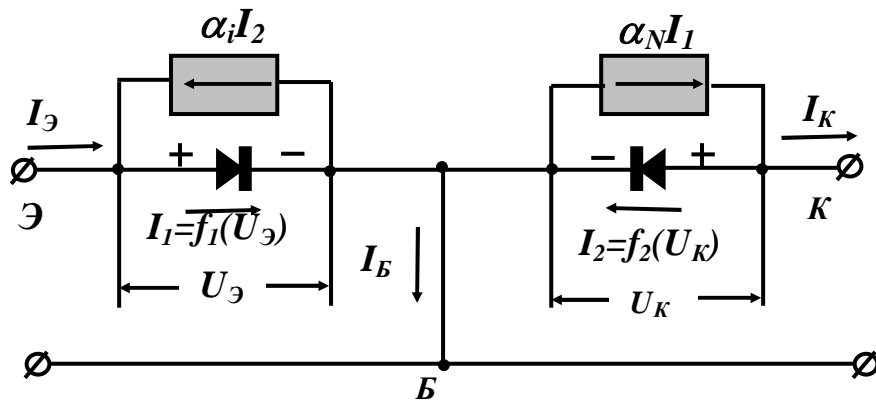


Рис. 4.9 — Эквивалентная схема идеализированного транзистора

Этот меньший ток обеспечивается на схеме генератором $\alpha_N I_1$, где $\alpha_N < 1$ — коэффициент передачи эмиттерного тока. Индекс N означает нормальное включение транзистора. Если триод работает в инверсном включении (положительное смещение на коллекторе и отрицательное на эмиттере), то прямому коллекторному току I_2 соответствует эмиттерный ток $\alpha_i I_2$, вытекающий из эмиттера. Коэффициент α_i есть коэффициент передачи коллекторного тока.

Таким образом, в общем случае токи эмиттера и коллектора складываются из двух компонентов: инжектируемого и собираемого

$$I_{\mathcal{E}} = I_1 - \alpha_i I_2; \quad (4.1a)$$

$$I_K = \alpha_n I_1 - I_2. \quad (4.1б)$$

Связь инжектируемых компонентов с напряжениями на переходах такая же, как и в отдельном диоде, т.е. в простейшем случае выражается формулами:

$$I_1 = I_{\mathcal{E}0}^1 \left(e^{U_{\mathcal{E}}/\varphi_T} - 1 \right); \quad (4.2a)$$

$$I_2 = I_{K0}^1 \left(e^{U_K/\varphi_T} - 1 \right), \quad (4.2б)$$

где $I_{\mathcal{E}0}^1$ и I_{K0}^1 — тепловые токи эмиттерного и коллекторного диодов (соответственно при $U_K = 0$ и $U_{\mathcal{E}} = 0$). Тепловые токи $I_{\mathcal{E}0}^1$ и I_{K0}^1 можно выразить через токи, которые обычно задаются в технической документации на транзистор, а именно через токи $I_{\mathcal{E}0}$ и I_{K0} , измеряемые при обрыве соответственно коллектора и эмиттера. Эти токи связаны соотношениями:

$$I_{K0}^1 = \frac{I_{K0}}{1 - \alpha_N \alpha_i}, \quad (4.3a)$$

аналогичным путем получим:

$$I_{\mathcal{E}0}^1 = \frac{I_{\mathcal{E}0}}{1 - \alpha_N \alpha_i}. \quad (4.3б)$$

Подставив токи из (4.2) в соотношения (4.1), получим уравнения, описывающие статические вольт-амперные характеристики транзистора:

$$I_{\mathcal{E}} = I_{\mathcal{E}0}^1 \left(e^{U_{\mathcal{E}}/\varphi_T} - 1 \right) - \alpha_i I_{K0}^1 \left(e^{U_K/\varphi_T} - 1 \right); \quad (4.4a)$$

$$I_K = (1 - \alpha_N) I_{\mathcal{E}0}^1 \left(e^{U_{\mathcal{E}}/\varphi_T} - 1 \right) - I_{K0}^1 \left(e^{U_K/\varphi_T} - 1 \right). \quad (4.4б)$$

Запишем ток базы:

$$I_{\bar{o}} = (1 - \alpha_N) I_{\bar{\Delta}0}^1 \left(e^{U/\varphi_T} - 1 \right) + (1 - \alpha_i) I_{K0}^1 \left(e^{U_K/\varphi_T} - 1 \right);$$

$$I_{\bar{o}} = (1 - \alpha_N) I_{\bar{\Delta}0}^1 \left(e^{U/\varphi_T} - 1 \right) + (1 - \alpha_i) I_{K0}^1 \left(e^{U_K/\varphi_T} - 1 \right). \quad (4.4\text{в})$$

Формулы Мола-Эберса (4.4), несмотря на их приближенность, очень полезны для анализа статических режимов, так как хорошо отражают основные особенности транзисторов при любых сочетаниях напряжений на переходах. Обратные токи у кремниевых транзисторов нельзя считать тепловыми, поскольку гораздо большую роль играют токи термогенерации. Поэтому количественные расчеты по формулам (4.4) в случае кремниевых транзисторов дают значительную погрешность.

В транзисторах выполняется соотношение:

$$\alpha_N I_{\bar{\Delta}0} = \alpha_i I_{K0}. \quad (4.5)$$

Это соотношение иногда позволяет упростить формулы (4.4) и выводы из них. В первом приближении можно полагать $I_{\bar{\Delta}0} = I_{K0}$.

Идеализированные статические характеристики. В гл. 2 было показано, что задать прямое напряжение на p - n переходе трудно. Поэтому в большинстве случаев целесообразно считать заданной величиной эмиттерный ток, а не эмиттерное напряжение.

$$I_K = \alpha_N I_{\bar{\Delta}} - I_{K0} \left(e^{U_K/\varphi_T} - 1 \right). \quad (4.6)$$

Это выражение представляет собой семейство коллекторных характеристик $I_K(U_K)$ с параметром $I_{\bar{\Delta}}$. Такое семейство показано на рис. 4.10, *a*.

Семейство эмиттерных характеристик $U_{\bar{\Delta}}(I_{\bar{\Delta}})$ с параметром U_K получается из выражения (4.4а), если разрешить его относительно $U_{\bar{\Delta}}$. Используя соотношение (4.5), получаем

$$U_{\bar{\Delta}} = \varphi_T \ln \left[\frac{I_{\bar{\Delta}}}{I_{\bar{\Delta}0}} + 1 + \alpha_N \left(e^{U_K/\varphi_T} - 1 \right) \right]. \quad (4.7)$$

Эмиттерное семейство характеристик показано на рис. 4.10, б. Из рис. 4.10, а ясно видны два резко различных режима работы транзистора: активный режим, соответствующий значениям $U_K < 0$ (первый квадрант), и режим насыщения, соответствующий значениям $U_K > 0$ (второй квадрант). Активный режим является основным в усилительной технике и будет подробно изучен в последующих параграфах. Режим насыщения характерен для ключевых импульсных схем и будет рассмотрен позже. Для активного режима характерны условия $U_K < 0$ и $|U_K| \gg \varphi_T$, при которых формулы (4.6) и (4.7) упрощаются

$$I_K = \alpha I_{\mathcal{E}} + I_{K0}; \quad (4.8)$$

$$I_{\mathcal{E}} = \varphi_T \ln \frac{I_{\mathcal{E}}}{I_{\mathcal{E}0}^1}. \quad (4.9)$$

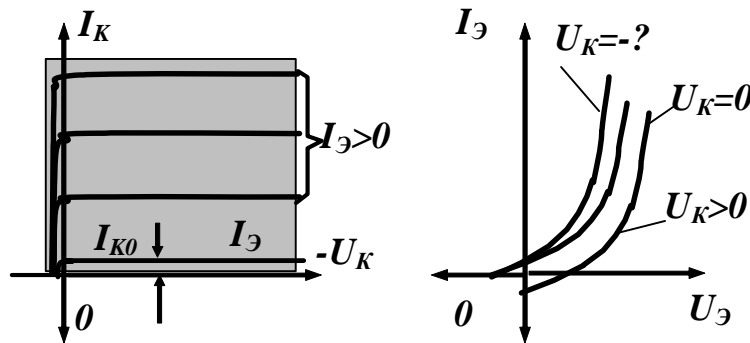


Рис. 4.10 — Статические характеристики идеализированного транзистора: а — коллекторные; б — эмиттерные

В формуле (4.8), широко используемой на практике, для простоты опущен индекс N при коэффициенте α_N , а при выводе формулы (4.9) для простоты положено $1 - \alpha_N = 0$, что вполне оправдано. Реальные характеристики имеют конечный наклон, обусловленный неучтенным в формулах (4.4) сопротивлением коллекторного перехода (следствие модуляции толщины базы). Относительно эмиттерного семейства (рис. 4.10, б) можно сделать следующие замечания. Кривая с параметром $U_K = 0$, естественно, является обычной диодной характеристикой. При значениях $U_K > 0$ кривые сдвигаются вправо и вниз в связи с нарастанием собираемого компонента эмиттерного тока. При значениях

$U_K < 0$ кривые несколько смещаются влево и вверх. На реальных характеристиках, как увидим ниже, влияние отрицательного напряжения U_K тоже невелико, но все же имеет место при любых значениях U_K из-за внутренней обратной связи по напряжению (следствие модуляции толщины базы).

Реальные статические характеристики. В формулах Мола-Эберса не учитывается целый ряд факторов, таких, как эффект Эрли, пробой перехода, зависимость α от тока и др. Поэтому характеристики рис. 4.10 в значительной степени идеализированы. Реальные коллекторные и эмиттерные характеристики показаны на рис. 4.11. Кривые коллекторного семейства имеют конечный, хотя и очень небольшой, наклон, который в области, близкой к пробую, резко увеличивается. Расстояние между кривыми немного уменьшается при больших токах из-за уменьшения α .

На рис. 4.11, *a* проведена гипербола допустимой мощности, рассеиваемой в основном на коллекторном переходе. При нагреве транзистора кривые смещаются вверх в область больших токов из-за роста тока I_{K0} . В активном режиме (1-й квадрант), усредняя нелинейное сопротивление r_K , можно характеризовать коллекторное семейство ОБ достаточно строгим соотношением

$$I_K = \alpha I_{\mathcal{E}} + I_{K0} + \frac{U_K}{r_K}. \quad (4.10)$$

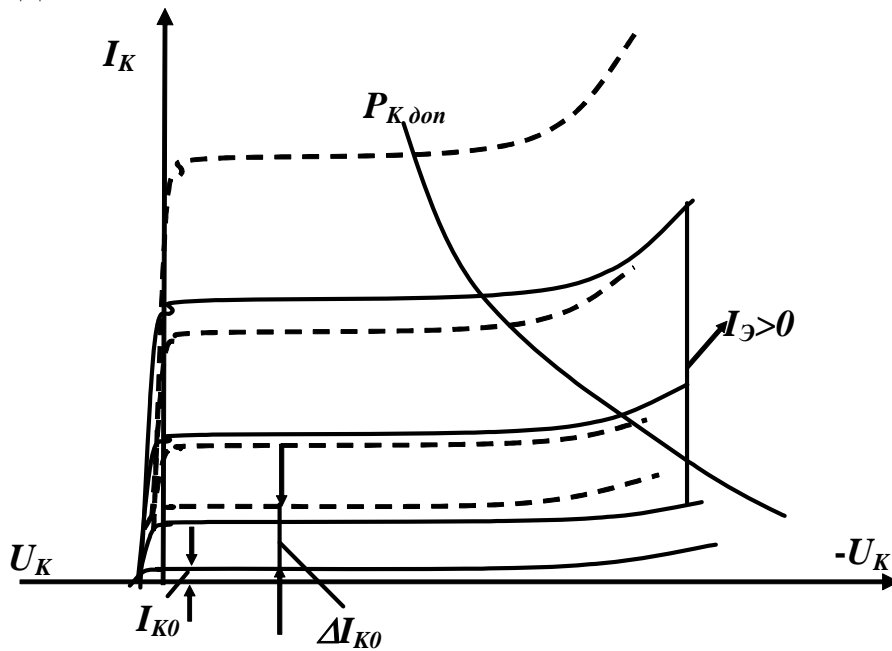
При нагреве транзистора кривые смещаются влево в область меньших напряжений. При одном и том же эмиттерном токе эмиттерные напряжения у кремниевых транзисторов на 0,3 — 0,4 В больше, чем у германиевых, но все же обычно не превышают 0,6 — 0,7 В. При достаточно большом токе входные вольт-амперные характеристики деформируются.

Рассмотрим особенности выходных характеристик в области пробоя. Если любой из переходов транзистора использовать в качестве диода, то пробой при обратном напряжении будет иметь такой же характер, как в изолированном переходе.

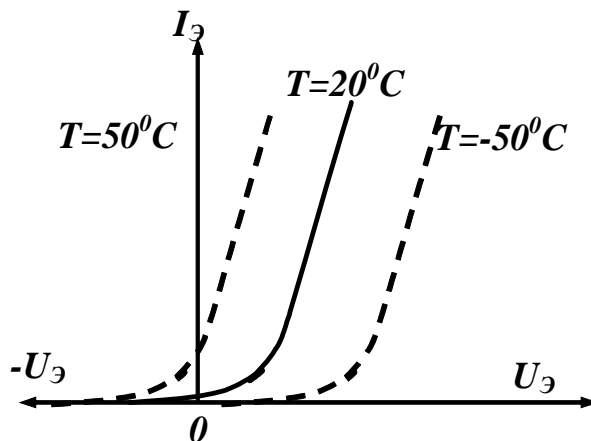
Например, при оборванном эмиттере коллекторный ток будет равен MI_{K0} (где M — коэффициент ударной ионизации), а напряжением лавинного пробоя будет величина U_M , при которой $M = \infty$. При любом конечном токе эмиттера ток I_K тоже увели-

чивается в M раз и, следовательно, напряжением пробоя опять будет величина U_M .

Необходимо, однако, заметить, что чисто лавинный пробой имеет место только при достаточно быстром повышении коллекторного напряжения, когда температура перехода не успеваеет заметно повыситься. В результате, учитывая (4.3а), получим эквивалентную схему, показанную на рис. 4.12, б и соответствующую выражению (4.8). Такая схема полезна для расчета режима усилительных каскадов.



а



б

Рис. 4.11 — Статические характеристики транзистора при включении по схеме с общей базой: а — выходные; б — входные

Сопротивления слоев коллектора и эмиттера здесь опущены, так как в усилительной технике они несущественны. Однако включение этих сопротивлений в схему не приводит ни к каким затруднениям, так как через них протекают заданные токи, и, значит, соответствующие напряжения легко рассчитать и добавить к напряжениям на переходах.

4.4 Статические параметры транзистора

Нелинейные эквивалентные схемы, показанные на рис. 4.12, используются при анализе вопросов, связанных с большим сигналом. При расчете малых переменных составляющих, характерных для усилительной техники, эти схемы целесообразно линеаризовать.

Возьмем за основу схему на рис. 4.12, б. Генератор постоянного тока I_{K0} исключим, поскольку нас интересуют переменные составляющие, и введем вместо него дифференциальное сопротивление коллекторного перехода r_K . Эмиттерный диод также заменим его дифференциальным сопротивлением r_E . Обратную связь по напряжению отразим генератором э.д.с., включенным последовательно с сопротивлением r_E .

Наконец, для учета частотных зависимостей включим параллельно сопротивлениям r_E и r_K барьерные емкости, а коэффициент α будем считать операторной или комплексной величиной. Тогда линейная эквивалентная схема транзистора будет такой, как показано на рис. 4.13. Ее легко дополнить паразитными емкостями, однако это редко необходимо.

Схема на рис. 4.13 хорошо отражает структуру транзистора и содержит физически обоснованные параметры. Точка B^1 на схеме называется внутренней базовой точкой в отличие от внешнего зажима базы.

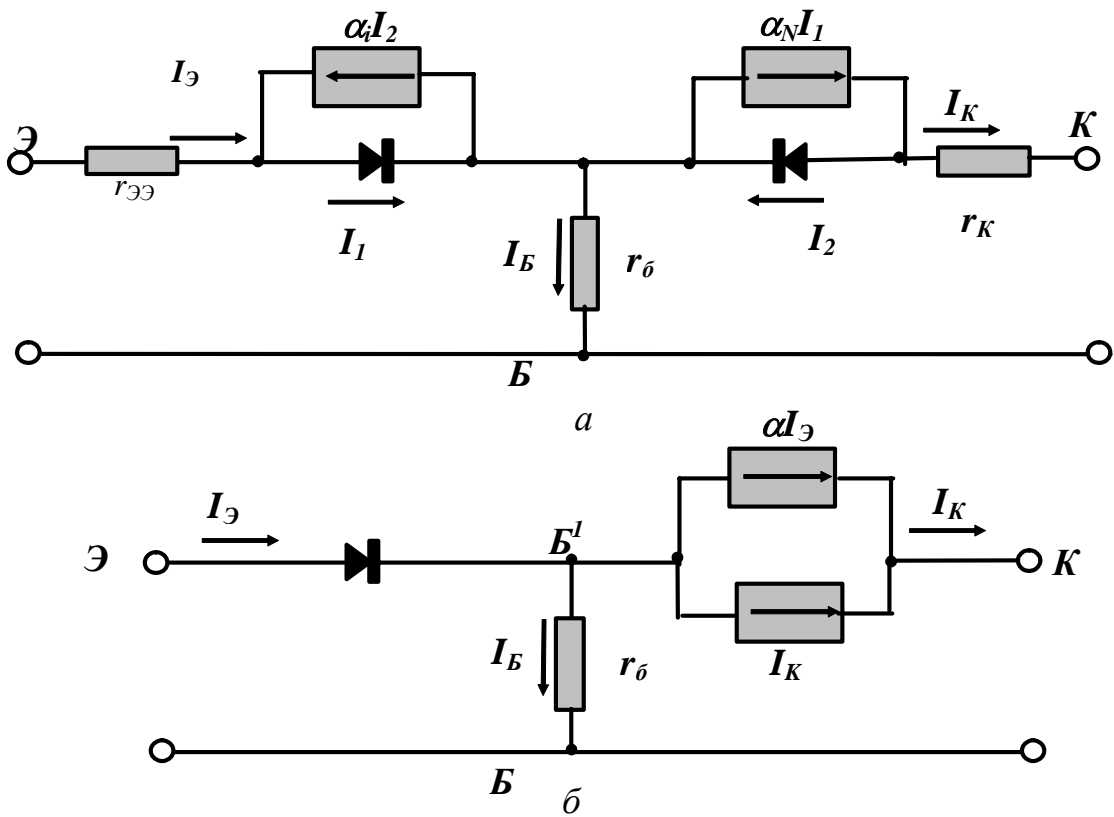


Рис. 4.12 — Эквивалентные схемы транзистора для постоянных составляющих: *a* — общая схема с учетом сопротивлений слоев; *б* — схема для нормального активного режима

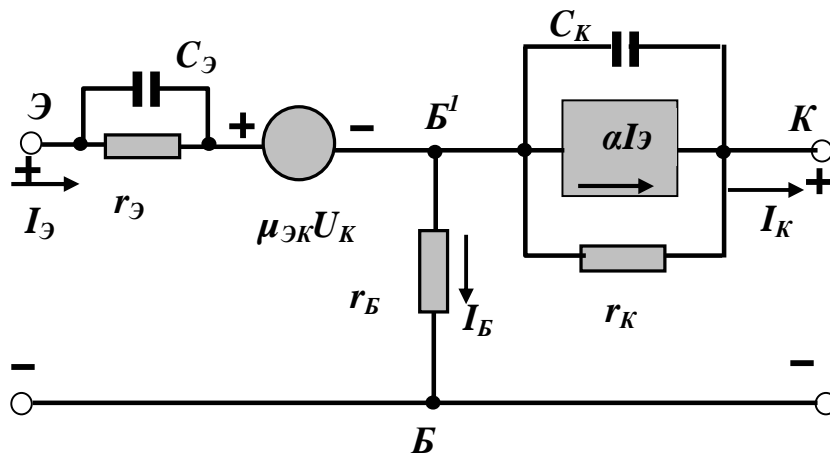


Рис. 4.13 — Эквивалентная схема транзистора для переменных токов

К числу основных параметров, необходимых при построении эквивалентной схемы транзистора (для переменных составляющих), относятся следующие: 1. Дифференциальный коэффициент передачи эмиттерного тока при нормальном включении α_N . В дальнейшем будем писать его без индекса N , так как инверсное включение относится к специальным случаям. Коэффициент α определяется следующим образом:

$$\alpha = \left(\frac{dI_K}{dI_{\mathcal{E}}} \right)_{U_K = \text{const}} . \quad (4.11a)$$

2. Дифференциальное сопротивление эмиттерного перехода

$$r_{\mathcal{E}} = \left(\frac{dU_{\mathcal{E}}}{dI_{\mathcal{E}}} \right)_{U_K = \text{const}} . \quad (4.11b)$$

3. Дифференциальное сопротивление коллекторного перехода

$$r_K = \left(\frac{dU_K}{dI_K} \right)_{U_{\mathcal{E}} = \text{const}} . \quad (4.11b)$$

4. Коэффициент внутренней обратной связи по напряжению, характеризующий влияние коллекторного напряжения на эмиттерный переход в связи с модуляцией толщины базы,

$$\mu_{\mathcal{E}K} = \left(\frac{dU_{\mathcal{E}}}{dU_K} \right)_{I_{\mathcal{E}} = \text{const}} . \quad (4.11g)$$

5. Объемное сопротивление базы r_B . В отличие от предыдущих параметров сопротивление базы должно определяться не для одномерной модели, а для реальной структуры транзистора. Ток базы протекает в направлении, перпендикулярном потоку дырок, и, следовательно, необходимо учитывать реальную конфигурацию базы, т.е. активную и пассивную ее части. Помимо перечисленных дифференциальных параметров, важную роль в работе транзистора играет тепловой ток I_{K0} , который определяется следующим образом:

$$I_{K0} = (I_K)_{I_{\mathcal{E}}=0}; \quad |U_K| \gg \varphi_T .$$

Коэффициент передачи эмиттерного тока. Величина α , стоящая в формуле (4.8), в отличие от величины α в формуле

(4.11a) является интегральной, так как связывает не приращения $\partial I_{\mathcal{E}}$ и ∂I_K , а полные токи I_K и $I_{\mathcal{E}}$.

$$\alpha_{ИН} = \frac{I_K - I_{K0}}{I_{\mathcal{E}}}. \quad (4.12a)$$

Если бы коэффициент α не зависел от тока $I_{\mathcal{E}}$, то, как следует из (4.8), дифференциальный коэффициент передачи был бы равен интегральному значению. На самом деле он является функцией эмиттерного тока, и поэтому, продифференцировав (4.8) по току $I_{\mathcal{E}}$, получим следующее соотношение:

$$\alpha = \alpha_{ИН} + \frac{dI_0}{dI_{\mathcal{E}}}. \quad (4.12b)$$

Выразим коэффициент передачи тока эмиттера через физические параметры транзистора. Эта задача будет линейной только в том случае, если пренебречь зависимостью $\alpha(I_{\mathcal{E}})$, считая $\alpha = \alpha_{ИН}$. Поэтому в дальнейшем мы будем различать интегральный и дифференциальный коэффициенты передачи лишь тогда, когда это принципиально необходимо. Коэффициент передачи эмиттерного тока можно записать в следующем виде:

$$\alpha = \gamma\chi, \quad (4.13)$$

где γ — коэффициент инжекции дырок; χ — коэффициент переноса дырок через базу, показывающий, какая доля инжектированных дырок доходит до коллектора. На низких частотах коэффициент инжекции в транзисторах не имеет специфики по сравнению с диодами и выражается известными формулами. Поэтому, прежде всего, проанализируем коэффициент переноса. Для этого решим уравнение диффузии (1.39a) для стационарного режима $\frac{\partial p}{\partial t} = 0$. В этом случае уравнение будет по форме таким же, как (2.16); запишем его для полной концентрации

$$\frac{dp}{dx^2} - \frac{p}{L^2} = -\frac{p_0}{L^2}. \quad (4.14)$$

Общее решение этого уравнения будет таким же, как (2.17), а частным решением будет p_0 , поэтому

$$p(x) = A_1 e^{\frac{x}{L}} + A_2 e^{-\frac{x}{L}} + p_0. \quad (4.15)$$

Граничные условия запишем, исходя из того, что в эмиттерной цепи задан дырочный ток, $I_{\text{ЭП}} = \gamma I_{\text{Э}}$, а на коллекторном переходе — напряжение U_K . Учитывая (1.34а) и (2.13а), получаем при $x = 0$ и $x = w$ соответственно:

$$-\frac{dp}{dx}\Big|_{x=0} = \frac{I_{\text{ЭП}}}{qDS}; \quad (4.16а)$$

$$p(w) = p_0 e^{\frac{U_K}{\Phi_T}}, \quad (4.16б)$$

где S — как и раньше, площади переходов. Используя граничные условия (4.16), определяем коэффициенты A_1 и A_2 , входящие в уравнение (4.15). После этого распределение концентрации можно привести к следующему виду:

$$p(x) = \frac{I_{\text{ЭП}}L}{qDS} \frac{\text{sh}\left(\frac{w-x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)} + p_0 \left[e^{\frac{U_K}{\Phi_T}} - 1 \right] \frac{\text{sh}\left(\frac{w-x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)}. \quad (4.17а)$$

При нормальном рабочем режиме ($U_K < 0$ и $|U_K| \gg \Phi_T$), а также при условии $w \ll L$ вторым членом в правой части (4.17а) можно пренебречь. Тогда:

$$p(x) = \frac{I_{\text{ЭП}}L}{qDS} \frac{\text{sh}\left(\frac{w-x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)}. \quad (4.17б)$$

Если пренебречь рекомбинацией в базе, положив $L \rightarrow \infty$, что выполняется далеко не всегда, то формула (4.17б) принимает вид

$$p(x) = \frac{I_{\text{ЭП}}L}{qDS} \left(1 - \frac{x}{w}\right). \quad (4.17в)$$

Как видим, стационарное распределение дырок в тонкой базе близко к линейному, на что уже обращалось внимание выше при анализе диода. Дифференцируя (4.17б) по x , умножая обе части на $-Dqs$ и принимая во внимание, что $x = w$, получаем для коллекторного тока выражение

$$I_k = \frac{I_{ЭП}}{ch(w/L)}.$$

Беря производную, находим коэффициент переноса

$$\chi = \frac{1}{ch\left(\frac{w}{L}\right)} = \operatorname{sech}\left(\frac{w}{L}\right). \quad (4.18)$$

Формула (4.18) является одной из фундаментальных в теории транзисторов. С увеличением толщины базы коэффициент переноса согласно (4.18) неограниченно уменьшается и при $w > L$ близок к нулю, т.к. неосновные носители успевают рекомбинировать в базе, не доходя до коллекторного перехода. Поэтому в транзисторе, как уже отмечалось, базу делают как можно более тонкой, так чтобы всегда выполнялось условие $w \ll L$. Тогда, разлагая $ch(z)$ в ряд с точностью до трех первых членов, приходим к выражению

$$\chi = \frac{1}{1 + \frac{1}{2}\left(\frac{w}{L}\right)^2 + \frac{1}{24}\left(\frac{w}{L}\right)^4}. \quad (4.19a)$$

Учитывая, что второй и третий члены знаменателя много меньше единицы, можно пользоваться приближением:

$$\chi = 1 - \frac{1}{2}\left(\frac{w}{L}\right)^2. \quad (4.19б)$$

Заметим, что для расчета коэффициента переноса второй член вообще несущественен (независимо от режима и толщины базы), поскольку в него не входит эмиттерный ток.

Коэффициент передачи тока α согласно (4.13) получается путем умножения любого из выражений (4.18) или (4.19) на коэффициент инжекции. Так, из наиболее строгого выражения (4.18) следует

$$\alpha = \gamma \operatorname{sech}\left(\frac{w}{L}\right), \quad (4.20)$$

а приближенные выражения (4.19a) и (4.19б) дают соответственно:

$$\alpha = \frac{1}{1 + \frac{1}{2} \left(\frac{w}{L} \right)^2}; \quad (4.21a)$$

$$\alpha \approx \gamma \left[1 - \frac{1}{2} \left(\frac{w}{L} \right)^2 \right]. \quad (4.21b)$$

Несмотря на ряд ограничений, общей тенденцией при разработке транзисторов является приближение коэффициента передачи α к единице, поскольку при этом, как увидим ниже, улучшаются усилительные свойства транзистора.

Объемное сопротивление базы. Идеализируем структуру бездрейфового транзистора так, как показано на рис. 4.14. Как видим, базовый ток (если считать, что он протекает от центра базы к периферии) встречает различные сопротивления на трех разных участках. Первый — внутренний — участок (активная область базы) является диском.

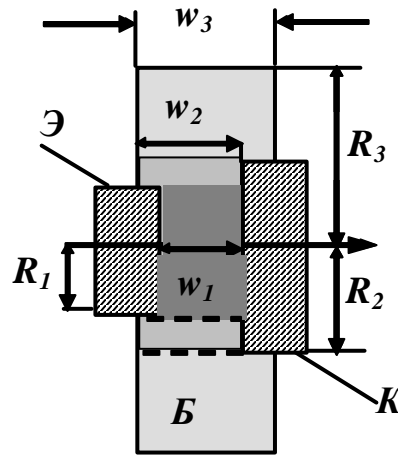


Рис. 4.14 — Идеализированная структура транзистора

Сопротивления колец в радиальном направлении выражаются формулами:

$$r_{\delta 2} = \frac{\rho_{\delta}}{2\pi w_2} \ln \frac{R_2}{R_1};$$

$$r_{\delta 3} = \frac{\rho_{\delta}}{2\pi w_3} \ln \frac{R_3}{R_2},$$

где ρ_B — удельное сопротивление базы. Сопротивление диска равно:

$$I_{\sigma}(R) = I_{\sigma} \frac{R^2}{R_1^2}.$$

4.5 Динамические параметры транзистора

Барьерные емкости. Эти емкости не имеют особой специфики по сравнению с емкостями $p-n$ перехода (диода) и могут быть рассчитаны по формулам (2.82). При этом, как уже отмечалось, формула (2.82б) оказывается неточной при сколько-нибудь значительных прямых смещениях. Учитывая, что при прямом смещении на переходе барьерная емкость во много раз меньше диффузионной, ее влиянием можно пренебречь. Относительно емкости C_k нужно заметить, что она представляет собой емкость коллектора, соответствующую активной области базы. Емкости коллектора, соответствующие промежуточной и пассивной областям базы, могут считаться соединенными параллельно и подключенными между внешними электродами коллектор — база. Эта межэлектродная емкость часто сравнима с величиной C_k , однако при расчетах ее можно относить к нагрузке транзистора. Необходимо отметить, что такое утверждение не всегда справедливо и зависит от конкретных схемных решений. Так, если в цепи базы транзистора имеется внешнее сопротивление, то эти емкости нельзя считать включенными параллельно нагрузке.

Особый интерес при анализе переходных и частотных характеристик транзистора представляет коэффициент передачи тока α . Поскольку согласно (4.13) коэффициент α состоит из двух множителей — коэффициента инжекции и коэффициента переноса, рассмотрим эти множители отдельно.

Коэффициент инжекции. Напомним, что ток, заряжающий барьерную емкость C_{Σ} , образуется основными носителями и поэтому не передается в цепь коллектора. Значит, увеличение емкостного тока на высоких частотах равносильно уменьшению коэффициента инжекции:

$$\gamma = \gamma \frac{-jX_{CЭ}}{-jX_{CЭ} + r_{Э}} = \frac{\gamma}{1 + j \frac{w}{w_{\gamma}}}. \quad (4.22a)$$

Соответственно изображение коэффициента инжекции имеет вид

$$\gamma = \frac{\gamma}{1 + p\tau_{\gamma}}. \quad (4.22б)$$

В формулах (4.30) постоянная времени τ_{γ} определяется выражением

$$\tau_{\gamma} = \frac{1}{w_{\gamma}} = r_{Э} C_{Э}. \quad (4.23)$$

Как видим, частотные и переходные свойства коэффициента γ улучшаются с увеличением рабочего тока, так как при этом уменьшается сопротивление $r_{Э}$, а увеличение барьерной емкости эмиттерного перехода происходит медленнее.

У дрейфовых транзисторов, характерных большими рабочими токами и меньшими емкостями, частота достигает сотен и тысяч мегагерц, и частотные свойства коэффициента передачи тока зависят в основном от коэффициента переноса, поэтому перейдем к рассмотрению α .

Коэффициент передачи тока. В общем случае динамические характеристики коэффициента α определяются коэффициентом переноса и коэффициентом инжекции

$$\alpha(p) = \gamma \operatorname{sech}\left(\frac{w}{L}\right) \sqrt{1 + p\tau}. \quad (4.24a)$$

Используя приближение, можно записать:

$$\alpha(p) = \frac{\alpha}{1 + p\tau_{\alpha}}, \quad (4.24б)$$

где

$$\tau_{\alpha} = \tau_{\chi} = \chi t_D \approx \frac{w^2}{2D}; \quad (4.25a)$$

$$\tau_{\alpha} = C_{Э.д.} r_{Э.д.} \quad (2.25б)$$

Оригиналом изображения (4.24б) является экспоненциальная функция

$$\alpha(t) = \alpha(1 - e^{-t/\tau_\alpha}). \quad (4.26)$$

Частотная характеристика, соответствующая изображению (4.24), получается заменой оператора p на jw и имеет вид

$$\alpha = \frac{\alpha}{1 + J \frac{w}{w_\alpha}}. \quad (4.27)$$

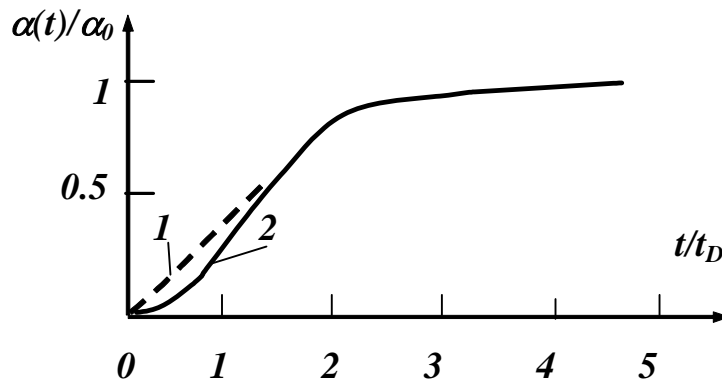


Рис. 4.15 — Переходные характеристики коэффициента передачи: 1 — первое приближение; 2 — строгое решение

Амплитудно-частотная и фазочастотная характеристики легко получаются из (4.27):

$$|\alpha(w)| = \frac{\alpha}{\sqrt{1 + (w/w_\alpha)^2}}; \quad (4.28a)$$

$$\varphi = -\text{arctg}(w/w_\alpha). \quad (4.28b)$$

Граничная частота (на уровне 0,7) равна:

$$w_\alpha = \frac{1}{\tau_\alpha} \approx \frac{2D}{w^2}. \quad (4.29)$$

Изображение (4.24) и связанные с ним функции (4.25) — (4.27) широко используются при анализе схем, если задержка фронта и сдвиг фазы не играют первоочередной роли.

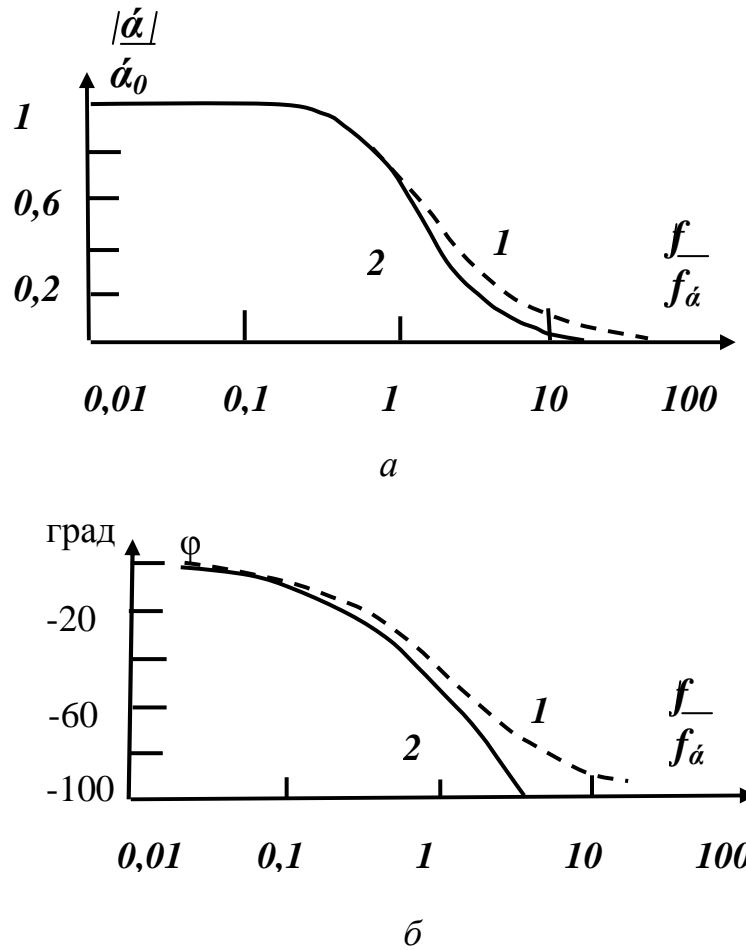


Рис. 4.16 — Частотные характеристики коэффициента передачи:
 а — амплитудно-частотные; б — фазочастотные.
 Кривые: 1 — приближенное решение; 2 — строгое решение

Диффузионные емкости. Понятие диффузионной емкости было введено при рассмотрении диодов как параметр, характеризующий зависимость приращения заряда в базе от приращения напряжения на переходе. Соответственно в транзисторах различают две диффузионные емкости — эмиттерную и коллекторную. Будем считать распределение дырок в базе линейным (рис. 4.17), что имеет место при условии $w < L$. Тогда неравновесный заряд дырок будет выражаться формулой (2.47а), если ток I заменить на $I_{\mathcal{E}}$.

$$\Delta Q = I_{\mathcal{E}} \frac{w^2}{2D} = I_{\mathcal{E}} t_D. \quad (4.30)$$

Дифференцируя (4.30) по напряжению $U_{\mathcal{E}}$ и учитывая (4.22), получаем диффузионную емкость эмиттера:

$$C_{ЭД} = \frac{t_D}{r_{Э}} = I \frac{t_D}{\Phi_T}. \quad (4.31)$$

Дифференцируя (4.30) по напряжению U_K (при $I_{Э} = const$) и используя выражения (4.23) и (4.24), можно привести диффузионную емкость коллектора к следующему виду:

$$C_{КД} = \sqrt{\frac{\xi_0 \xi}{2qN_D}} \frac{I_{Э} w}{D \sqrt{U_K}} = \frac{\tau}{r_K}. \quad (4.32)$$

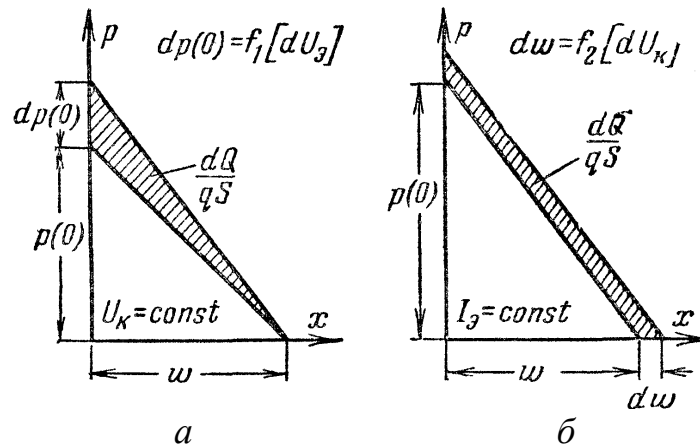


Рис. 4.17 — Изменение заряда в базе, приводящее к образованию диффузионных емкостей эмиттерного (а) и коллекторного (б) переходов

Диффузионная емкость коллектора $C_{КД}$ играет относительно малую роль, поскольку ее величина значительно меньше барьерной емкости C_K . Поэтому при анализе эквивалентных схем диффузионная емкость коллекторного перехода не учитывается. Диффузионная емкость эмиттерного перехода не изображается на эквивалентных схемах, а учитывается в коэффициенте передачи тока эмиттера.

Постоянная времени базы

До сих пор сопротивление базы мы не учитывали, т.е. считали равным нулю. U_K считалось приложенным непосредственно к коллекторному переходу. Ток генератора $\alpha I_{Э}$ (см. рис. 4.13) не полностью идет во внешнюю цепь: часть тока ответвляется во внутреннюю цепь $r_K C_K$. В области низких и средних частот влиянием емкости коллекторного перехода можно пренебречь. В этом

случае величина коллекторного тока, I_K строго говоря, зависит от дифференциального сопротивления коллекторного перехода r_K , объемного сопротивления базы r_B и внешнего сопротивления нагрузки. Если транзистор работает в активном режиме, влиянием сопротивлений r_K и r_B на ток коллектора можно пренебречь. Таким образом, в реальном транзисторе при заданном токе $I_{\mathcal{E}}$ ток I_K зависит не только от коэффициента α , но и от соотношения сопротивлений коллекторного перехода, объемного сопротивления базы и внешнего сопротивления нагрузки. В области высоких частот (малых времен) барьерная емкость коллекторного перехода оказывает существенное влияние на величину коллекторного тока вследствие уменьшения реактивного сопротивления C_K , а это приводит к тому, что часть коллекторного тока ответвляется во входную цепь. Величина коллекторного тока уменьшается. Ток коллектора будет связан с током эмиттера очевидным соотношением

$$I_K = \alpha(p)I_{\mathcal{E}} \frac{X_C(p)}{X_C(p) + r_B} = I_{\mathcal{E}} \frac{\alpha(p)}{1 + p\tau_{\delta}}, \quad (4.33)$$

где

$$\tau_{\delta} = r_B C_K. \quad (4.34)$$

Выражение (4.33) справедливо при выполнении условия: тока короткого замыкания на выходе.

Как правило, транзистор является элементом некоторого устройства, например, усилительного каскада и в цепи коллектора включено сопротивление нагрузки. В этом случае емкость C_K является элементом цепи, которая создает отрицательную обратную связь, т.е. часть выходного тока в противофазе ответвляется на вход, что приводит к уменьшению тока коллектора на высоких частотах (малых временах). Таким образом, в области высоких частот выходной ток I_K уменьшается как за счет инерционных свойств коэффициента передачи тока эмиттера, так и паразитной емкости коллекторного перехода, которая создает отрицательную обратную связь.

4.6 Зависимость параметров транзистора от режима и температуры

Зависимость от параметров транзистора режима. Параметры транзистора зависят от рабочей точки. Будем считать величинами, определяющими режим транзистора по постоянному току: ток $I_{\mathcal{E}}$ и напряжение U_K .

Рассмотрим, как могут меняться параметры транзистора в зависимости от этих величин. Коэффициент передачи тока α согласно (4.20) зависит от напряжения U_K из-за модуляции толщины базы. Чем больше (по модулю) коллекторное напряжение, тем уже база и тем ближе к единице коэффициент переноса дырок. Следовательно, коэффициент α увеличивается с ростом U_K . Вторым фактором, приводящим к зависимости $\alpha(U_K)$, является ударная ионизация в коллекторном переходе. Поскольку относительные изменения α невелики и их трудно отобразить на графике, на рис. 4.18, б показана кривая $1/(1-\alpha)$, которая более наглядно иллюстрирует эти изменения, сохраняя качественный характер зависимостей. Зависимость α от тока эмиттера (рис. 4.18, б) обусловлена главным образом изменением коэффициента инжекции.

Уменьшение коэффициента α в области малых токов объясняется двумя причинами. Во-первых, уменьшается эквивалентный коэффициент диффузии. Во-вторых, возрастает влияние тока рекомбинации в области эмиттерного перехода. Ток рекомбинации обусловлен уходом электронов из базы, общий коэффициент инжекции уменьшается, что приводит к уменьшению коэффициента передачи тока эмиттера. Первая причина приводит к уменьшению α всего на единицы процентов, т.е. не оказывает принципиального влияния на коэффициент передачи тока эмиттера. Вторая причина, которая особенно сильно выражена у кремниевых транзисторов, может привести к значительному уменьшению α . В настоящее время, особенно в связи с разработкой маломощных интегральных схем, поведение транзисторов в области малых токов и напряжений привлекает особое внимание.

В области микротоков можно величину $1/(1-\alpha)$ (рис. 4.18), которая более наглядно характеризует изменения α , представить в виде:

$$\frac{1}{1-\alpha} = 1 + \frac{1}{b} \sqrt{I_{\text{Э}}}; \quad b = \frac{\varphi_T}{\tau} \sqrt{\frac{2\zeta_0 \zeta_w S}{\Delta\varphi_0 D}},$$

где τ — время жизни носителей.

Приведенные соображения подтверждают и конкретизируют спад α с уменьшением тока, а также иллюстрируют зависимость этого спада от ряда факторов.

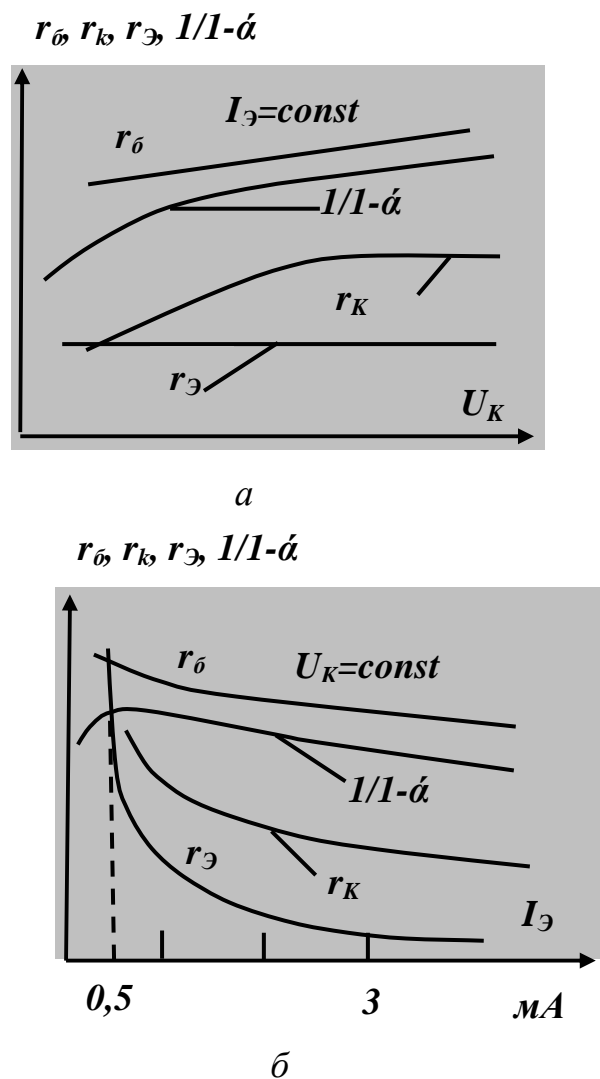


Рис. 4.18 — Зависимость статических параметров транзистора от режима:
 а — при постоянном токе эмиттера,
 б — при постоянном коллекторном напряжении

В первую очередь следует отметить прямую зависимость α от времени жизни τ и тем самым от свойств и состояния поверхности в области перехода. Спад коэффициента передачи в области больших и малых токов приводит к наличию максимума на кривой α , который имеет место при небольшом токе.

Зависимость от температуры. Параметры транзистора зависят от температуры при неизменной рабочей точке ($I_{\text{Э}} = \text{const}$ и $U_{\text{К}} = \text{const}$). Коэффициент передачи α согласно (4.20) и (2.35) зависит от температуры через параметры τ , D и $\rho_{\text{Б}}$. Из этих параметров главную роль играет время жизни, которое существенно возрастает с температурой.

Поэтому коэффициент α растет при нагревании транзистора и уменьшается при его охлаждении. На рис. 4.19 для большей ясности показана температурная зависимость величины $1/(1-\alpha)$. Сопротивление эмиттерного перехода $r_{\text{Э}}$ согласно (4.22) линейно зависит от температуры. Легко показать, что величина $r_{\text{Э}}$ меняется приблизительно на 0,33 %/град. Сопротивление коллекторного перехода согласно (4.24) зависит от температуры в основном через диффузионную длину L (т.е. через время жизни) и должно увеличиваться при нагреве транзистора.

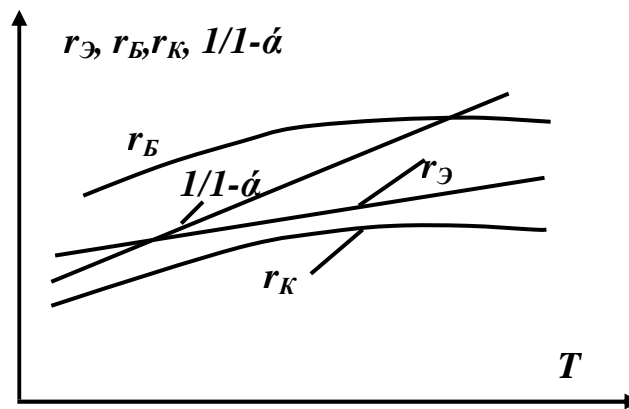


Рис. 4.19 — Зависимость статических параметров транзистора от температуры

Объемное сопротивление базы $r_{\text{Б}}$ меняется с температурой в результате изменения удельного сопротивления материала.

Основное значение для стабильности работы транзисторных схем имеет температурная зависимость тепловых токов коллек-

торного и эмиттерного переходов. Будучи небольшими при комнатной температуре, они сильно возрастают при нагреве транзистора, а это, как видно из рис. 4.11, *a*, смещает все кривые коллекторного семейства характеристик. Характеристики для повышенной температуры на рисунке показаны пунктиром. В результате получается косвенная температурная зависимость параметров. Подробно зависимость теплового тока от температуры анализировалась ранее.

Необходимо отметить, что нестабильность параметров транзистора от режимов и особенно от температуры не является препятствием для создания высокостабильных транзисторных устройств. Применение дополнительных элементов, например термисторов (резисторы, сопротивление которых изменяется с изменением температуры), введение отрицательных обратных связей, ряд других методов позволяют создавать высокостабильные устройства и большие системы. Зависимость тепловых токов от температуры позволяет использовать транзистор в качестве датчика в электронных термометрах.

4.7 Характеристики и параметры транзистора при включении с общим эмиттером

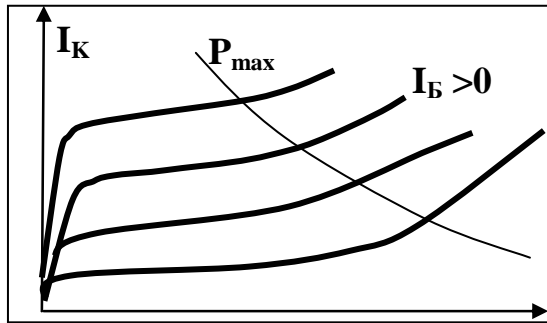
Статические характеристики и параметры. При включении транзистора по схеме ОЭ (рис. 4.4, *б*) входным током является ток базы, который и примем за параметр коллекторного семейства характеристик. Статические вольт-амперные характеристики показаны на рис. 4.20, *a* и *б*, эквивалентная схема на рис. 4.21.

Главные отличительные черты включения ОЭ, вытекающие из сравнения рис. 4.21 и 4.11, сводятся к следующему.

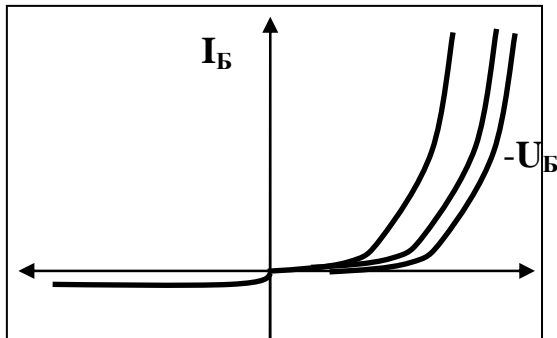
1. Кривые коллекторного семейства не пересекают ось ординат и полностью расположены в I квадранте.

2. Кривые коллекторного семейства менее регулярны, чем в схеме ОБ, они имеют гораздо больший и неодинаковый наклон и заметно сгущаются при больших токах. **Ток при оборванной базе (когда $I_B = 0$)** намного больше тока I_{K0} , имеющего место при включении транзистора в схеме с общей базой. **Это связано с тем, что в этом случае тепловой ток коллекторного перехода**

протекает через эмиттерный переход и изменяет величину инжектированных носителей из эмиттера. Если сопротивление в цепи базы сделать равным нулю, тепловые токи для схем включения с ОБ и ОЭ практически равны.



а



б

Рис. 4.20 — Статические вольт-амперные характеристики транзистора ОЭ: *а* — выходные статические вольт-амперные характеристики, *б* — входные статические вольт-амперные характеристики

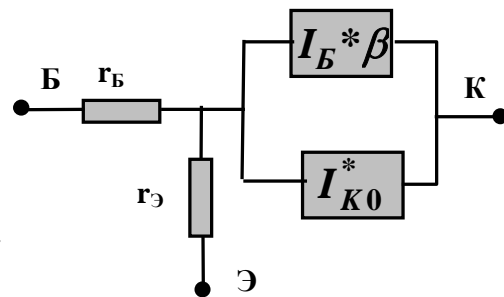


Рис. 4.21 — Эквивалентная схема транзистора

В аналитическом виде семейство коллекторных характеристик **ОЭ** при работе транзистора в активном режиме ($U_K < 0$) получается из выражения (4.10), если в правую часть подставить очевидное соотношение $I_Э = I_K + I_б$ и выразить ток коллектора через ток базы:

$$I_K = \frac{\alpha}{1-\alpha} I_б + \frac{I_{K0}}{1-\alpha} + \frac{U_K}{r_K(1-\alpha)}.$$

Коэффициент при токе I_B является интегральным коэффициентом передачи базового тока. Для этого важнейшего параметра введем специальное обозначение:

$$\beta = \frac{\alpha}{1 - \alpha}. \quad (4.35)$$

Тогда ток коллектора можно записать в следующей форме,

$$I_K = \beta I_{\bar{o}} + I_{K0}^* + \frac{U_K}{r_K^*}, \quad (4.36)$$

где

$$I_{K0}^* = (1 + \beta) I_{K0}; \quad (4.37)$$

$$r_K^* = \frac{r_k}{1 + \beta}. \quad (4.38)$$

Часто последним членом в (4.36) пренебрегают; тогда получается аналог выражения (4.10)

$$I_K = \beta I_{\bar{o}} + I_{K0}^*. \quad (4.39)$$

Величина β , входящая в формулы (4.35) и (4.39), является интегральной. Из формулы (4.39) легко получить интегральный коэффициент передачи

$$\beta_{ИН} = \frac{I_K - I_{K0}^*}{I_{\bar{o}} + I_{K0}^*}. \quad (4.40a)$$

Дифференциальный коэффициент передачи определяется по аналогии с (4.11 а) как

$$\beta = \frac{dI_K}{dI_{\bar{o}}} [U_K = const. \quad (4.40b)$$

Интегральный и дифференциальный коэффициенты передачи тока базы связаны формулой:

$$\beta = \beta_{ин} + (I_{\bar{o}} + I_{K0}^*) \frac{d\beta}{dI_{\bar{o}}}. \quad (4.41)$$

Динамические параметры. При включении транзистора по схеме ОЭ частотные и временные зависимости свойственны не только коэффициенту, β но и коллекторному сопротивлению, которое согласно (4.36) зависит от β .

Для выяснения инерционных свойств коэффициента передачи тока базы подставим в формулу (4.35) изображение $\alpha(p)$ из

(4.22) и комплексную величину $\alpha(j\omega)$ из (4.27). Тогда после несложных преобразований получим:

$$\beta(p) = \frac{\beta}{1 + p\tau_\beta}; \quad (4.42a)$$

$$\beta(\omega) = \frac{\beta}{1 + j\frac{\omega}{\omega_\beta}}; \quad (4.42б)$$

$$\tau_\beta = \frac{\tau_\alpha}{1 - \alpha} = (1 + \beta)\tau_\alpha; \quad (4.43a)$$

$$\omega_\beta = (1 - \alpha)\omega_\alpha = \frac{\omega_\alpha}{1 + \beta}. \quad (4.43б)$$

Эквивалентная схема с ОЭ для высоких частот приведена на рис. 4.24.

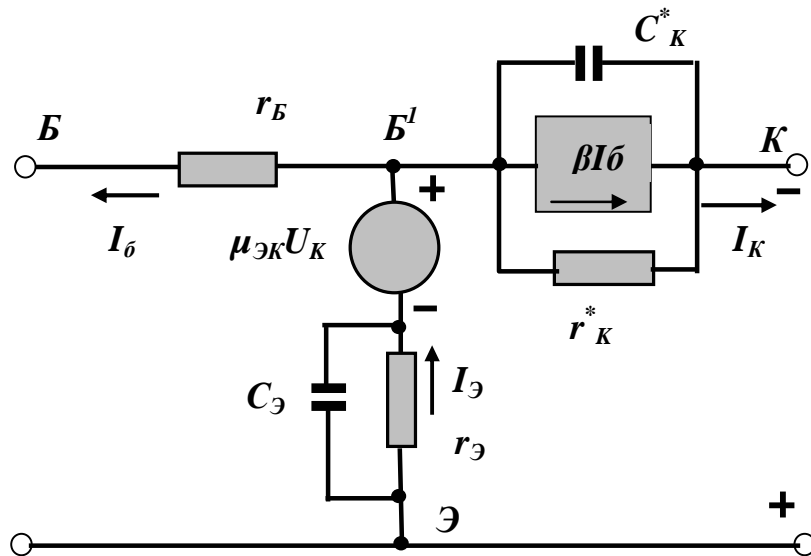


Рис. 4.22 — Эквивалентная схема ОЭ для высоких частот

В принципе, эквивалентную схему с общим эмиттером можно получить из схемы с общей базой, применяя правила преобразования одной схемы в другую.

На рис. 4.23 приведена зависимость коэффициента передачи тока базы $\beta(f)$ от частоты. Пунктирной линией показана зависимость $\beta(f)$, вычисленная по формуле (4.42б). Использование линейной аппроксимации позволяет просто вычислить коэффи-

циент передачи тока базы на любой частоте. В справочниках приводится модуль коэффициента передачи тока базы, измеренный на высокой частоте $f_{ИЗМ}$. Частота, на которой коэффициент передачи $|\beta| = 1$, называется частотой среза.

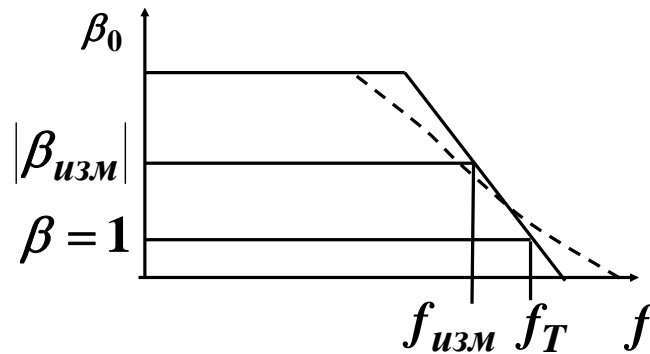


Рис. 4.23 — Зависимость коэффициента передачи тока базы от частоты

Схема с общим коллектором.

Эта схема (см. рис. 4.4, в) имеет много общего со схемой ОЭ, потому что в обеих управляющим током является ток базы, а выходные токи ($I_{Э}$ или $I_{К}$) различаются незначительно.

Поэтому семейство выходных характеристик будет практически таким же, как на рис. 4.20, а, если ток коллектора заменить током эмиттера. Что касается входного семейства, то оно представляется кривыми, показанными на рис. 4.20, б, но сдвинутыми вправо на величину напряжения на коллекторном переходе, которое в данном случае играет роль входного напряжения.

Переходные и частотные свойства схемы ОК почти совпадают со свойствами схемы ОЭ, так как они определяются коэффициентом передачи

$$\frac{dI_{Э}}{dI_{б}} = 1 + \beta,$$

мало отличающимся от β .

В заключение главы рассмотрим параметры усилительных каскадов при различных схемах включения транзисторов. Будем считать выходные статические вольт-амперные характеристики

для всех схем включения одинаковыми Анализ схем проведем для области низких частот.

На рис. 4.18 приведены принципиальные схемы усилительных каскадов при включении транзисторов с ОЭ, ОК, ОБ соответственно. На рис. 4.19 — эквивалентные схемы в области низких частот для каждого усилительного каскада. Чаше под сопротивлением $R_{Г}$ имеется в виду выходное сопротивление предыдущего каскада.

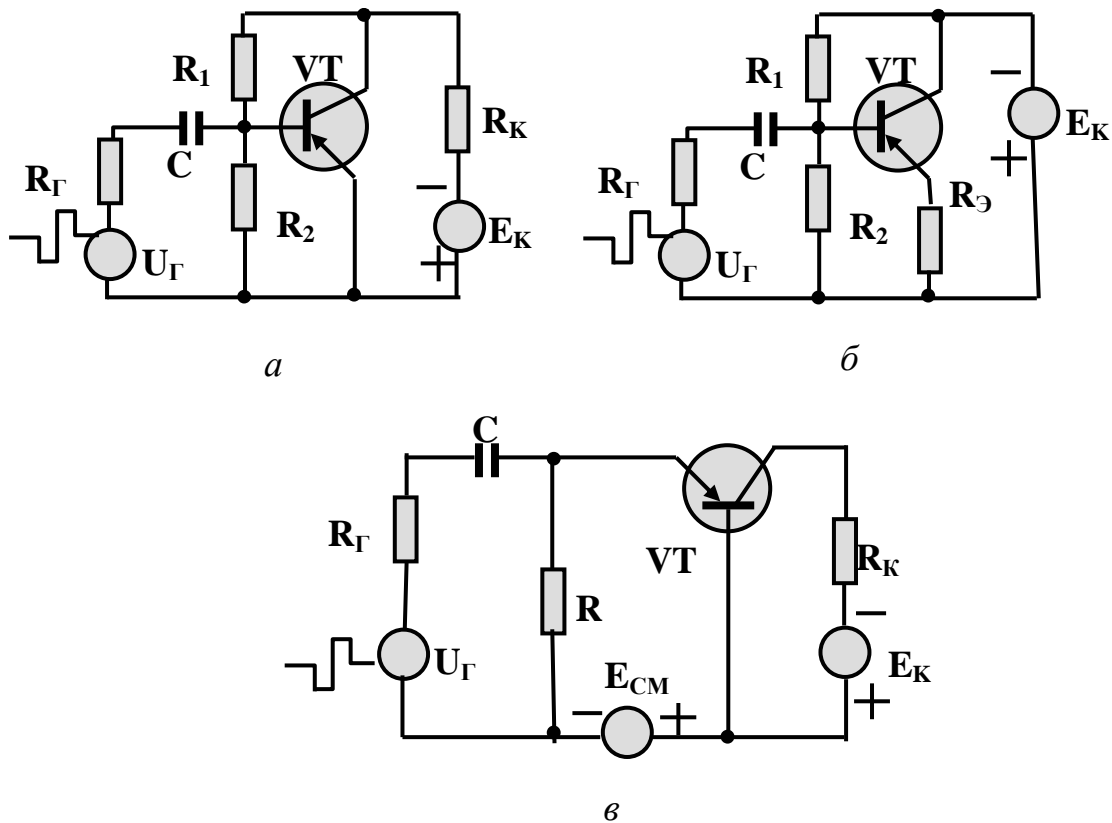


Рис. 4.24 — Принципиальные схемы усилительных каскадов при включении транзисторов ОЭ, ОК, ОБ соответственно

Схема с общей базой

Входное сопротивление каскада, если учесть, что влиянием дифференциального сопротивления коллекторного перехода можно пренебречь, запишется в виде:

$$R_{BX} = \frac{I_{\text{Э}} r_{\text{Э}} + I_{\text{Б}} r_{\text{Б}}}{I_{\text{Э}}}. \quad (4.44)$$

Ток базы связан с током эмиттера $I_{\mathcal{E}} = \frac{I_B}{1-\alpha}$. Подставив значение тока базы в формулу (4.44), получим

$$R_{BX} = r_{\mathcal{E}} + r_B(1-\alpha). \quad (4.45)$$

Выходное сопротивление каскада:

$$R_{ВЫХ} = r_K. \quad (4.46)$$

Коэффициент усиления по току можно записать как отношение выходного тока к входному

$$K_I = \frac{\alpha I_{\mathcal{E}}}{I_B} = \alpha. \quad (4.47)$$

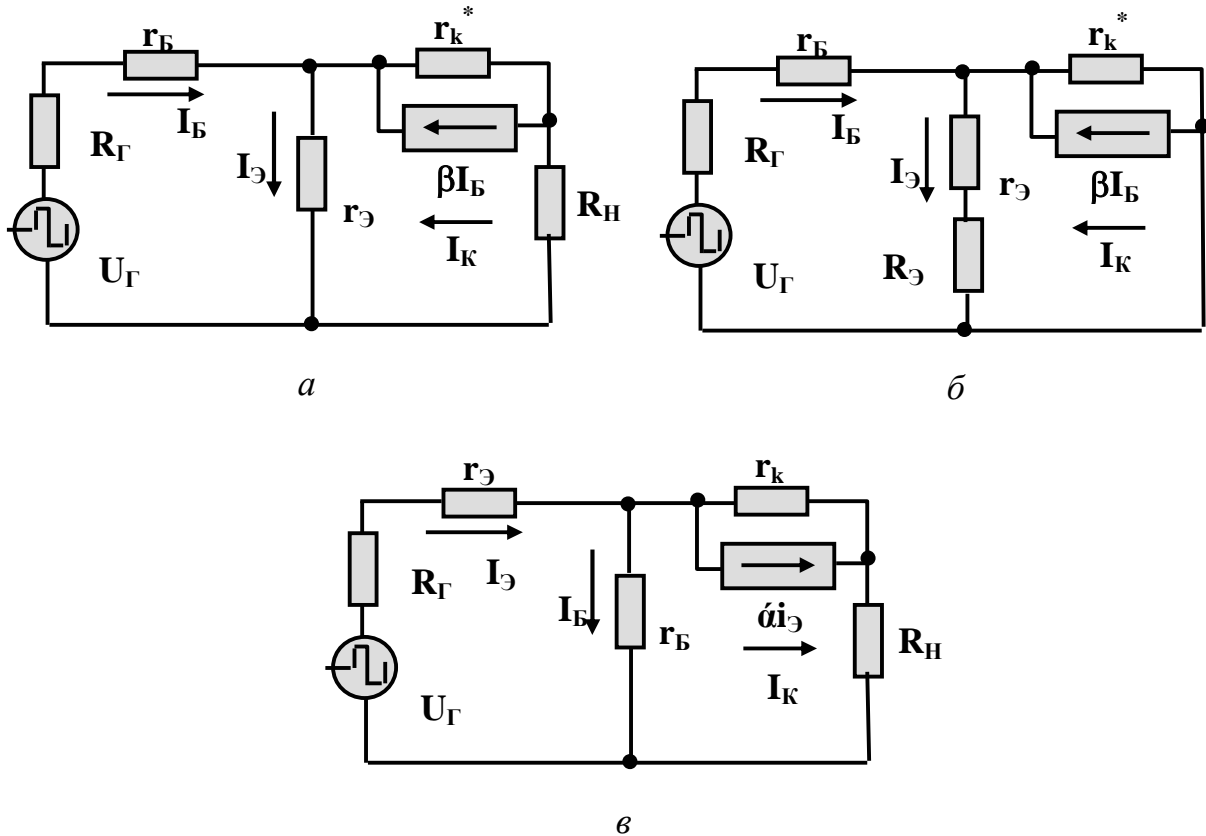


Рис. 4.25 — Эквивалентные схемы усилительных каскадов для области низких частот при включении транзисторов с ОЭ, ОК и ОБ

Коэффициент усиления по напряжению:

$$K_U = \frac{U_{ВЫХ}}{U_{\Gamma}} = \frac{\alpha R_H I_{\mathcal{E}}}{I_{\mathcal{E}}(r_{\mathcal{E}} + R_{\Gamma} + r_B(1-\alpha))} = \frac{\alpha R_H}{r_{\mathcal{E}} + R_{\Gamma} + r_B(1-\alpha)}. \quad (4.48)$$

Если в формуле 4.48 выполняется условие

$R_{\Gamma} \gg r_{\mathcal{E}} + r_B(1 - \alpha)$ с учетом, что $\alpha \approx 1$, выражение для коэффициента передачи по напряжению запишется в виде: $K_U = \frac{R_H}{R_{\Gamma}}$, т.е.

не зависит от параметров транзистора и его режима. Однако при этом следует помнить, что для обеспечения усиления по мощности необходимо выполнять условие $R_H > R_{\Gamma}$. Таким образом, увеличение сопротивления R_{Γ} повышает стабильность коэффициента усиления, но снижает усиление мощности входного сигнала.

Схема с общим эмиттером

Входное и выходное сопротивления, учитывая допущение которое мы сделали выше, равны

$$R_{BX} = \frac{U_{\Gamma}}{I_B} = \frac{I_B r_B + I_{\mathcal{E}} r_{\mathcal{E}}}{I_B} = r_B + r_{\mathcal{E}}(1 + \beta); \quad (4.49)$$

$$R_{BIX} = r_K^*. \quad (4.50)$$

Коэффициенты усиления по току и напряжению соответственно равны

$$K_I = \beta; \quad (4.51)$$

$$K_U = \frac{\beta R_H}{r_{\mathcal{E}}(1 + \beta) + r_B + R_{\Gamma}} = \frac{\alpha R_H}{r_{\mathcal{E}} + (r_B + R_{\Gamma})(1 - \alpha)}. \quad (4.52)$$

Из формулы 4.52 видно, что сопротивление R_{Γ} при включении транзистора с ОЭ не обеспечивает высокой стабильности коэффициента усиления по напряжению.

Если в цепи эмиттера включить дополнительный резистор $R_{\mathcal{E}}$, коэффициент усиления по напряжению можно записать в виде:

$$K_U = \frac{\alpha R_H}{R_{\mathcal{E}} + r_{\mathcal{E}} + (r_B + R_{\Gamma})(1 - \alpha)}.$$

Если выполнить условие $R_{\mathcal{E}} \gg r_{\mathcal{E}} + (r_B + R_{\Gamma})(1 - \alpha)$, можно записать: $K_U \approx \frac{R_H}{R_{\mathcal{E}}}$, т.е. коэффициент усиления не зависит от параметров транзистора.

Схема с общим коллектором

Входное и выходное сопротивления, учитывая допущение которое мы сделали выше, равны

$$R_{BX} = \frac{I_B r_B + I_{\mathcal{E}}(r_{\mathcal{E}} + R_{\mathcal{E}})}{I_B} = r_B + (r_{\mathcal{E}} + R_{\mathcal{E}})(1 + \beta); \quad (4.53)$$

$$R_{BIX} = r_K^*. \quad (4.54)$$

Коэффициенты усиления по току и напряжению соответственно равны

$$K_I = \beta + 1; \quad (4.55)$$

$$K_U = \frac{\beta R_{\mathcal{E}}}{(r_{\mathcal{E}} + R_{\mathcal{E}})(1 + \beta) + r_B} = \frac{\alpha R_{\mathcal{E}}}{r_{\mathcal{E}} + R_{\mathcal{E}} + r_B(1 - \alpha)}. \quad (4.56)$$

Анализ усилительных схем при различном включении транзисторов показывает (область низких и средних частот):

1. Схема с ОБ имеет самое низкое входное сопротивление. Самое высокое входное сопротивление имеет схема с ОК. Однако следует иметь в виду, что, если в схеме с ОЭ в цепи эмиттера включить сопротивление $R_{\mathcal{E}}$, как в схеме с ОК, входные сопротивления для обоих каскадов будут равны.

2. Самое низкое выходное сопротивление имеет схема с ОК. Как правило, оно лежит в пределах десятков Ом. Благодаря низкому выходному сопротивлению каскад с ОК (эмиттерный повторитель) очень часто используется при работе на низкоомную нагрузку. Два других каскада имеют высокое выходное сопротивление — десятки и сотни кОм, т.е. можно считать, что выходной ток не зависит от величины сопротивления коллекторного резистора (режим генератора тока).

3. Все рассмотренные схемы принципиально могут обеспечить усиление по мощности, причем схема с ОЭ обеспечивает максимальное усиление по мощности на постоянном токе.

4. Очень часто на основании п.3 делают ошибочный вывод, что схема с ОЭ является более предпочтительной по отношению к двум другим. Мы анализировали область низких частот, что является частным случаем.

Для современной радиоэлектроники наиболее характерным является работа устройства в широкой полосе частот, и тогда все три схемы с точки зрения усиления по мощности становятся равноценными.

4.8 Разновидности эквивалентных схем

Т-образные эквивалентные схемы, которые мы рассмотрели и которыми будем пользоваться в дальнейшем, не являются единственно возможными. В литературе можно встретить ряд других схем, из которых чаще всего встречается П-образная эквивалентная схема. Кроме того, нередко (особенно в справочной литературе) транзистор рассматривается как бесструктурный четырехполюсник с той или иной системой параметров.

П-образные эквивалентные схемы. Основная П-образная схема для включения ОЭ показана на рис. 4.25, а, а ее вариант, в котором выделено сопротивление базы, на рис. 4.25, б. Как видим, в обеих схемах используются проводимости, комплексные (Y) или активные (g), а в качестве усилительного параметра используется крутизна S . Остановимся подробнее на «гибридной» П-образной схеме (рис. 4.25, б) как более распространенной и более специфичной для транзисторов. Установим связь между параметрами П- и Т-образной схем (рис. 4.25, б и 4.24). Для этого сначала рассмотрим область низких частот, пренебрегая емкостями и частотной зависимостью параметров.

Кроме того, исключим из рассмотрения сопротивление r_B , одинаковое для обеих схем. Тогда в каждой схеме останется по четыре параметра: $g_{BЭ}^1$, $g_{BЛ}^1$, $g_{BЭ}$, S и $r_Э$, $\mu_{ЭК}$, r_K^* , β .

Чтобы выразить одни параметры через другие, необходимы четыре уравнения. Такие уравнения получаются, если, например, приравнять входные (базовые) и выходные (коллекторные) токи в обеих схемах при заданном входном напряжении (между точками Б' и Э) и коротком замыкании на выходе, а затем приравнять выходные (коллекторные) токи и входные (базовые) напряжения при заданном выходном напряжении и холостом ходе на входе. Точные решения несколько громоздки, поэтому имеет смысл принять некоторые несущественные ограничения, а именно: $r_Э < r_K$; $\mu_{ЭК} < 1$. То-

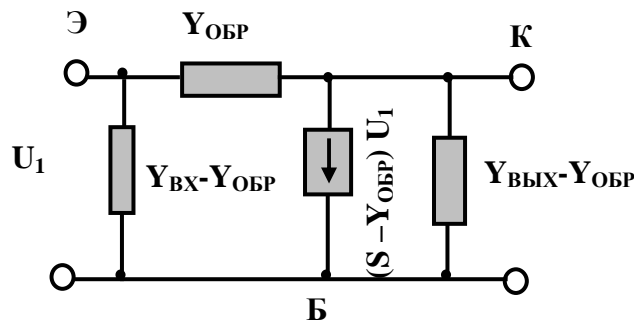
гда с учетом соотношения $2\mu_{\text{ЭК}}(\gamma - \alpha)r_K = r_{\text{Э}}$ связь между параметрами оказывается весьма простой:

$$s = \frac{\alpha}{r_{\text{Э}}}; \quad (4.57\text{а})$$

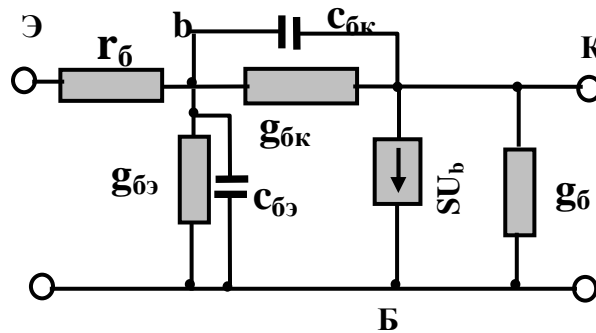
$$g_{\text{бЭ}} = \frac{1}{(1 + \beta)r_{\text{Э}}}; \quad (4.57\text{б})$$

$$g_{\text{бК}} = \frac{\mu_{\text{ЭК}}(1 - \alpha)}{r_{\text{Э}}} = \frac{1}{2r_K}; \quad (4.57\text{в})$$

$$g_{\text{КЭ}} = \frac{\mu_{\text{ЭК}}}{r_{\text{Э}}} = \frac{1}{2r_K^*}. \quad (4.57\text{г})$$



а



б

Рис. 4.25 — П-образные эквивалентные схемы ОЭ:
а — основная; б — модифицированная

Подобно тому, как внутренние параметры Т-образной схемы связаны соотношением $2\mu_{\text{ЭК}}(\gamma - \alpha)r_K = r_{\text{Э}}$, параметры П-образной схемы, как легко убедиться, связаны соотношением:

$$g_K g_{\text{бЭ}} = s g_{\text{бК}}. \quad (4.58)$$

Теперь рассмотрим область высших частот. Тогда методика, аналогичная предыдущей, приводит к следующим результатам

$$s(w) = \frac{\alpha(w)}{r_{\text{Э}}} = \frac{\alpha / r_{\text{Э}}}{1 + jw\tau_{\alpha}}; \quad (4.59a)$$

$$Y_{\text{бЭ}} \approx g_{\text{бЭ}}(1 + jw\tau_{\beta}); \quad (4.59б)$$

$$Y_{\text{бК}} = g_{\text{бК}}(1 + jw\tau_{\text{К}}); \quad (4.59в)$$

$$Y_{\text{КЭ}} \approx g_{\text{КЭ}} \frac{1 + jw\tau_{\text{К}}}{1 + jw\tau_{\beta}}. \quad (4.59г)$$

4.9 Составные транзисторы

Для того чтобы повысить значение коэффициента усиления β , нужно уменьшить толщину базы, что, конечно, представляет некоторые трудности технологического характера, а самое главное, снижает допустимое значение на коллекторном переходе. Повысить коэффициент передачи тока базы можно, соединяя определенным образом два транзистора, рассматривая два транзистора как единое целое. Такая комбинация (иногда выполняемая на одной пластине с внутренними соединениями и тремя внешними выводами) называется составным транзистором или схемой Дарлингтона (рис. 4.26, а). Покажем, что составной транзистор действительно имеет коэффициент β , значительно больший, чем у обоих его компонентов. Задавая приращение тока $dI_{\text{Б}} = dI_{\text{Б1}}$, получаем:

$$dI_{\text{Э1}} = (1 + \beta_1)dI_{\text{б}};$$

$$dI_{\text{К}} = dI_{\text{К1}} + dI_{\text{К2}} = \beta_1 dI_{\text{б}} + \beta_2 [(1 + \beta_1)dI_{\text{б}}].$$

Деля $dI_{\text{К}}$ на $dI_{\text{Б}}$, находим результирующий дифференциальный коэффициент передачи:

$$\beta_{\Sigma} = \beta_1 + \beta_2 + \beta_1\beta_2. \quad (4.60a)$$

Поскольку всегда $\beta \gg 1$, можно считать

$$\beta_{\Sigma} \approx \beta_1\beta_2. \quad (4.60б)$$

Легко видеть, что величина β_{Σ} может составлять несколько тысяч при использовании рядовых транзисторов. Для этого оба транзистора должны находиться в активном режиме.

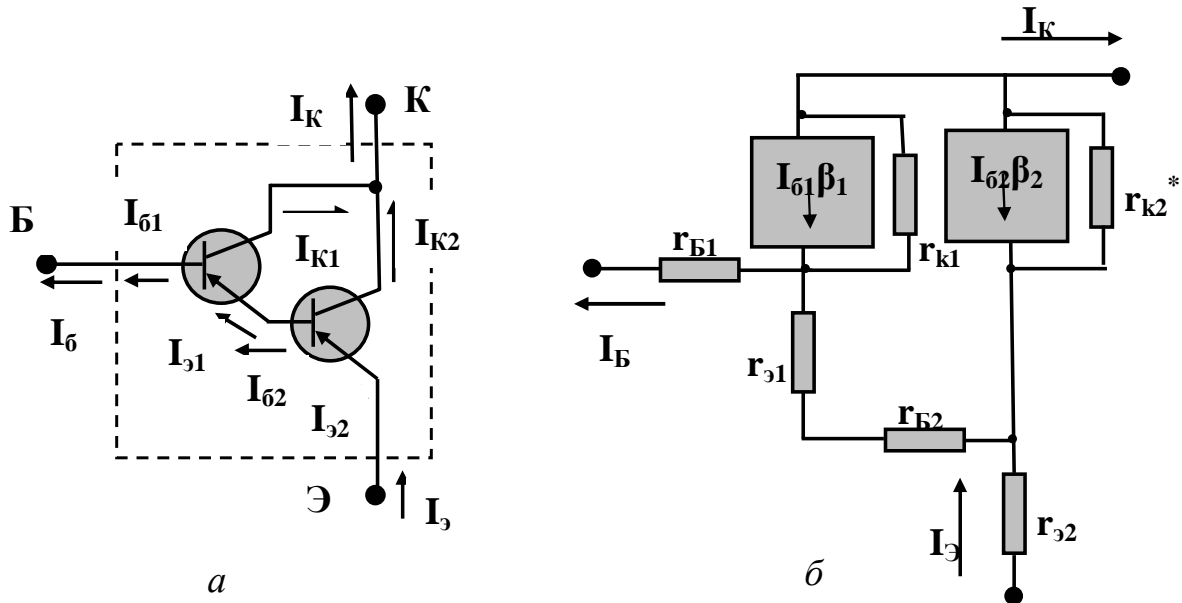


Рис. 4.26 — Составной транзистор (а) и его эквивалентная схема (б)

4.10 Допустимая мощность

Поскольку в активном режиме токи I_E и I_K почти одинаковы, а напряжение U_K значительно больше, чем U_E , то основная часть мощности потерь выделяется в области коллекторного перехода. Каждый транзистор характеризуется предельно допустимой температурой перехода, при превышении которой параметры резко ухудшаются. Исходя из соотношения (2.41) легко прийти к следующей зависимости между допустимой мощностью рассеяния, допустимой температурой перехода и температурой окружающей среды:

$$P_{дон} = \frac{T_{пер.дон} - T_{окр}}{R_t}, \quad (4.61)$$

где R_t — тепловое сопротивление переход — среда, которое, как и величина $T_{пер.дон}$, указывается в справочниках. Типичной величиной R_t для маломощных транзисторов является 0,5 — 0,7 град/МВт. Для мощных транзисторов эта величина в десятки раз меньше. Типичными значениями $T_{перех.дон}$ являются 90 — 100° С для германия и 150 — 200° С для кремния. Из формулы (4.119) следует, что допустимая мощность уменьшается с ростом окружающей

температуры и что главным путем повышения мощности является уменьшение величины R_T , т. е. улучшение теплоотвода. Мощные транзисторы характерны большими рабочими токами и соответственно большими площадями p - n переходов (до 1 см^2), оба фактора отражаются на величине основных параметров и придают мощным транзисторам определенную специфику. Так, при большой площади переходов трудно реализовать тонкую базу, особенно сплавным методом, а это приводит к сравнительно низкой граничной частоте. У мощных сплавных транзисторов частота не превышает $100 — 200 \text{ кГц}$, но даже у дрейфовых мощных транзисторов, у которых частота достигает $20 — 40 \text{ мГц}$, она все же ниже, чем у маломощных дрейфовых транзисторов, у которых в настоящее время достигнуты значения в несколько гигагерц.

Кроме того, коллекторная емкость у мощных транзисторов обычно составляет сотни, а иногда и тысячи пикофард, так что в целом мощные транзисторы являются сравнительно низкочастотными. Большие рабочие токи приводят к резкому уменьшению сопротивлений r_E и r_K . Из выражения (4.22) следует, что при токе больше 100 ма сопротивление эмиттерного перехода ничтожно мало, и с ним практически можно не считаться. Поэтому в схеме ОЭ входным сопротивлением будет по существу только сопротивление базы r_B , которое при высоких уровнях инжекции модулируется и обычно лежит в пределах до 10 Ом . Малая величина входного сопротивления не является препятствием для применения мощных транзисторов, если связь с источником сигнала осуществляется через трансформатор. Сопротивление коллекторного перехода при токах порядка 1 А составляет всего несколько килоом, а сопротивление r_K^* в схеме ОЭ — сотни Ом. Тепловой ток коллектора, пропорциональный площади перехода, доходит у мощных транзисторов до десятков миллиампер. С приближением напряжения к максимально допустимой величине тепловой ток увеличивается в несколько раз в связи с возрастающей ролью термогенерации и ударной ионизации в переходе, а также самогрева. Как известно, при больших эмиттерных токах наблюдается уменьшение коэффициентов передачи α и β .

У мощных транзисторов эта зависимость весьма существенна: коэффициент β в рабочем диапазоне токов может уменьшиться в несколько раз.

В связи с этим характеристики коллекторного семейства заметно сближаются в области больших токов, особенно в схеме ОЭ (рис. 4.27).

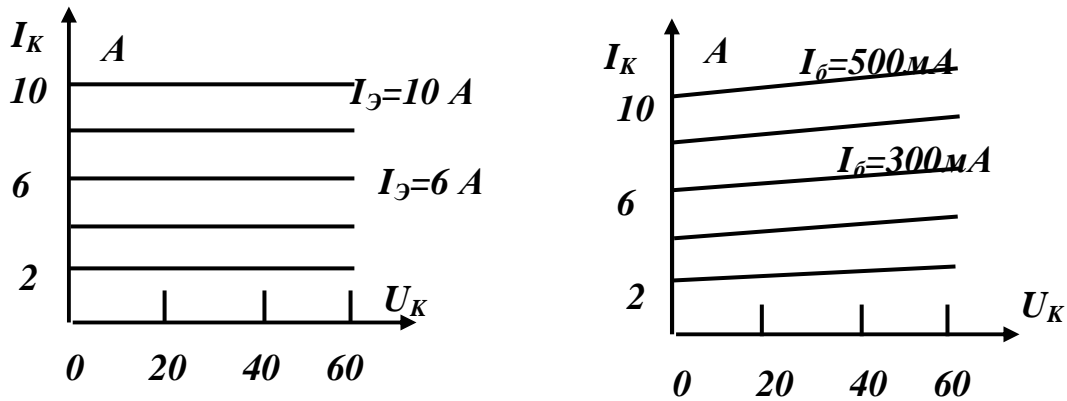


Рис. 4.27 — Коллекторные характеристики мощного транзистора:
а — в схеме ОБ; б — в схеме ОЭ

Результатом такой неоднородности характеристик являются, конечно, нелинейные искажения при усилении сигналов. Одной из главных проблем при конструировании и эксплуатации мощных транзисторов является обеспечение теплоотвода. Эта проблема решается путем непосредственного контакта коллектора с корпусом транзистора и применением внешних радиаторов.

В заключение отметим одно явление, которое характерно для мощных транзисторов в связи с большим диаметром эмиттерного перехода и малой длиной «пассивного» участка базы (между ее внешним выводом и эмиттером) (рис. 4.28). Это явление выражается в неодинаковой плотности тока на разных участках эмиттера, плотность тока максимальна на периферии эмиттерного перехода и минимальна в центре перехода. Причиной такой неравномерности является протекание базового тока в радиальном направлении — от центра к внешнему кольцевому электроду. В результате потенциал базы вдоль радиуса уменьшается, а напряжение эмиттер — база соответственно растет. У мало-мощных транзисторов неравномерность инжекции обычно несущ-

щественна, так как значительная часть падения напряжения приходится на «пассивный» участок базы, который у этих транзисторов значительно больше диаметра эмиттера по конструктивным соображениям.

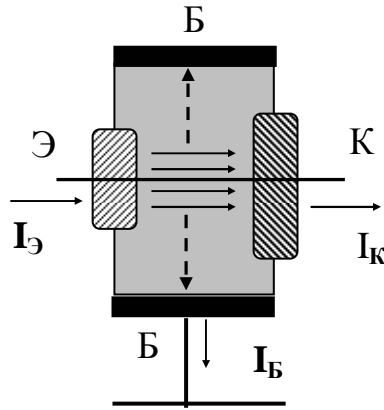


Рис. 4.28 — Структура мощного сплавленного транзистора; пунктирными стрелками показано протекание базового тока, обуславливающее неравномерную инжекцию

4.11 Дрейфовые транзисторы

В неоднородной базе имеется собственное электрическое поле независимо от уровня инжекции. Поэтому механизм движения инжектированных носителей у транзисторов с такой базой преимущественно дрейфовый, откуда следует их общее название — дрейфовые транзисторы.

Поскольку, однако, неоднородность базы достигается путем диффузии примесного материала, можно встретиться с терминами «триод с диффузионной базой» или даже просто «диффузионный транзистор», которые отражают технологию изготовления, а не механизм движения инжектированных носителей.

Особенности дрейфовых транзисторов

На рис. 4.29 изображены две структуры дрейфового транзистора, различие между которыми обусловлено технологическими особенностями.

Сравнивая структуры дрейфового и бездрейфового транзи-

сторгов (см. рис. 4.3), видим, что они заметно различаются. А именно, дрейфовый транзистор значительно более несимметричен: коллекторный слой много толще двух других слоев, велика разница в площадях эмиттера и коллектора, по существу отсутствует пассивная область базы. Все эти особенности обусловлены тем, что в случае дрейфовых транзисторов исходная пластина полупроводника служит основой для создания слоев базы и эмиттера, а сама она в дальнейшем играет роль коллектора, тогда как в случае бездрейфовых (сплавных) транзисторов исходная пластина служит основой для создания слоев эмиттера и коллектора, а сама она в дальнейшем играет роль базы. Резкая асимметрия дрейфовых транзисторов делает их практически необратимыми приборами.

Кроме того, она осложняет анализ тех случаев, когда существенна неоднородность транзистора. Однако при анализе нормального усилительного режима использование одномерной модели вполне оправдано, поскольку движение инжектированных носителей в активной области базы происходит почти строго по оси x .

Как уже отмечалось, база дрейфовых транзисторов получается путем диффузии примесей в глубь исходной пластины (на рис. 4.29 вдоль оси x). Согласно законам диффузии распределение примесных атомов в слое базы должно выражаться дополнительной функцией ошибок.

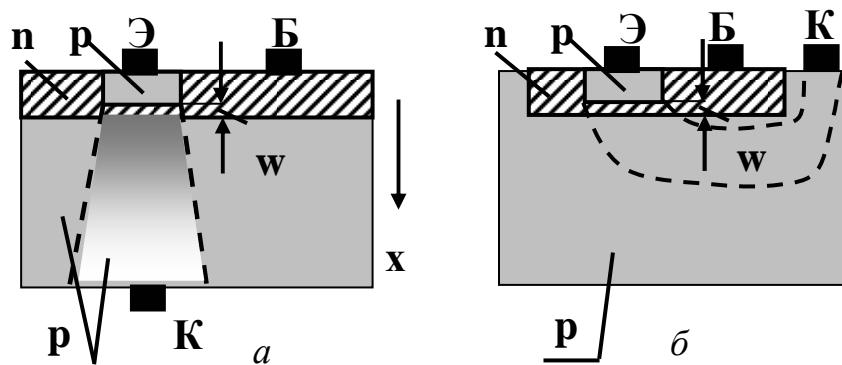


Рис. 4.29 — Упрощенные структуры дрейфовых транзисторов: *a* — выводы электродов в разных плоскостях (меза технология и сплавно-диффузионная технология); *б* — выводы электродов в одной плоскости (планарная технология). Пунктиром показаны линии тока от эмиттера к коллектору

$$N_{\bar{o}}(x) \approx N_{\bar{o}}(0) \operatorname{cerf}\left(\frac{x}{L_N}\right),$$

где x — расстояние, отсчитываемое от эмиттера в глубь базы; L_N — средняя длина диффузии примеси. Поскольку функция ошибки близка к спадающей экспоненте, можно записать приближенное выражение:

$$N_{\bar{o}}(x) \approx N_{\bar{o}}(0) e^{-x/L_N}. \quad (4.62)$$

Это выражение широко используется при анализе дрейфовых транзисторов.

Диффузионная технология позволяет получить очень тонкую базу, что само по себе (даже без учета распределения примесей) приводит к ряду важных следствий. А именно при прочих равных условиях существенно уменьшается время диффузии и увеличивается коэффициент передачи β , поскольку оба эти параметра зависят от квадрата толщины базы. Толщина базы у дрейфовых транзисторов в 5 — 10 раз меньше, чем у сплавных. Поэтому время диффузии и постоянная времени τ_{α} оказываются меньше в десятки раз; соответственно увеличивается граничная частота f_{α} . Коэффициент передачи β по тем же соображениям должен был бы доходить до 1000 и больше. На самом деле он значительно меньше и обычно не превышает 100 — 500. Это объясняется тем, что величины α и β зависят не только от толщины базы, но также от времени жизни и коэффициента инжекции.

В связи с повышенной концентрацией примесей вблизи эмиттера время жизни в базе дрейфового транзистора значительно меньше, чем у сплавного. Коэффициент инжекции более заметно отличается от единицы. Тем не менее у специальных типов дрейфовых транзисторов удается получить значения β до 5000 и более, уменьшая толщину базы до долей микрона. Однако такие значения β получаются за счет резкого уменьшения рабочих напряжений: напряжение смывания составляет у этих транзисторов единицы вольт.

Частотные свойства дрейфовых транзисторов могут ограничиваться не временем пролета, а постоянной времени τ_{γ} . Для того чтобы уменьшить влияние емкости C_{Σ} , часто используют

дрейфовые транзисторы при большом токе эмиттера, например 4 — 5 мА вместо 1 мА. Тогда сопротивление $r_{э}$ уменьшается и постоянная времени $r_{э}C_{э}$ оказывается достаточно малой.

Конечно, величины барьерных емкостей C_k и $C_{э}$ зависят не только от соотношения удельных сопротивлений слоев, но и от площадей переходов. Технология изготовления дрейфовых транзисторов обеспечивает гораздо меньшие площади, чем сплавная технология, и это способствует уменьшению барьерных емкостей.

Вопросы для самопроверки

1. Назначение переходов в биполярном транзисторе.
2. Режимы работы биполярных транзисторов.
3. Схемы включения биполярных транзисторов.
4. Физические основы работы биполярных транзисторов.
5. Распределение носителей в базе транзистора.
6. К чему приводит модуляция толщины базы?
7. Начертите эквивалентную схему идеализированного биполярного транзистора.
8. Выведите уравнения Эберса-Молла.
9. Запишите уравнения идеализированных статических характеристик.
10. Начертите статические характеристики реального транзистора для схемы с общей базой.
11. Нарисуйте эквивалентную схему транзистора при малом сигнале при включении с ОБ для постоянных составляющих.
12. Нарисуйте эквивалентную схему транзистора при малом сигнале для переменных составляющих при включении с ОБ.
13. Запишите статические параметры транзистора.
14. Зависимость барьерной емкости коллекторного перехода от напряжения на переходе.
15. Почему частотные свойства коэффициента инжекции улучшаются с увеличением тока эмиттера?
16. От каких параметров зависит коэффициент передачи тока эмиттера α ?
17. Что означает постоянная времени базы?
18. Зависимость параметров транзистора от температуры.

19. Основные параметры транзистора при включении с ОЭ.
20. Статические характеристики транзистора с ОЭ.
21. Эквивалентная схема транзистора на постоянном токе при малых сигналах в схеме с ОЭ.
22. Эквивалентная схема транзистора для переменных составляющих с ОЭ.
23. Какие разновидности эквивалентных схем Вы знаете?
24. Запишите выражения для входного сопротивления при включении транзистора с ОБ, ОЭ и ОК.
25. Выведите формулы коэффициентов усиления по току для всех схем включения транзистора.
26. Выведите формулы коэффициентов усиления по напряжению для всех схем включения биполярного транзистора.
27. Основные параметры составного транзистора и его эквивалентная схема.
28. Особенности мощных транзисторов.
29. Основные отличия дрейфовых транзисторов от диффузионных?

5. ПОЛЕВЫЕ ТРАНЗИСТОРЫ

Полевыми транзисторами называют такие полупроводниковые приборы, работа которых основана на использовании носителей заряда одного знака: только дырок или только электронов. С этой точки зрения обычные транзисторы, рассмотренные ранее, можно назвать биполярными, так как в них важную роль играют оба типа носителей: инжекция носителей одного знака сопровождается компенсацией образующегося заряда носителями другого знака.

Второй термин — «полевые транзисторы» характеризует механизм управления током: с помощью электрического поля (а не тока, как в биполярных транзисторах). Строго говоря, управление выходным током в обоих транзисторах осуществляется зарядом.

Униполярные полевые транзисторы имеют несколько разновидностей. Простейшую из них, предложенную в 1952 г. Шокли, — назовем полевой транзистор с управляющим p - n переходом.

5.1 Полевой транзистор с управляющим p - n переходом

На рис. 5.1, *а* показана упрощенная структура, которая представляет собой пластинку полупроводника типа n (p), у которой на торцах имеются омические контакты, а на обоих больших гранях — слои типа p . Последние образуют с пластинкой p - n переходы. Оба p слоя соединены между собой и образуют единый электрод, который называется затвором. Двумя другими электродами являются омические контакты.

При этом контакт, от которого движутся основные носители (в нашем случае — электроны), называется истоком, а тот, к которому они движутся, — стоком. Следовательно, в унитроне типа n сток имеет положительную полярность относительно истока. Оба p - n перехода работают в обратном направлении. Для этого на затвор подается отрицательное смещение относительно истока, как показано на рис. 5.1, *б* (далее под U_3 будет пониматься модуль напряжения).

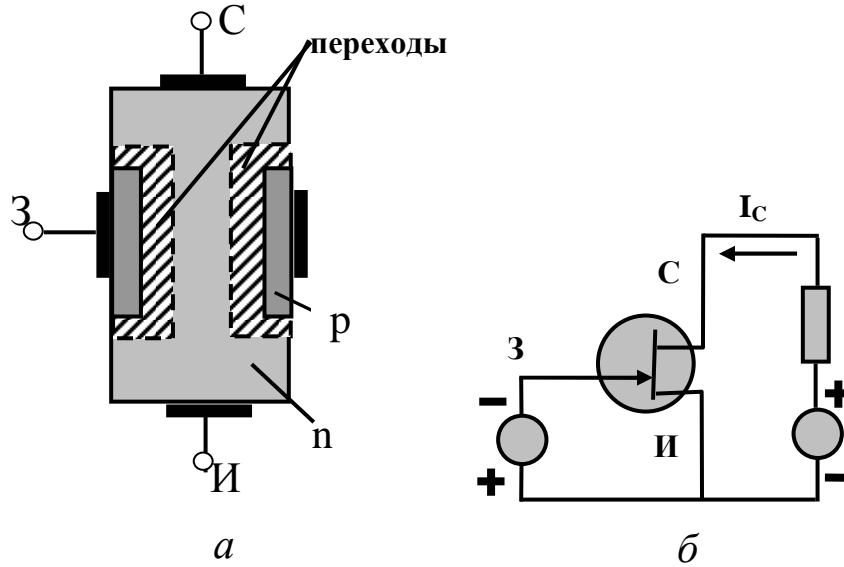


Рис. 5.1 — Конструкция ПТ с p - n переходом (а) и схема его включения (б)

Принцип действия очень прост и заключается в том, что при изменении потенциала затвора меняется ширина p - n переходов, а значит, и рабочее сечение пластинки. В результате меняются ее сопротивление и соответственно токи истока и стока. Поскольку p - n переходы работают в обратном включении, их сопротивление для входного сигнала велико и входная мощность мала.

Полезная мощность, определяемая величиной питающего напряжения и соотношением сопротивлений пластинки и нагрузки, может значительно превышать входную мощность. Таким образом, унитрон является усилительным прибором типа управляемого активного сопротивления, причем он имеет много общего с электронной лампой. Это сходство выражается не только в большом входном сопротивлении, но и в том, что при некотором отрицательном смещении затвора расширившиеся переходы перекрывают все сечение пластинки; что вызовет отсечку тока стока в канале — явление, аналогичное запираению лампы.

Назовем каналом рабочий (переменный) объем пластинки, расположенный между p - n переходами. Пренебрегая пока участками пластинки, прилегающими к истоку и стоку, можно представить структуру унитрона в упрощенном виде (рис. 5.2). Обозначим максимальную толщину канала через a , его ширину через z и длину через L . Пусть $U_c = 0$, тогда канал будет эквипотенциальным слоем и напряжение на p - n переходах будет равно U_A на

протяжении всего канала. Соответственно в любой точке ширина перехода равна l , a – толщина канала.

$$w = a - 2l.$$

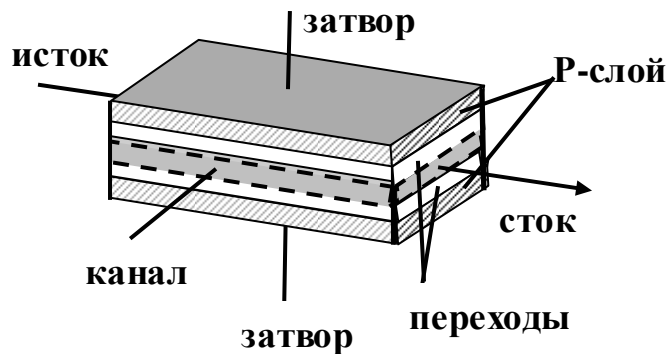


Рис. 5.2 — Рабочая часть полевого транзистора с управляющим p - n переходом

Подставляя сюда выражение (2.12) для ширины l и полагая $qN_D = qn = \frac{1}{\rho\mu_n}$, получаем:

$$w = a - 2\sqrt{2\xi_0\xi\rho\mu_n U_3}.$$

Из условия $w = 0$ легко найти напряжение отсечки:

$$U_{30} = \frac{a^2}{8\xi_0\xi\rho\mu_n}. \quad (5.1)$$

Используя (5.1), запишем толщину канала в более компактной форме:

$$w = a\left(1 - \sqrt{\frac{U_3}{U_{30}}}\right). \quad (5.2)$$

В рабочем режиме, когда $U_C \neq 0$ канал не является эквипотенциальным слоем; в разных точках x потенциал различен: он меняется от $U_x=0$ около истока до $U_x=+U_C$ около стока. Поэтому обратное напряжение на p - n переходах увеличивается в направлении от истока к стоку. Соответственно ширина перехода в этом направлении растет, а канал сужается (рис. 5.3, а) В наиболее узком месте (около стока) напряжение на переходе равно U_3+U_C . С ростом U_C это напряжение в конце концов делается настолько большим, что переходы почти смыкаются (рис. 5.3, б), но это не приводит к отсечке тока, так как само смыкание является следствием увеличения

тока. Вместо отсечки тока происходит отсечка его приращений, т.е. резкое возрастание дифференциального сопротивления канала. При этом на кривой $I_C(U_C)$, начиная с некоторой точки Н, получается практически горизонтальный участок (рис. 5.4).

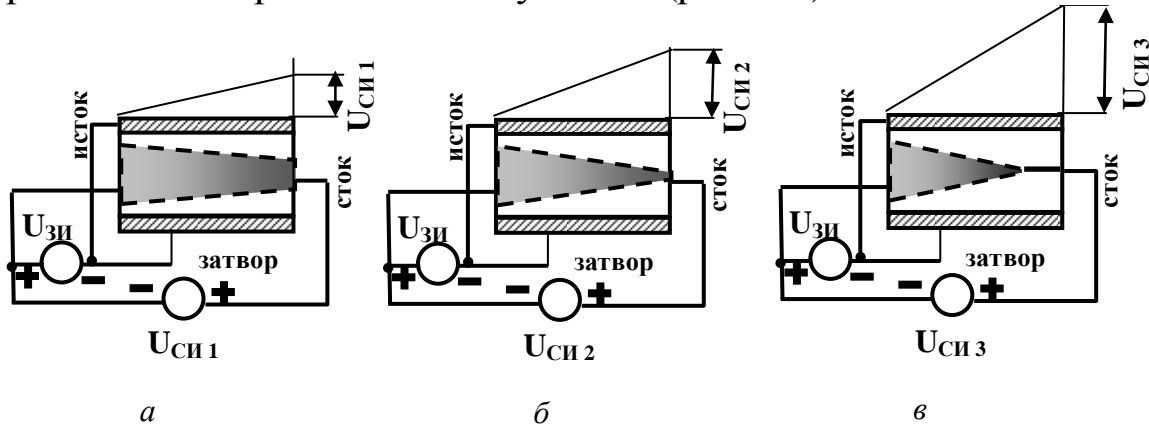


Рис. 5.3 — Сечение канала при ненасыщенном режиме полевого транзистора с управляющим $p-n$ переходом (а), на границе насыщения (б) и в насыщенном режиме (в)

Такой режим можно назвать насыщением тока стока. Таким образом, в режиме насыщения происходит модуляция длины канала. При расчете характеристик полевого транзистора с управляющим $p-n$ переходом следует учесть, что сопротивление канала меняется вдоль оси x , поскольку меняется толщина w . Падение напряжения на элементарном участке dx составляет

$$dU_x = I_C dR_x = I \frac{\rho dx}{L_a} \left(1 - \sqrt{\frac{U_3 + U_x}{U_{30}}}\right)^{-1}.$$

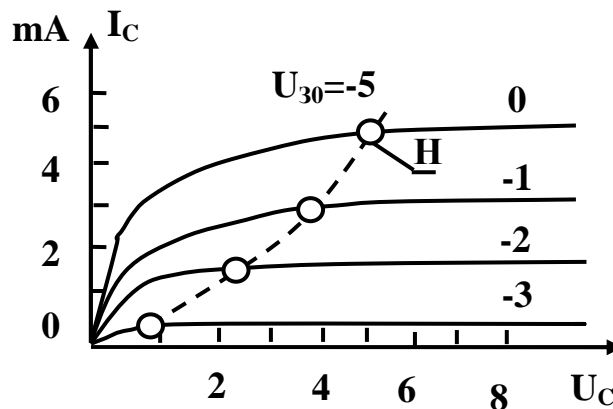


Рис. 5.4 — Семейство выходных характеристик полевого транзистора с $p-n$ переходом

Ток I_C остается неизменным в любом сечении. Поэтому, разделяя переменные и интегрируя обе части соответственно в пределах от 0 до U_C и от 0 до L , можно представить искомую функцию $I_C(U_C)$ в следующем виде

$$I_C = \frac{1}{R_{K0}} \left[U_C + \frac{2}{3} \frac{U_3^{3/2} - (U_3 + U_C)^{3/2}}{U_{30}^{1/2}} \right]. \quad (5.3)$$

Здесь через R_{K0} обозначено минимальное дифференциальное сопротивление канала.

$$R_{K0} = \frac{\rho L}{aZ}. \quad (5.4)$$

Ток насыщения получается равным:

$$I_{C0} = \frac{1}{R_{K0}} \left[\frac{1}{3} U_{30} - U_3 \left(1 - \frac{2}{3} \sqrt{\frac{U_3}{U_{30}}} \right) \right]. \quad (5.5)$$

Поскольку участок насыщения является основным рабочим участком, определим крутизну S именно в этой области. Дифференцируя (5.5) по U_3 , получаем:

$$S = -\frac{\partial I_{CH}}{\partial U_3} = \frac{1}{R_{K0}} \left(1 - \sqrt{\frac{U_3}{U_{30}}} \right). \quad (5.6)$$

Дифференциальное сопротивление r_C в области насыщения согласно (5.5) равно бесконечности. Практически оно может составлять несколько мОм.

Переходя к оценке быстродействия унитрона, напомним, что его работа не связана с инъекцией неосновных носителей и распространением их до стока. Инерционность унитрона обусловлена зарядом барьерных емкостей переходов.

Емкости переходов заряжаются через сопротивление канала. При этом разные участки емкостей заряжаются через разные сопротивления в зависимости от расстояния данного участка от истока (рис. 5.5). Чтобы не усложнять анализа, примем, что канал имеет одинаковое сечение на всем протяжении. Тогда средняя емкость одного из р-п переходов и среднее сопротивление канала будут иметь следующие значения:

$$C_3 = \frac{\xi_0 \xi(LZ)}{0,5(a-w)}; \quad (5.7a)$$

$$R_K = \frac{0,5\rho L}{Zw}. \quad (5.76)$$

Постоянная времени затвор-исток будет равна удвоенному произведению этих параметров, так как через сопротивление канала заряжаются одновременно обе емкости:

$$\tau_3 = 2R_K C_3 = \frac{2\xi_0 \xi \rho L^2}{w(a-w)}. \quad (5.8a)$$

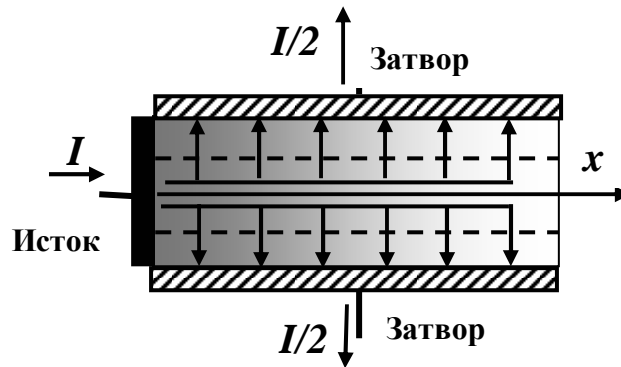


Рис. 5.5 — Процесс заряда емкости затвора в полевом транзисторе с p - n переходом

Постоянная времени τ_3 зависит от толщины w .

$$\tau_{3\text{МИН}} = 8\xi_0 \xi \rho \frac{L^2}{a^2}. \quad (5.8b)$$

В заключение заметим, что в режиме насыщения вольт-амперная характеристика (5.5) хорошо аппроксимируется более простой квадратичной характеристикой.

$$I_C = \frac{1}{R_{K0}} \frac{(U_{30} - U_3)^2}{3U_{30}}. \quad (5.9a)$$

Дифференцируя (5.9a), получаем выражение для крутизны

$$s = \frac{2/3}{R_{K0}} \frac{U_0 - U_3}{3U_{P0}}. \quad (5.9b)$$

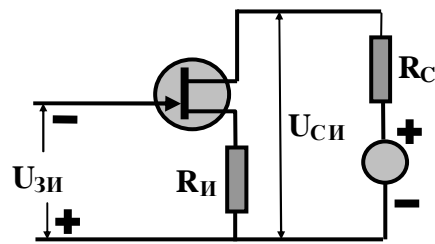


Рис. 5.6 — Схема включения полевого транзистора с $p-n$ переходом

Эквивалентная схема. Аналогия между полевыми транзисторами и электронными лампами объясняет сходство их эквивалентных схем (рис. 5.7); канал и управляющий $p-n$ переход представлены сосредоточенными параметрами. Канал представлен дифференциальным сопротивлением и межэлектродной емкостью, величина которой определяется геометрией и материалом прибора.

Управляющий $p-n$ переход представлен сопротивлением и усредненной емкостью. Сопротивлением затвора-истока часто пренебрегают в связи с его большой величиной; поэтому на рис. 5.7 оно показано пунктиром. Усилительные свойства транзистора отражены генератором тока SU_3 . Сопротивления истока и стока показаны условно, так как их влияние автоматически учитывается при измерении параметров.

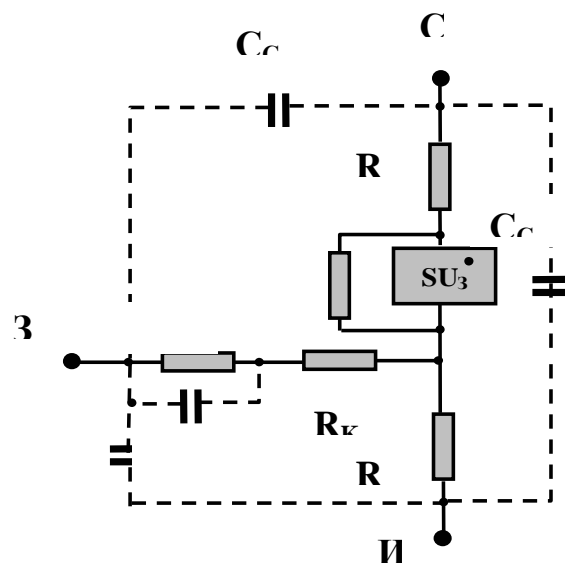


Рис. 5.7 — Эквивалентная схема униполярного транзистора для переменных составляющих токов и напряжений

5.2 Полевые транзисторы с изолированным затвором

Структура и классификация. Спецификой унитарных является максимальная проводимость канала при нулевом смещении на затворе. С ростом смещения (по модулю) проводимость канала уменьшается вплоть до полной отсечки. Смещение может иметь только одну полярность, соответствующую отсутствию инжекции через переход.

У полевых транзисторов с изолированным затвором последний представляет собой металлический слой, изолированный от полупроводника тонкой диэлектрической пленкой. Наличие диэлектрика снимает ограничение на полярность смещения: она может быть как положительной, так и отрицательной, причем в обоих случаях ток затвора отсутствует. Структура таких транзисторов (металл — диэлектрик — полупроводник) лежит в основе широко распространенного их названия МДП-транзисторы. В том весьма распространенном случае, когда диэлектриком является окисел (двуокись кремния), их называют МОП-транзисторами. Две основные структуры МДП-транзисторов показаны на рис. 5.8. Первая из них (рис. 5.8, *а*) характерна наличием специально осуществленного (собственного или встроенного) канала, проводимость которого модулируется смещением на затворе. В случае канала *p*-типа положительный потенциал U_3 «отталкивает» дырки из канала (режим обеднения), а отрицательный — «притягивает» их (режим обогащения). Соответственно проводимость канала либо уменьшается, либо увеличивается по сравнению с ее значением при нулевом смещении.

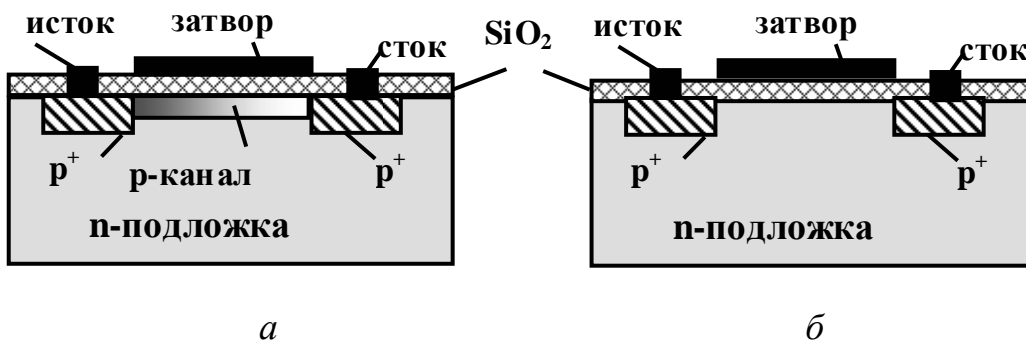


Рис. 5.8 — Структуры МДП-транзисторов с собственным (*а*) и индуцированным (*б*) каналами

Вторая структура (рис. 5.8, б) характерна отсутствием структурно выраженного канала. Поэтому при нулевом смещении на затворе проводимость между истоком и стоком отсутствует: исток и сток образуют с подложкой встречно включенные $p-n$ переходы. Тем более не может быть проводимости между истоком и стоком при положительной полярности смещения, когда к поверхности полупроводника притягиваются дополнительные электроны.

Однако при достаточно большом отрицательном смещении, когда приповерхностный слой сильно обогащается притянутыми дырками, между истоком и стоком образуется своего рода индуцированный (наведенный полем) канал, по которому может протекать ток. В настоящее время транзисторы с индуцированным каналом имеют наибольшее распространение, главным образом из-за простоты их изготовления.

Оба типа МДП-транзисторов могут изготавливаться как с p -, так и с n -каналом, что дополнительно расширяет возможности их схемного применения.

Физические процессы. Рассмотрим работу МДП-транзистора с индуцированным p -каналом. Весь анализ справедлив и для транзисторов с n -каналом.

Ниже все напряжения и потенциалы записываются в виде их модулей, чтобы избежать знаков минуса, характеризующих отрицательную полярность смещений в транзисторах (p - или n -типа).

В отсутствие смещений приповерхностный слой полупроводника обычно обогащен электронами из-за наличия ловушек на границе $Si-SiO_2$ и наличия положительных ионов в пленке диэлектрика. Соответственно энергетические зоны искривлены вниз, и начальный поверхностный потенциал ϕ_{30} отрицателен (рис. 5.9, а).

По мере роста отрицательного смещения U_3 электроны отталкиваются от поверхности. При этом энергетические зоны сначала спрямляются, а затем искривляются вверх, т.е. поверхностный потенциал делается положительным.

Однако до тех пор, пока приповерхностная область сохраняет электронный тип проводимости или, точнее, пока сток и исток образуют с этой областью выпрямляющие контакты, проводи-

мость рабочей цепи остается крайне малой и протекание тока в этой цепи невозможно. Существует некоторое пороговое напряжение $U_3=U_0$, при превышении которого энергетические зоны искривляются настолько сильно, что вблизи поверхности образуется инверсионный дырочный слой. Именно этот слой играет роль индуцированного канала (рис. 5.9, б). Принято считать, что пороговое напряжение соответствует поверхностному потенциалу $\varphi_{cm} = 2\varphi_F$ (электростатический потенциал φ_E принимается равным нулю).

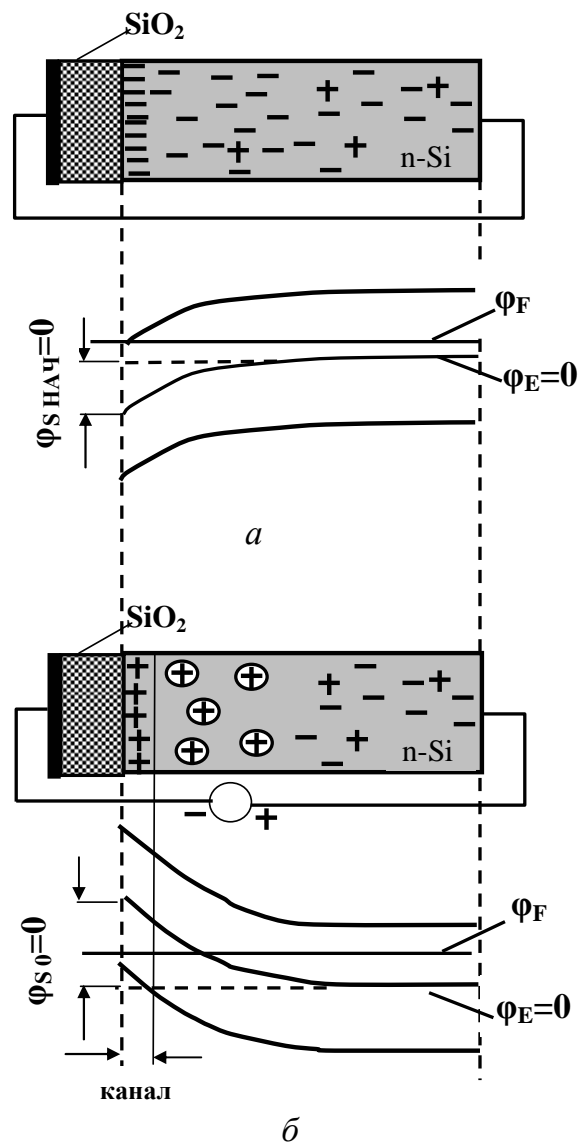


Рис. 5.9 — Распределение носителей и зонные диаграммы в МОП-транзисторе с индуцированным *p*-каналом:
a — равновесное состояние; *б* — при отрицательном смещении на затворе

Дальнейший рост напряжения U_3 слабо влияет на величину φ_S , поскольку изменения последней всего на несколько φ_T достаточно для изменения концентрации дырок в десятки раз.

Наряду с образованием дырочного канала под затвором образуется также обедненный слой, в котором положительный заряд обусловлен обнаженными ионами доноров (рис. 5.9, б). Образование обедненного слоя вызвано отталкиванием основных носителей подложки — электронов — от поверхности.

Инверсионный слой значительно тоньше обедненного слоя. Толщина последнего обычно составляет сотни ангстрем, а толщина индуцированного канала составляет всего 10 — 20 Å.

Как видим, дырки буквально «прижаты» к поверхности полупроводника. Отсюда ясно, что структура и свойства границы полупроводник — диэлектрик играют в МДП-транзисторах исключительно важную роль.

Если в начальном состоянии все электроды МДП-транзистора находились при нулевом потенциале, а затем на затвор была подана ступенька напряжения $U_3 > U_0$, то в первый момент поле будет иметь примерно такую конфигурацию, как показано на рис. 5.10, а.

Под действием этого поля электроны перемещаются в сторону от поверхности, обнажая ионы доноров и образуя обедненный слой, а дырки движутся к поверхности, накапливаются вблизи границы с диэлектриком.

С увеличением напряжения на стоке (при постоянном потенциале $U_{3И}$) ток I_C сначала нарастает почти линейно (пока дырочный канал вблизи стока слабо деформирован), затем нарастание тока замедляется.

Семейство характеристик МДП-транзистора показано на рис. 5.11.

В целом семейство вольт-амперных характеристик МДП-транзистора напоминает семейство характеристик унитрона (рис. 5.4).

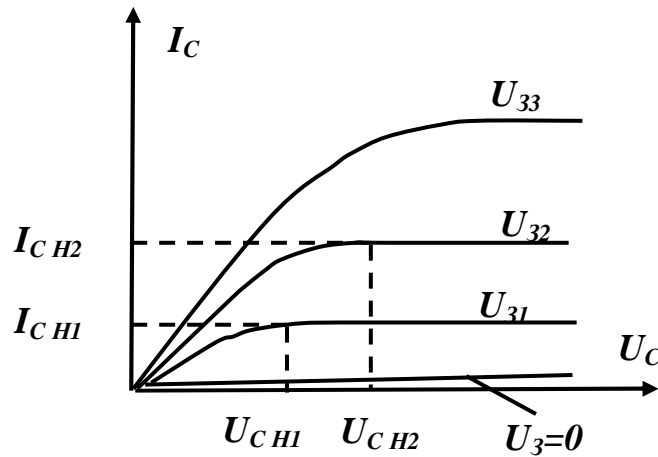


Рис. 5.10 — Статические вольт-амперные характеристики МДП-транзистора

Характеристики и параметры в 1-м приближении. Рассмотрим самую простую аппроксимацию выражения, которой удобно пользоваться при инженерных расчетах.

$$I_C = b[(U_3 - U_0)U_C - \frac{1}{2}U_C^2], \quad (5.10)$$

где $b = \frac{\mu_{zz}c_0}{L} = \frac{\zeta\zeta_0z}{Ld}$ — параметр, определяющий «масштаб» характеристики, называется удельной крутизной.

Такая аппроксимация всегда действительна при малых сток-ковых напряжениях ($U_C < \varphi_{Sm}$).

Из выражения (5.10) легко найти напряжение насыщения, полагая $\frac{dI_C}{dU_C} = 0$.

$$U_{CH} = U_3 - U_0. \quad (5.11)$$

Как уже отмечалось, величина тока $I_C(U_{3H})$ сохраняется при всех значениях $U_C > U_{CH}$.

Поэтому, подставляя (5.11) в (5.10), получаем вольт-амперную характеристику МДП-транзистора в режиме насыщения:

$$I_C = \frac{1}{2}b(U_3 - U_0)^2. \quad (5.12)$$

В усилительной технике МДП-транзисторы чаще используются в режиме насыщения, поскольку им свойственны наименьшие нелинейные искажения и оптимальные значения дифферен-

циальных параметров: крутизны $S = \partial I / \partial U$, внутреннего сопротивления $R_i = \partial U_c / \partial I_c$ и собственного коэффициента усиления $\mu = dU_c / dU_3$. Эти параметры связаны между собой «ламповым» соотношением:

$$\mu = sR_i. \quad (5.13)$$

Крутизна в режиме насыщения определяется из выражения (5.12):

$$s = b(U_3 - U_0). \quad (5.14)$$

Как видим, крутизна линейно зависит от эффективного напряжения на затворе $U_3 - U_0$, а при заданном эффективном напряжении пропорциональна параметру b . Крутизна МДП-транзистора однозначно связана с током. Эту связь легко установить с помощью выражений (5.12) и (5.14):

$$s = \sqrt{2bI_c}. \quad (5.15)$$

Из выражения (5.14) ясно, что крутизна МДП-транзистора возрастает с уменьшением толщины диэлектрика, а также с увеличением ширины канала Z .

Однако оба эти фактора одновременно способствуют увеличению емкости между затвором и каналом, которая описывается выражением

$$C_3 = C_0(ZL) = \frac{\xi_0 \xi ZL}{d}. \quad (5.16)$$

Поэтому такой глобальный параметр всякого усилительного прибора, как добротность, определяемая отношением $S / 2,2C$, не зависит от величин Z и d . Поделив (5.14) на (5.16), получаем добротность МДП-транзистора в следующем виде:

$$D = \frac{\mu}{2,2L^2} (U_3 - U_0). \quad (5.17)$$

Отсюда следует, что основным средством повышения добротности является уменьшение длины канала L . В отличие от крутизны внутреннее сопротивление и коэффициент усиления МДП-транзистора не могут быть определены из выражения (5.12), поскольку оно не содержит напряжения. Согласно (5.12)

характеристики в режиме насыщения должны быть горизонтальными, а значит, $R_i = \infty$, $\mu = \infty$.

Разумеется, в области достаточно больших напряжений U_C наступают предпробойные явления, а затем и пробой, сопровождающийся резким возрастанием тока I_C и столь же резким уменьшением сопротивления R_i . В импульсных схемах МДП-транзистор работает в качестве ключа, и основным интерес представляют две крайние рабочие точки, соответствующие запертому и максимально открытому состоянию ключа (рис. 5.12).

Запертое состояние (точка 1) характеризуется условием $U_3 < U_0$. При этом в цепи стока протекает лишь некоторый остаточный ток, обусловленный утечками по поверхности, а также обратным током р-п перехода стока (если подложка находится под нулевым или положительным потенциалом). В качественных МОП-транзисторах этот ток не превышает нескольких наноампер. Максимально открытое состояние ключа (точка 2) характерно большими значениями эффективного напряжения $U_3 - U_0$ и расположением рабочей точки на крутом участке соответствующей вольт-амперной кривой. Обычно в открытом состоянии ток I_C оказывается заданным внешней цепью, а интерес представляет остаточное напряжение U_{CO} . Это напряжение легко определить из формулы (5.10), однако в общем виде выражение получается громоздким. На практике остаточное напряжение мало.

Тогда

$$U_{CO} = \frac{I_C}{b(U_3 - U_0)}. \quad (5.18)$$

Выражение (5.18) действительно при условии

$$U_3 - U_0 > (2 - 3)U_{CO}.$$

Еще одним параметром, важным для характеристики открытого состояния ключа, является сопротивление на начальном участке кривой $I_C(U_C)$. Дифференцируя (5.10) по U_C , получаем:

$$R_0 = \frac{1}{b(U_3 - U_0)}. \quad (5.19)$$

Это же выражение можно получить непосредственно из формулы (5.18).

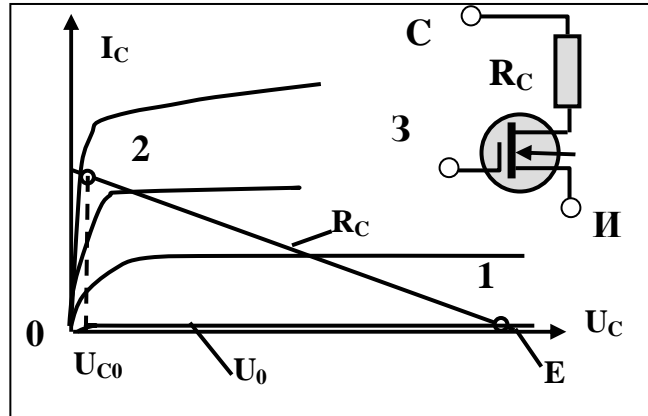


Рис. 5.11 — Рабочие точки МДП-транзистора в ключевом режиме

Поскольку при малых значениях U_C сечение дырочного слоя и концентрация дырок в нем почти не зависят от x , величину R_0 обычно называют сопротивлением канала. Из (5.18) и (5.19) очевидно, что остаточное напряжение на МДП-транзисторном ключе имеет чисто омический характер:

$$U_{C0} = I_C R_0.$$

Это обстоятельство является важным преимуществом МДП-транзисторов, так как в случае биполярных транзисторов даже при нулевом токе имеется вполне конечное остаточное напряжение.

Эквивалентная схема. Для МДП-транзисторов в отличие от унитронов эквивалентная схема для переменных составляющих должна отражать влияние подложки (независимо от наличия потенциала). Одна из таких схем, предложенная В.Н. Кононовым, показана на рис. 5.12. Несмотря на внешнюю симметрию, обе половины схемы имеют разные значения. Переходные и частотные характеристики МДП-транзисторов обусловлены перезарядкой межэлектродных емкостей через внешние резисторы, а также перезарядкой емкости затвор-канал через сопротивление канала. Последний процесс накладывает принципиальное ограничение на быстродействие транзистора. Строго говоря, канал МДП-транзистора следует рассматривать как распределенную систему.

Однако удобнее в первом приближении охарактеризовать инерционность этой системы постоянной времени, т.е. уподобить канал простейшей RC-цепочке. В качестве элементов такой цепочки можно принять емкость затвора (5.16) и сопротивление ка-

нала (5.19). Тогда постоянная времени канала (она же постоянная времени крутизны) будет иметь вид:

$$\tau_S = R_0 C_3 = \frac{L^2}{\mu(U_3 - U_0)}. \quad (5.20)$$

Как видим, это по существу обратная величина добротности (5.17).

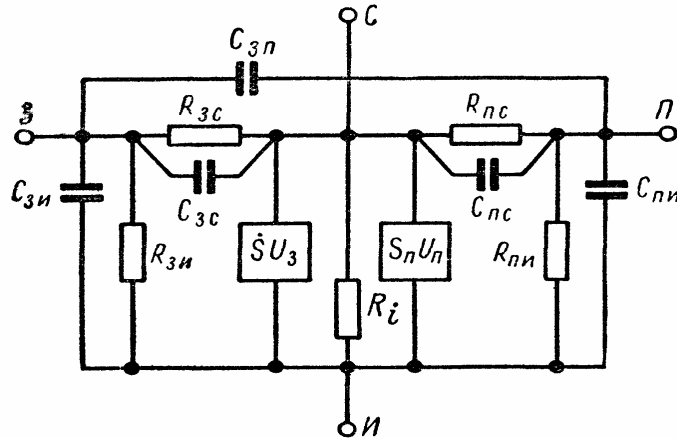


Рис. 5.12 — Эквивалентная схема МДП-транзистора для малых переменных составляющих

Для улучшения переходных и частотных свойств МДП-транзистора нужно в первую очередь уменьшать длину канала, а также увеличивать приповерхностную подвижность. При обычной длине канала 5 — 10 мкм граничная частота крутизны ($1/2\pi\tau$) лежит в пределах 100 — 300 МГц. Однако при ультракоротких (доли микрона) каналах удастся обеспечить граничные частоты до 10 — 20 ГГц и выше. На эквивалентной схеме параметры C_3 и R_0 не показаны, так как их влияние отражено операторным или комплексным характером крутизны S . В типичных случаях емкость затвора составляет десятые доли пикофарады (иногда 1 — 2 пф), а сопротивление канала — сотни Ом.

Межэлектродные емкости зависят от конструкции и геометрии прибора и обычно не превышают 1 пф. Интересной особенностью МДП-транзисторов является наличие критического значения рабочего тока I_C , при котором этот ток почти не зависит от температуры, поскольку взаимно компенсируются влияния температуры на удельную крутизну S (через подвижность μ) и пороговое напряжение U_0 , от которых в первую очередь зависит величина тока.

Вопросы для самопроверки

1. В чём принципиальные отличия в работе полевых транзисторов от биполярных?
2. Объясните физические принципы работы полевого транзистора с $p-n$ переходом.
3. Почему канал у стока сужается?
4. Какие типы каналов у полевых транзисторов с $p-n$ переходом Вы знаете?
5. Нарисуйте статические выходные и передаточные характеристики транзистора с $p-n$ переходом.
6. Почему нормальная работа полевого транзистора с $p-n$ переходом обеспечивается только при обратном смещении на переходе?
7. Основные параметры унитрона.
8. Начертите эквивалентную схему унитрона.
9. В чём особенность полевых транзисторов с изолированным затвором?
10. Какие типы полевых транзисторов с изолированным затвором Вы знаете?
11. Объясните физические принципы работы полевого транзистора со встроенным каналом?
12. Начертите статические вольт-амперные характеристики полевого транзистора со встроенным каналом.
13. Основные физические принципы работы полевого транзистора с индуцированным или наведенным каналом.
14. В чем принципиальное отличие передаточных характеристик полевого транзистора с индуцированным каналом от аналогичных для встроенного канала?
15. Начертите передаточные характеристики транзистора с индуцированным каналом для p - и n -типа.
16. Основные параметры полевых транзисторов с изолированным затвором.
17. Объясните, что означает понятие «насыщение тока стока» и с чем это связано?
18. Начертите эквивалентную схему для полевого транзистора с изолированным затвором.

6. ТРАНЗИСТОРНЫЕ КЛЮЧИ

6.1 Введение

Будем называть ключом такую схему, основное назначение которой состоит в замыкании и размыкании цепи нагрузки с помощью управляющих входных сигналов. По аналогии с механическим ключом (контактом) качество транзисторного ключа определяется в первую очередь минимальным падением напряжения на нем в замкнутом состоянии, минимальным током в разомкнутом состоянии, а также скоростью перехода из одного состояния в другое. Ключевые схемы лежат в основе более сложных импульсных схем; поэтому материал данной главы будет широко использован в дальнейшем. Кроме того, некоторые типы транзисторных ключей имеют самостоятельное значение в качестве бесконтактных прерывателей. Транзисторный ключ имеет ряд схемных вариантов, простейший из которых приведен на рис. 6.1. Активная нагрузка R_k включена в коллекторную цепь, а управляющие импульсы поступают от генератора E_B через сопротивление R_B , которое, в частности, может быть внутренним сопротивлением генератора.

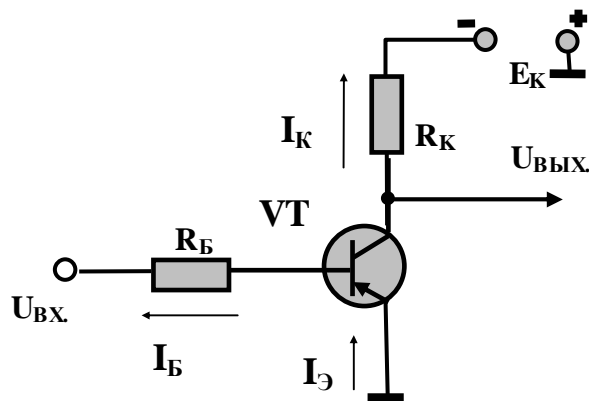


Рис. 6.1 — Простейший ключ на транзисторе, включенном по схеме ОЭ

Включение транзистора по схеме ОЭ имеет в импульсной технике столь же широкое распространение, как и в области усилителей. Поэтому такое включение положено в основу после-

дующего анализа, а особенности других включений и других схемных вариантов ключей будут охарактеризованы позднее. В отличие от усилителей, в которых (за исключением мощных каскадов класса В) транзисторы работают в активном режиме, в ключах (импульсных устройствах вообще) транзисторы должны работать в нескольких качественно различных режимах. Эти режимы характеризуются полярностями напряжений на переходах транзистора.

А именно различают:

1. Режим отсечки ($U_{БЭ} > 0$; $U_{КБ} < 0$).
2. Нормальный активный режим ($U_{БЭ} < 0$; $U_{КБ} < 0$).
3. Инверсный активный режим ($U_{БЭ} > 0$; $U_{КБ} > 0$).
4. Режим насыщения ($U_{БЭ} < 0$; $U_{КБ} > 0$). Последний было бы правильнее называть режимом двойной инжекции, так как насыщение (тока), как увидим ниже, есть лишь результат такого режима в ключевых схемах.

6.2 Статические характеристики ключа ОЭ

Рассмотрим семейство характеристик транзистора, включенного по схеме ОЭ (рис. 6.2). Проведем на характеристиках линию нагрузки R_k , или построим статическую вольт-амперную характеристику резистора R_k , уравнение которой имеет вид

$$I = \frac{E_k}{R_k}, \text{ т.е. закон Ома. Учитывая, что это уравнение прямой ли-}$$

нии, для её построения достаточно найти две точки. При токе, равном нулю, — точка E_K . При падении напряжения на резисторе R_k напряжения E_K ток равен I_K . Минимальный ток ключа соответствует точке А, а минимальное падение напряжения на транзисторе — точке 1. В точке 2 транзистор заперт, так как на его базе действует положительное смещение; в точке 1 транзистор открыт и насыщен. Рассмотрим подробнее эти два состояния, положив в основу формулы (4.4), описывающие идеализированный германиевый транзистор в режиме большого сигнала.

Режим отсечки. Пусть обе э. д. значительно превышают величину температурного потенциала ϕ_T . Положим, кроме того, что сопротивления R_B и R_K выбраны таким образом, что падением

напряжения на них от протекания обратных токов базы и коллектора можно пренебречь. Тогда можно считать $|U_B| \approx E_B$ и $|U_K| \approx E_K$, т.е. пренебречь экспонентами в формулах (4.4). Переходя в последних от α к β , используя соотношение (4.5) и полагая $\beta_N \gg \beta_i$, получаем следующие значения токов в закрытом транзисторе.

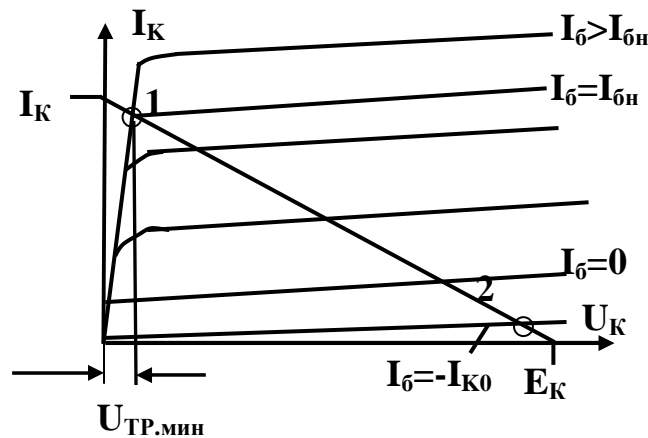


Рис. 6.2 — Рабочие точки ключа ОЭ на семействе коллекторных характеристик

$$I_{\mathcal{E}} = -\frac{1 + \beta_N}{1 + \beta_I + \beta_N} \frac{\beta_I}{\beta_N} I_{K0} \approx -\frac{\beta_I}{\beta_N} I_{K0}; \quad (6.1a)$$

$$I_K = \frac{1 + \beta_N}{1 + \beta_I + \beta_N} I_{K0} \approx I_{K0}; \quad (6.1б)$$

$$I_B = -\frac{1 + \beta_N}{1 + \beta_I + \beta_N} \left(1 + \frac{\beta_I}{\beta_N}\right) I_{K0} \approx -I_{K0}. \quad (6.1в)$$

Отсюда следует, что ток эмиттера значительно меньше остальных двух токов, поскольку $I_{\mathcal{E}} = I_{K0}(1 - \alpha)$, поэтому в режиме отсечки имеем:

$$I_K \approx -I_{\mathcal{E}} \approx I_{K0}.$$

Рассмотренный случай называют режимом отсечки, так как выходной ток имеет минимальное значение (**не путать с режимом отсечки тока стока в полевом транзисторе**).

С уменьшением положительного смещения на базе до нуля транзистор продолжает оставаться запертым, но его токи несколько изменяются.

Когда напряжение на базе станет отрицательным, транзистор перейдет в активный режим. Дальнейшее увеличение тока приводит к увеличению тока коллектора. Когда ток базы станет равным току базы насыщения (I_{BH}), ток коллектора достигнет максимальной величины и станет равным току коллектора насыщения (I_{HAC}).

Режим насыщения. В режиме насыщения оба $p-n$ перехода смещены в прямом направлении и напряжения на них не превышают нескольких φ_T . Говорят, транзистор «стянулся в точку».

Эквивалентная схема ключа в режиме насыщения приведена на рис. 6.3.

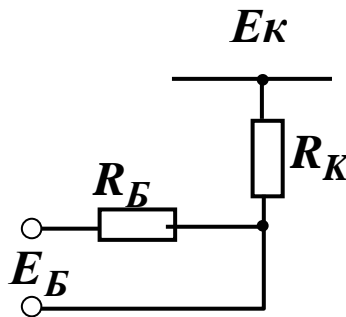


Рис. 6.3 — Эквивалентная схема ключа в режиме насыщения

$$I_{\bar{\sigma}} = \frac{E_B - U_{\bar{\sigma}\bar{\epsilon}}}{R_{\bar{\sigma}}} \approx \frac{E_{\bar{\sigma}}}{R_B}; \quad (6.2a)$$

$$I_K = \frac{E_K - U_{\bar{\sigma}\bar{\epsilon}}}{R_K} \approx \frac{E_K}{R_K}. \quad (6.2b)$$

Из формул (6.2) следует, что в режиме насыщения токи через транзистор определяются внешними элементами схемы.

Целесообразно, во всяком случае в статическом режиме, иметь критерий насыщения, выраженный через токи.

$$\frac{I_B \beta}{I_{KH}} > 1. \quad (6.3)$$

Как видим, режим насыщения определяется не величиной токов, а их соотношением и может иметь место при весьма малых токах, например порядка десятков микроампер и меньше. Опыт показывает, что удобно иметь не только критерий насыщения, но и количественную характеристику «глубины» насыщения, который можно записать в виде

$$S = \frac{I_B}{I_{BH}} = \frac{I_B \beta}{I_{KH}}. \quad (6.4)$$

Коэффициент S называют степенью насыщения, а произведение $I_B \beta = I_{K.КАЖ}$ — кажущимся током коллектора. Понятие кажущийся ток коллектора является условным и означает, что в цепи коллектора протекал бы такой ток, если бы не происходило его ограничения резистором R_K и источником питания E_K .

В заключение анализа насыщенного ключа в статическом режиме отметим следующее:

1. В запертом состоянии токи через переходы мы приняли равными нулю, а это значит, что мощность, которая рассеивается на транзисторе, равна нулю.
2. В режиме насыщения транзистор представляет эквипотенциальную поверхность, а это значит, что и в этом случае рассеиваемая мощность равна нулю.

Однако при переходе транзистора из режима отсечки токов и обратно он в течение переходного процесса работает в активном режиме, и на транзисторе рассеивается мощность. Чем выше частота переключения, тем больше мощность рассеивается на транзисторе.

6.3 Переходный процесс в насыщенном ключе при открывании транзистора

При анализе переходных процессов для упрощения анализа будем считать, что емкость коллекторного перехода равна нулю, а её влияние учтем позднее.

На рис. (6.4) приведены эпюры токов в ключе при открывании транзистора.

Пусть в момент времени t_0 на входе ключа действует отрицательное напряжение, которое смещает эмиттерный переход в

прямом направлении, амплитудой достаточной, чтобы транзистор зашел в режим насыщения.

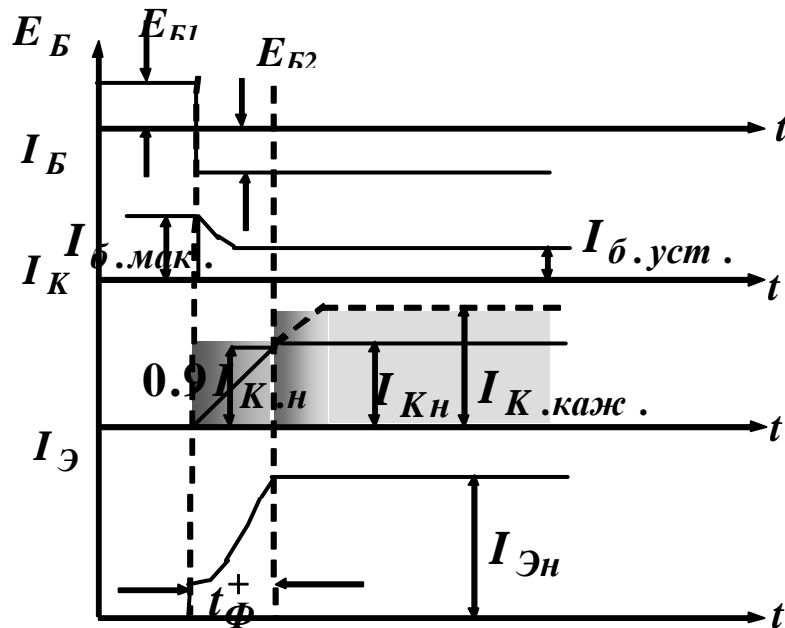


Рис. 6.4 — Эпюры изменения токов в насыщенном транзисторном ключе при открывании (формирование положительного фронта)

В начальный момент времени барьерная емкость эмиттерного перехода препятствует смещению его в прямом направлении. Ток базы при $t_0 \geq 0$ равен $I_{BМАК} = \frac{E_{B2}}{R_B}$. По мере заряда ем-

кости ток базы уменьшается и стремится к установившемуся значению, которое равно $I_B = \frac{E_B}{R_B + r_Э}$, где $r_Э$ — дифференциальное

сопротивление эмиттерного перехода. Плавное нарастание коллекторного тока связано с перераспределением неосновных носителей в области базы (в нашем случае дырок) или зарядом диффузионной ёмкости эмиттерного перехода $C_Э$. По мере заряда этой емкости ток коллектора возрастает и стремится к своему установившемуся значению I_{KH} .

Время, за которое ток коллектора увеличился до значения $0.9I_{KH}$, будем называть длительностью фронта $t_φ^+$. После того как ток коллектора достиг значения I_{KH} , переходный процесс не за-

канчивается, т.к. $I_B > I_{BH}$. После формирования фронта идет накопление избыточных носителей в базе у коллекторного и эмиттерного переходов. Количество избыточного заряда (дырок) в базе растет, а следовательно, положительный потенциал в базе возрастает, что приводит к уменьшению инжекции дырок из эмиттера. Наступает равновесное состояние, когда количество генерируемых носителей равно числу рекомбинированных. Величина избыточного заряда в базе равна $Q_{ИЗБ.} = A(I_{K.КАЖ.} - I_{KH})$, где A — постоянный коэффициент.

Проведем качественную оценку влияния входных и выходных параметров ключа на время формирования длительности фронта.

Влияние выходных параметров на длительность t_{ϕ}^+

Из рис. 6.5 видно, что при изменении параметров выходной цепи, например, увеличения сопротивления R_K или уменьшения напряжения E_K не изменяет скорости переходного процесса, т.к. ток базы остаётся постоянным.

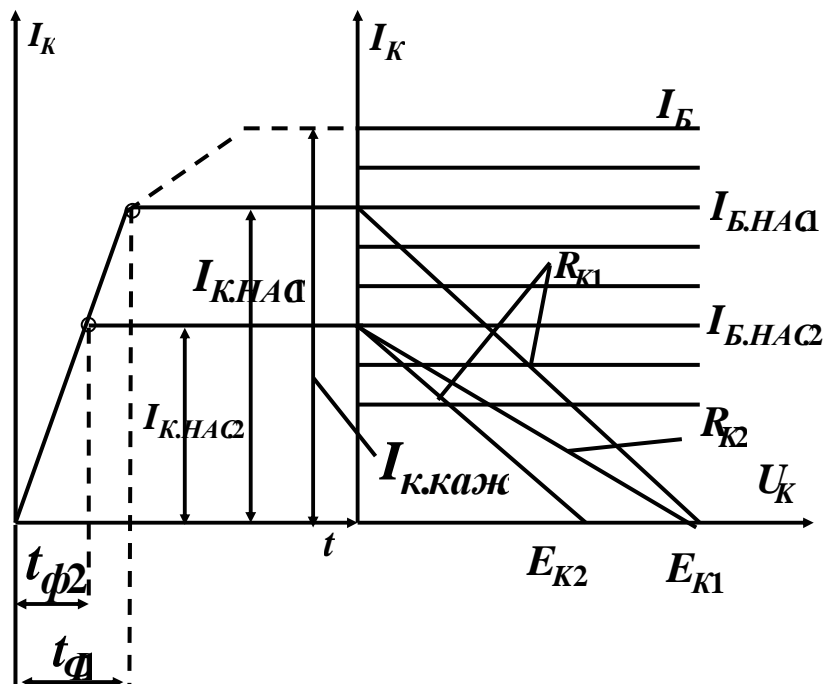


Рис. 6.5 — Переходные процессы в насыщенном ключе формирования длительности фронта при изменениях параметров коллекторной цепи

Ток коллектора уменьшается и раньше достигает установившегося значения, следовательно, длительность фронта сокращается. Учитывая, что количество инжектированных носителей из эмиттера остаётся неизменным, а величина заряда, ушедшего через коллекторную цепь, уменьшилась, следовательно, увеличился объемный заряд в базе.

Таким образом, можно сделать вывод, что, выбирая соответствующим образом параметры выходной цепи в насыщенном ключе, можно без изменения скорости переходного процесса управлять длительностью фронта.

Из рис. 6.6 следует, что изменение входного тока, например его увеличение, приводит к уменьшению t_{ϕ}^+ за счет повышения скорости переходного процесса при неизменном $I_{K,НАС}$. Величина избыточного заряда в базе при этом возрастает.

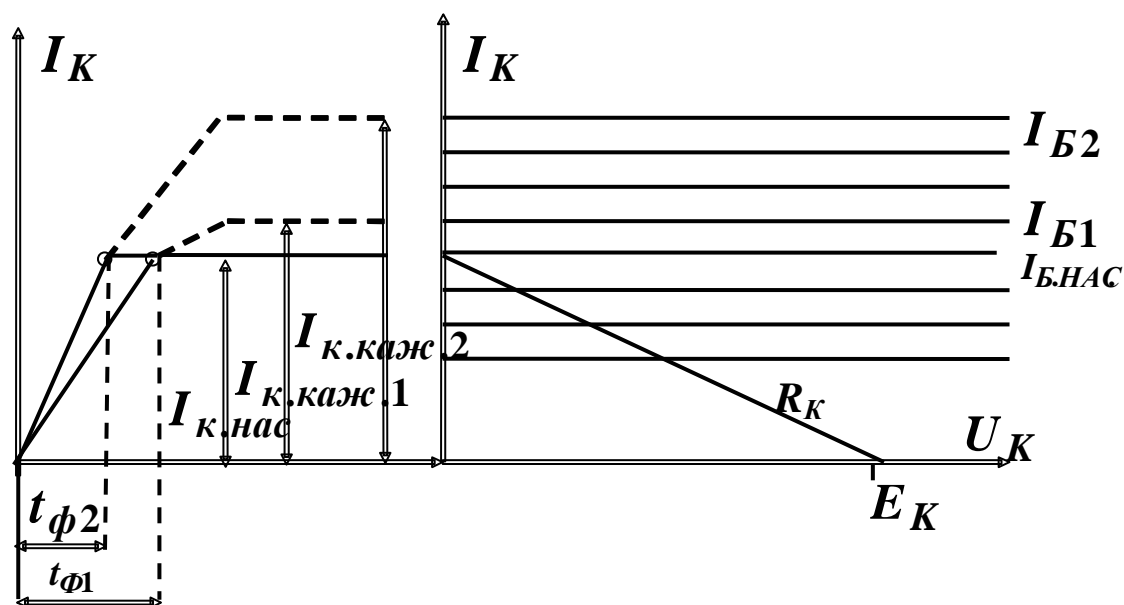


Рис. 6.6 — Переходные процессы в ключе при формировании длительности фронта при изменении параметров входной цепи

Анализ переходного процесса при открывании транзистора показывает, что длительность фронта можно изменять, выбирая соответствующим образом параметры входной или выходной цепи. При

этом необходимо помнить, что уменьшение длительности фронта приводит к увеличению избыточного заряда в базовой области.

Влияние входных параметров ключа на длительность фронта видно из рис. 6.6.

Количественная оценка длительности фронта

Изображение коллекторного тока с учетом формул Эберса-Молла равно:

$$I_K(p) = \frac{I_B}{p} \beta(p), \quad (6.5)$$

где p — оператор Лапласа.

Переходя от изображения к оригиналу, получим

$$I_K(t) = I_B \beta_0 \left(1 - e^{-t/\tau_\beta} \right). \quad (6.6)$$

Формирование длительности фронта заканчивается, когда ток коллектора станет равным $I_K(t_\Phi) = 0,9 I_{K.HAC.}$. Тогда (6.6) можно записать:

$$0,9 I_{K.HAC.} = I_B \beta_0 \left(1 - e^{-t_\Phi/\tau_\beta} \right). \quad (6.7)$$

Решив (6.7) относительно длительности фронта, получим

$$t_\Phi = \tau_\beta \ln \frac{I_B \beta_0}{I_B \beta_0 - 0,9 I_{K.HAC.}}. \quad (6.8)$$

Анализируя (6.8), можно сделать выводы, что длительность фронта уменьшается, если выбрать более высокочастотный транзистор (τ_β), увеличивать ток базы или уменьшать $I_{K.HAC.}$.

При этом необходимо помнить, что если τ_β не изменяет величину заряда в области базы, то изменения параметров входной и выходной цепей влияют на количество объёмного заряда. Используя степень насыщения (5.8), можно записать:

$$t_\Phi = \tau_\beta \ln \frac{s}{s - 0,9}. \quad (6.9)$$

Переходные процессы в насыщенном ключе при закрытии транзистора. Пусть в момент времени t_0 напряжение

входного генератора с отрицательной полярности изменилось на положительную. Ток базы скачком изменится на величину ΔI_B и станет отрицательным. Отрицательное направление тока обусловлено тем, что в базе у эмиттерного и коллекторного переходов имеется избыточный заряд, который мгновенно рассосаться не может, и оба перехода смещены в прямом направлении. Ток коллектора остаётся практически постоянным, ток эмиттера уменьшается на ΔI_B .

К моменту времени t_1 избыточный заряд рассасывается у коллекторного перехода, и с этого момента уменьшаются токи коллектора и эмиттера. При t_2 ток коллектора становится равным $0,1I_{K.НАС.}$. Время изменения коллекторного тока от t_1 до t_2 называют длительностью фронта t_{ϕ^-} .

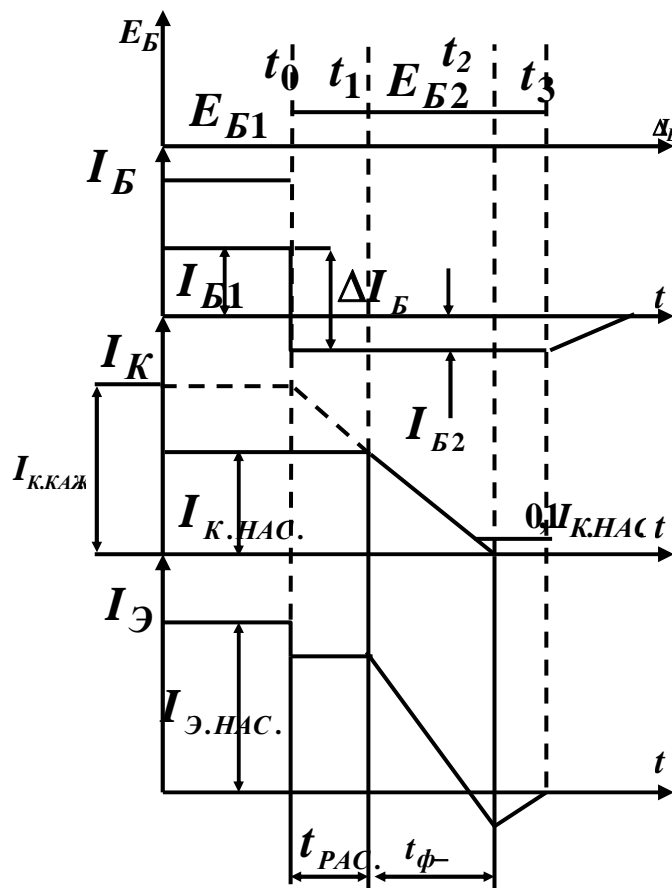


Рис. 6.7 — Переходные процессы в насыщенном ключе при закрывании транзистора

К моменту времени t_2 избыточные заряды полностью рассасутся у эмиттерного перехода, и он будет смещаться в обратном направлении, токи базы и эмиттера уменьшаются. Видимо, чем больше величина избыточного заряда в базе, тем больше время рассасывания.

Для анализа процессов рассасывания и формирования фронта воспользуемся уравнением (6.5), но учтем, что начальный ток отличен от нуля.

Подставив ток ΔI_B и заменив τ_β на τ_H — постоянная времени накопления, получим изображение коллекторного тока

$$I_K(p) = I_B \beta_0 - \frac{\Delta I_B}{p} \beta(p), \quad (6.10)$$

которому соответствует оригинал:

$$I_K(t) = I_{B1} \beta_0 - \Delta I_B \beta_0 \left(1 - e^{-t/\tau} \right). \quad (6.11)$$

Рассасывание закончится, когда избыточный заряд в базе исчезнет, и будет иметь место равенство $I_{K.КАЖ.} = I_{K.НАС.}$. Подставив $I_{K.НАС.}$ в левую часть (6.11) и заменив τ на τ_H , найдем время рассасывания в виде:

$$t_{PAC} = \tau_H \ln \frac{\Delta I_B \beta_0}{\Delta I_B \beta_0 - I_{KH.} (S - 1)}. \quad (6.12)$$

Из формулы (6.12) видно, что увеличение степени насыщения приводит к росту времени рассасывания, а длительность фронта при этом наоборот уменьшается.

Для сокращения суммарного времени переходного процесса необходимо уменьшать время рассасывания и длительность фронта.

Методы повышения быстродействия рассмотрим позже.

Формирование длительности отрицательного фронта ($t_{\phi-}$)

Формирование фронта заканчивается, когда коллекторный ток будет равен $0,1 I_{K.НАС.}$.

Учитывая, что к началу формирования среза $I_{B1} = I_{НАС.}$ и заменив τ на τ_β , т.к. длительность фронта формируется, когда

транзистор работает в активном режиме, формулу (6.11) можно записать в виде

$$0,1I_{KH} = I_{BH}\beta_0 - \Delta I_B\beta_0 \left(1 - e^{-t_\phi/\tau_\beta} \right). \quad (6.13)$$

Решив (6.13) относительно длительности t_ϕ^- , получим

$$t_\phi^- = \tau_\beta \ln \frac{\Delta I_B\beta_0}{\Delta I_B\beta_0 - 0,9I_{KH}}. \quad (6.14)$$

Анализ переходного процесса проводился без учета ёмкости коллекторного перехода. Учесть ёмкость коллекторного перехода можно, если в (6.9) и (6.14) заменить τ_β на $\tau_\beta + C_K R_K (1 + \beta_0)$.

На принципиальной схеме (рис. 6.8) предполагается, что источник сигнала E_B генерирует импульсы двух полярностей.

В реальном случае импульсы однополярные. В этом случае лучше применять принципиальную схему, которая приведена на рис. 6.8.

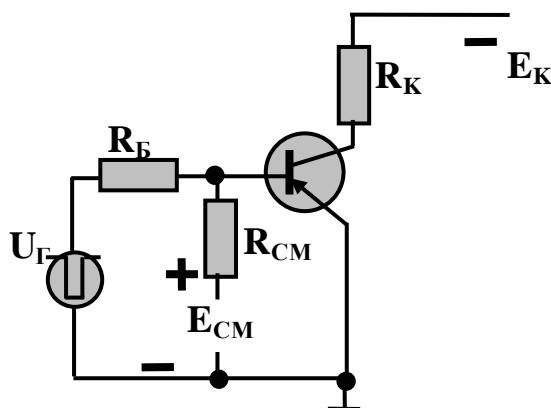


Рис. 6.8 — Принципиальная схема насыщенного ключа

6.4 Методы сокращения времени переходного процесса

Ключ с ускоряющей ёмкостью. Для сокращения времени переходного процесса на всех его стадиях используется ключ с

ускоряющей ёмкостью, принципиальная схема которого приведена на рис. 6.9 (R_{Γ} — внутреннее сопротивление источника сигнала).

Время переходного процесса можно сократить в несколько раз, если выполняется условие $R_B \gg R_{\Gamma}$.

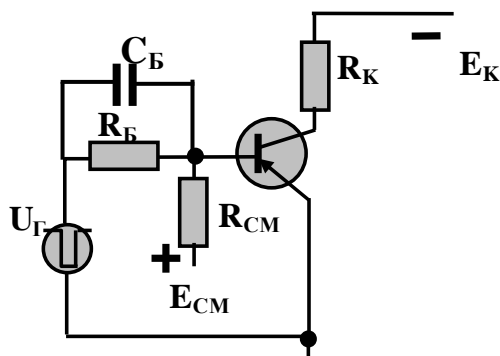


Рис. 6.9 — Ключ с ускоряющей ёмкостью

В момент отпирания ключа вместо тока

$$I_B = \frac{U_{\Gamma}}{R_{\Gamma} + R_B},$$

в первый момент получается гораздо больший ток

$$I_{B..мак.} = \frac{U_{\Gamma}}{R_{\Gamma}}.$$

Учитывая, что $R_B \gg R_{\Gamma}$, имеем $I_{B..мак.} \gg I_B$.

Конечно, по мере заряда конденсатора ток базы уменьшается, но фронт и срез формируются током, близким к максимальному значению. В то же время величина избыточного заряда в базе соответствует значительно меньшему току.

Применяя теорему об эквивалентном генераторе, изображение тока базы запишем в виде:

$$I_B(p) = I_{B0} \frac{1 + p\tau_C}{1 + p\tau_C^1},$$

где

$$I_{B0} = \frac{U_{\Gamma}}{R_{\Gamma} + R_B};$$

$$\tau_C = CR;$$

$$\tau_C^1 = \frac{R_{\Gamma}R_B}{R_{\Gamma} + R_B};$$

$$\tau_C^1 \ll \tau_C.$$

Подставляя $I_B(p)$ в формулу (6.5), получим изображение тока коллектора:

$$I_K(p) = I_{B0}\beta_0 \frac{1 + p\tau_C}{(1 + p\tau_\beta)(1 + p\tau_C^1)}.$$

Оригинал этой функции при $\tau_C > \tau_\beta$ имеет одиночный выброс, что часто бывает нежелательно. Поэтому постоянную времени τ_C выбирают из условия $\tau_C = \tau_\beta$. Тогда

$$I_K = I_{B0}\beta_0 \frac{1}{p(1 + p\tau_C^1)}. \quad (6.15)$$

Сравнивая (6.5) и (6.15), видим, что они совпадают до постоянных времени. Можно показать, что и при запертии ключа изображения будут отличаться только постоянными времени.

Учитывая, что $\tau_\beta \gg \tau_C$, длительности фронтов в насыщенном ключе существенно сокращаются, а время рассасывания стремится к нулю.

Недостатком схемы с ускоряющей ёмкостью является наличие конденсатора, который трудно реализовать при интегральной технологии.

Ненасыщенный ключ. Простейший вариант ненасыщенного ключа с отрицательной нелинейной обратной связью показан на рис. 6.10.

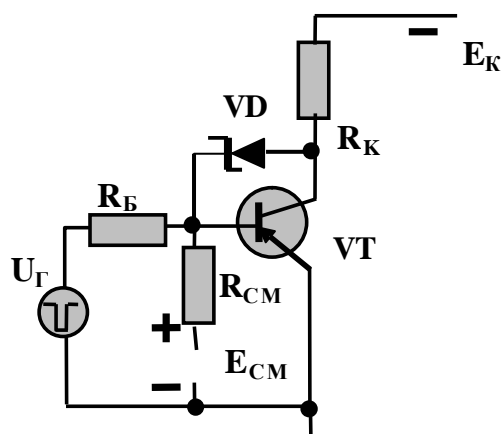


Рис. 6.10 — Ключ с отрицательной нелинейной обратной связью

При достаточно больших коллекторных напряжениях диод VD смещен в обратном направлении и его сопротивление велико. Цепь обратной связи разорвана. Когда напряжение на коллекторе станет близким к нулю, диод откроется и приращения коллекторного тока в противофазе с током генератора будут поступать в цепь базы, что приведёт к уменьшению инжекции со стороны эмиттера.

Следовательно, в базе не будут накапливаться избыточные носители и время рассасывания окажется равным нулю, а формирование фронта и среза происходит при больших токах базы.

Вопросы для самопроверки

1. Начертите принципиальную схему насыщенного ключа.
2. Начертите статическую вольт-амперную характеристику ключа и отметьте на ней точки, соответствующие режиму насыщения и отсечки.
3. Начертите эквивалентную схему ключа, соответствующую режиму отсечки, и расставьте токи, протекающие в транзисторе.
4. Начертите эквивалентную схему ключа в режиме насыщения и расставьте токи, протекающие в транзисторе в этом режиме.
5. Запишите выражения, определяющие мощность, рассеиваемую на транзисторе в режимах насыщения и отсечки.
6. Объясните причины, приводящие к конечному времени переходного процесса в насыщенном ключе.
7. Начертите эпюры напряжений и токов при открывании транзистора.
8. Объясните, почему ток базы в начальный момент времени имеет максимальное значение, а затем уменьшается до установившегося значения.
9. От чего зависят токи транзистора в режиме насыщения? Запишите их значения.
10. Запишите формулу длительности фронта при открывании транзистора.
11. С помощью каких параметров можно изменять длительность фронта в насыщенном ключе?

12. Нарисуйте эпюры напряжений и токов в насыщенном ключе при запираании транзистора.

13. Почему при запираании транзистора ток базы становится отрицательным?

14. Выведите выражение для длительности фронта при запираании транзистора.

15. Запишите формулу для времени рассасывания избыточного заряда при запираании транзистора.

16. Зависит ли длительность отрицательного фронта импульса от степени насыщения?

17. Как можно уменьшить время рассасывания в насыщенном ключе.

18. Нарисуйте принципиальную схему ключа с ускоряющей ёмкостью.

19. Объясните, почему включение конденсатора позволяет сократить время переходного процесса.

20. Запишите выражение изменения коллекторного тока в ключе с ускоряющей емкостью и докажите, что время переходного процесса сокращается.

21. Можно ли уменьшить время переходного процесса в ключе с ускоряющей емкостью, если источник сигнала является генератором тока?

22. Начертите схему ключа с отрицательной нелинейной обратной связью.

23. Объясните работу ключа с нелинейной обратной связью.

24. Почему включение диода Шотки в цепи отрицательной обратной связи позволяет сократить время переходного процесса в ключе?