

А. Н. БУБЕННИКОВ

МОДЕЛИРОВАНИЕ ИНТЕГРАЛЬНЫХ МИКРО- ТЕХНОЛОГИЙ, ПРИБОРОВ И СХЕМ



Допущено
Государственным комитетом СССР
по народному образованию
в качестве учебного пособия
для студентов вузов,
обучающихся по специальности
«Физика и технология
материалов и компонентов
электронной техники»



Москва
Высшая школа 1989

ББК 32.844.1

Б90

УДК 621.382

Рецензенты: кафедра проектирования и конструирования ИМС Московского института электронной техники (зав. кафедрой — д-р техн. наук, проф. Г. Г. Казеннов); кафедра микроэлектроники Московского инженерно-физического института (зав. кафедрой — д-р техн. наук, проф. А. В. Шальнов)

Бубенников А. Н.

Б90 Моделирование интегральных микротехнологий, приборов и схем: Учеб. пособие для спец. «Физика и технология материалов и компонентов электронной техники». — М.: Высш. шк., 1989. — 320 с.: ил.

ISBN 5-06-000123-7

В книге рассмотрены математические модели интегральных микротехнологий, приборов и схем; описаны основные принципы построения моделей; технологических, физико-топологических, схемных, элементарно-фрагментных и методы идентификации их параметров; рассмотрены основы математического обеспечения численного расчета БИС.

Б 2304040000(4309000000)—440 185—89
(001)01—89

ББК 32.844.1

6Ф0.3

ISBN 5-06-000123-7

© Издательство «Высшая школа», 1989

В решениях XXVII съезда КПСС содержится концепция генерального стратегического направления — ускорения социально-экономического развития страны на основе научно-технического прогресса. К приоритетным направлениям, определяющим суть научно-технического потенциала СССР, отнесены разработка и широкое внедрение в производство перспективной элементной базы высокопроизводительных ЭВМ высших поколений, а также методов и средств автоматизации бездефектного проектирования элементной базы.

Разработка и ускоренное внедрение в производство перспективной элементной базы, в первую очередь больших интегральных схем (БИС) с высокими быстродействием и степенью интеграции, в значительной степени определяются созданием систем автоматизированного проектирования (САПР) и подготовкой высококвалифицированных инженерных кадров — разработчиков БИС, в совершенстве владеющих инструментом машинного моделирования интегральных микротехнологий, приборов и схем.

Поскольку современный этап развития интегральной электроники связан с переходом к субмикронным микротехнологиям и приборам, определяющим уже на начальном уровне разработки основные технико-эксплуатационные характеристики БИС и ЭВМ, требования к научному вооружению высококвалифицированного специалиста новой формации в области разработки и машинного моделирования БИС резко возрастают.

В основу изложения материала положен комплексный системный подход многоуровневого моделирования БИС от технологии их изготовления до функционирования приборов и схем, с необходимым описанием соответствующих моделей, алгоритмов, программного обеспечения, методов идентификации, а также оптимизации элементов и фрагментов БИС.

Ознакомление широкого круга студентов и аспирантов радиоэлектронных и инженерно-физических специальностей с методами сквозного моделирования БИС позволит сократить расстояние от изобретения и модификаций перспективных интегральных микротехнологий, приборов, микросхемотехники до внедрения и промышленного освоения элементной базы ЭВМ.

Автор выражает признательность коллегам по лаборатории моделирования МФТИ и ИПК АН СССР за помощь в обсуждении учебного пособия, в первую очередь канд. техн. наук А. Д. Садовникова, совместно с которым написаны § 3.3 и 5.6.

Отзывы о книге можно направлять по адресу: 101430, Москва, ГСП-4, Неглинная ул., д. 29/14, издательство «Высшая школа».

Требования к элементно-технологической базе высокопроизводительных ЭВМ 4-го и 5-го поколений — быстродействующим БИС — обуславливают актуальность дальнейшего поиска наиболее эффективных структурно-технологических решений реализации устройств ЭВМ с максимально высокими уровнями интеграции, быстродействия, выхода годных схем.

Повышение эффективности разработок быстродействующих кремниевых БИС, улучшение их технико-эксплуатационных показателей в настоящее время определяются исключительно автоматизацией проектирования БИС.

Решение ряда острейших проблем разработки быстродействующих кремниевых БИС требует применения новых методов автоматизированного проектирования, основой которых является математическое моделирование всех этапов изготовления и разработки БИС, объединенных в некоторую «супер-САПР» БИС. Одной из основных подсистем «супер-САПР» БИС является сквозное электронное моделирование БИС, представленное согласованными друг с другом уровнями: 1) технологическим; 2) физико-топологическим; 3) схмотехническим; 4) функционально-логическим.

В настоящее время в руках разработчиков БИС появился инструмент теоретического и практического исследования беспрецедентного по своей сложности объекта (БИС из 10^3 — 10^7 активных элементов) методами математического моделирования, тонкого вычислительного эксперимента на ЭВМ.

Парадокс вычислительного эксперимента с БИС заключается в том, что детализация объекта исследования связана со значительным усложнением математической формулировки и методов решения. Если на верхнем функционально-логическом уровне БИС из миллионов элементов описывается простейшими логико-временными соотношениями (типа элементарных булевых уравнений или соотношений для регистровых передач), а на схмотехническом уровне БИС из сотен и тысяч элементов — законами Кирхгофа в алгебраической и дифференциальной формах (линейные алгебраические и обыкновенные дифференциальные уравнения), то на нижних физических уровнях описание процессов в полупроводнике с помощью в общем случае нелинейных дифференциальных уравнений в частных производных в двух- и трехмерном приближении представляет качественно другой, значительно более сложный и тонкий тип описания объекта исследования.

Основной этап каждого уровня моделирования — математическая формулировка задачи исследования, выбор математической модели объекта. Выбору математической модели объекта предшествует очень важный процесс выбора физических приближений, ранжирования учитываемых и пренебрегаемых факторов, уточнения, иерархии моделей в зависимости от типов (и количества) различных режимов эксплуатации объекта. Объект исследования на нижних физических уровнях — полупроводниковый прибор, интегральная неоднородная микроструктура со сложной геометрией — при его полном детальном многомерном математическом описании в виде сложных нелинейных дифференциальных уравнений в частных производных с соответствующим многообразием граничных и начальных условий не может быть полностью исследован с приемлемыми вычислительными затратами на современных высокопроизводительных ЭВМ.

Более того, погоня за излишней сложностью и детализацией моделей структур в отрыве от возможностей современных ЭВМ может привести к практически бесполезному машинному «перемалыванию» объемов информации, оторванному от реальных задач разработки элементарно-технологической базы. Без дальнейшего экспериментального подтверждения приемлемой адекватности выходных характеристик и параметров каждого уровня моделирования бессмысленно использовать формальный математический аппарат решения исходных уравнений моделей.

Подчеркнем первостепенную важность формулировки и синтеза математических моделей микротехнологии изготовления приборов, собственно прибора и схем из активных и пассивных полупроводниковых структур. В формулировке математической модели с учетом разумного компромисса «точность — сложность» модели объекта указывается необходимая и достаточная, в целом ограниченная группа входных структурных параметров и группа всех требуемых выходных параметров. Между группами параметров задается связь в виде алгебродифференциальных уравнений, а также информация о соответствующих граничных и начальных условиях (для замкнутости сформулированной задачи). Задачу синтеза математической модели объекта в общем виде, вероятно, бессмысленно поручать ЭВМ, а целесообразно максимально использовать интеллект исследователя с допуском к самой разнообразной информации о технологии изготовления БИС, физике и электронике приборов и схем. При этом в работе интеллектуального разработчика подразумеваются не только действие сдерживающих факторов, компромиссов, но и выявление новых закономерностей, связей искомым и базовых переменных модели, фундаментализация их описания.

После составления замкнутых систем уравнений, описывающих электрофизический процесс, исследования математической модели методами общей теории алгебраических, интегродифференциальных уравнений необходимо установить возможность эффективного единственного решения, алгоритмизации задачи, оптимизации вычисли-

тельных аспектов проблемы. Взаимосвязь формальных и неформальных методов в синтезе модели и ее расчете — узел всех проблем адекватного математического моделирования и дальнейшей оптимизации параметров сложных объектов. Может оказаться, что математическая модель слишком груба, результаты расчетов не согласуются с физическими экспериментами (которые тоже проводятся с существенными погрешностями). С другой стороны, может оказаться, что модель слишком сложна и решение с достаточной степенью точности может быть получено с помощью более простых моделей при весьма экспериментальном подтверждении. Необходимость согласования и стыковки всех этапов моделирования, даже в пределах одного уровня (технологического или физико-топологического), подразумевает действие обратных связей: расчет — экспериментальное подтверждение — новый расчет.

Отметим, что при бурном развитии технологии, новых физических принципов приборов, систем- и микросхемотехники БИС наблюдается процесс быстрой модификации моделей и качества расчетов и даже достаточно быстрого их «морального старения».

В свете новых требований к ускорению научно-технического прогресса, к уровню подготовки специалистов — инженеров-физиков, инженеров-конструкторов в области разработки перспективной элементной базы — на первый план выдвигаются актуальные проблемы дальнейшего повышения интеллектуализации этапов машинного моделирования и проектирования БИС, «усиления интеллекта» разработчика, конструктора сложнейших объектов интегральной электроники. Современный специалист в области микротехнологии, физики полупроводниковых приборов, микросхемотехники, вооруженный методологией сквозного многоуровневого моделирования БИС, знанием возможностей вычислительной техники и ее математического обеспечения, должен решать задачи оптимизации БИС, выявления возможностей дополнительного увеличения их быстродействия, степени интеграции, выхода годных схем, снижения стоимости, сложность которых оказалась бы не по плечу специалистам, взявшимся за решение этих задач «вручную».

ПУТИ И ПЕРСПЕКТИВЫ РАЗВИТИЯ КРЕМНИЕВОЙ ИНТЕГРАЛЬНОЙ ЭЛЕКТРОНИКИ И СКВОЗНОГО МНОГОУРОВНЕВОГО МОДЕЛИРОВАНИЯ БИС

1.1. КРЕМНИЕВАЯ КОМПЬЮТЕРНАЯ ТЕХНОЛОГИЯ

Вехи интегральной микроэлектроники. Ни в одной из областей цивилизации немислим прогресс без микроэлектроники и вычислительной техники, одних из самых перспективных научно-технических направлений. Темпы развития микроэлектроники, высокопроизводительных ЭВМ, систем обработки информации, широта и эффективность их внедрения во все области человеческой деятельности не имеют аналогов в истории науки и техники.

Конец 40-х годов — изобретение биполярного транзистора (БТ), внедрение в промышленное производство дискретных БТ.

Конец 50-х — начало 60-х годов — изобретение транзистора со структурой металл — окисел — полупроводник (МОПТ); разработка первых интегральных микросхем (ИМС) со степенью интеграции до нескольких единиц — десятков компонентов в кремниевом кристалле.

Конец 60-х годов — в интегральных схемах средней степени интеграции (СИС) достигнуто размещение около 10^2 активных компонентов — БТ или МОПТ.

Середина 70-х годов — степень интеграции больших интегральных схем (БИС) превысила 10^3 компонентов.

Начало 80-х годов — степень интеграции в сверхбольших интегральных схемах (СБИС) выросла до 10^5 активных компонентов.

Середина 80-х годов — степень интеграции в СБИС и ультра-БИС превысила 10^6 — 10^7 активных компонентов.

Прогноз до 2000 г. — микроэлектронные высокопроизводительные средства обработки и передачи информации будут органически вплетены во все области человеческой деятельности, в развитии цивилизации появится новый тип инфраструктуры — информационный, резко увеличится производительность труда в научно-практической деятельности и будет обеспечен значительный социально-экономический прогресс.

Развитие кремниевых БИС, СБИС, сверхскоростных больших и сверхбольших интегральных схем (ССБИС, С³БИС), элементно-технологической и конструктивной базы высокопроизводительных ЭВМ 4-го и 5-го поколений в конце 70-х — середине 80-х годов происходило чрезвычайно бурно. Основной проблемой разработки БИС,

ССБИС, С³БИС является совершенствование технологии, схемотехники для достижения противоречивых требований: увеличения системного быстродействия, степени интеграции, выхода годных схем.

Изготовление кремниевых БИС произвело так называемую информационную революцию, не уступающую по важности индустриальной революции. Информационная революция существенно отличается от индустриальной использованием бурно развивающейся компьютерной технологии по сравнению с ранее известной технологией, направленной на механическое взаимодействие с окружающей средой.

Одной из характерных особенностей развития БИС и С³БИС, в первую очередь логических, матричных, стало резкое обострение проблем межсоединений элементов внутри БИС. Проблемы межсоединений стали настолько острыми, что современные и тем более перспективные С³БИС следует рассматривать уже не как совокупность активных компонентов, взаимодействующих через межсоединения, а скорее как созданную в рамках компьютерной кремниевой технологии систему сигнальных проводников, возбуждаемую активными компонентами. Тем не менее тот непреложный факт, что активный компонент (БТ и МОПТ), преобразующий информационные сигналы за счет своих усилительных и переключательных способностей, является по-прежнему основной составной частью БИС, есть ключевая особенность, по которой кремниевую компьютерную технологию выделяют среди других индустриальных технологий. Теснейшая связь между активным компонентом (БТ и МОПТ), технологией его изготовления и технико-эксплуатационными параметрами схем является сутью компьютерной технологии всех прогрессивных разработок С³БИС с субмикронными размерами, ориентированных на достижение предельного быстродействия и степени интеграции элементно-технологической и конструктивной базы высокопроизводительных ЭВМ.

Первый БТ, изобретенный в 1948 г. сотрудниками американской фирмы «Бэлл» Бардином и Брэттенем и теоретически описанный Шокли, не был еще твердотельным прибором в современном смысле слова: он представлял собой ненадежную конструкцию из фрагментов полупроводникового кристалла и прижатых к нему электродов. Однако в процессе развития микротехнологии полупроводниковых приборов и БИС, объединив достижения методов литографии, легирования, локальной защиты поверхности кристалла, нанесения пленок, окисления, эпитаксии и др., микротехнология при формировании сотен и тысяч БТ и МОПТ стала групповой, высокопроизводительной, а параметры интегральных приборов — строго воспроизводимыми.

В настоящее время значимость микротехнологий в производстве БИС особенно велика, поскольку бурный прогресс и совершенствование микротехнологий характеризует достижение качественно новых реализаций микроэлектронных устройств с максимально высо-

ким выходом годных схем, уникальными электрическими характеристиками, возможностями обеспечения сверхинтеграции и сверхбыстродействия. Прогресс субмикронной микроэлектроники основан на совершенствовании микротехнологий БИС за счет использования новых методов изготовления сверхчистых материалов, бездефектных кристаллов, значительного сокращения многочисленных химических операций по очистке поверхности, уменьшения числа технологических операций, в первую очередь высокотемпературных и «мокрых», перехода к качественно новым, перспективным процессам, в основе которых лежит взаимодействие излучений и потоков частиц с поверхностью кристалла (электронная, ионная, рентгеновская литография, а также ионное легирование, лазерный отжиг, молекулярно-лучевая эпитаксия и др.).

Роль кремния в компьютерных технологиях. Название «кремний» (silicon) происходит от лат. silicis — кремний. Кремний — наиболее распространенный после кислорода элемент на Земле — составляет около 26% земной коры, его кристаллы имеют структуру кристаллической решетки алмаза. Благодаря уникальной способности термически выращенного слоя двуокиси кремния играть роль маски на этапах изготовления ИМС, возможности использования в микротехнологиях нитрида кремния в качестве высококачественного пассивирующего слоя, а также вследствие низкой температурной чувствительности и высокой технологичности процессов кремний находит исключительно широкое применение в производстве БИС, С³БИС.

Известно, что для каждого уровня технологии и заданного типа схем существует оптимальная степень интеграции, при которой групповое серийное производство наиболее экономично. Например, в середине 80-х годов широкими сериями выпускаются БИС со степенью интеграции до 10^4 элементов, малыми сериями — 10^5 — 10^6 элементов и уникальной серией — свыше 10^6 — 10^7 . Стандартная ширина линий в БИС в серийном производстве составляет около 2—2.5 мкм, в самых лучших образцах — менее 1 и даже 0.5 мкм. Однако при прогрессивной тенденции увеличения уровня интеграции БИС и уменьшения горизонтальных (вертикальных) размеров элементов вероятность того, что хотя бы один элемент или небольшой фрагмент схемы, а с ними вся БИС уйдут в брак, резко увеличивается при минимальном выходе годных схем. При этом производительность серийных БИС падает, брак может превышать 99% продукции, увеличиваются расход материалов, стоимость и время проектируемых разработок. Такая ситуация вполне реальна при промышленном освоении перспективных БИС и С³БИС со степенью интеграции 10^5 — 10^6 компонентов (10^4 — 10^5 вентиляей). Однако уже при степени интеграции 10^6 — 10^7 компонентов технологические, функционально-эксплуатационные, надежность экономические проблемы переплетаются настолько тесно, что их невозможно решать независимо друг от друга. Сколь ни совершенны будут производство и уровень схемно-технологиче-

ской реализации С³БИС со степенью интеграции 10^6 — 10^8 компонентов, вероятность возникновения дефектов в отдельных элементах, ошибок синтеза, выхода их из строя в ходе эксплуатации будет очень большой. Для сложнейших систем С³БИС, ориентированных на обработку огромной информации, главными критериями становятся совершенство внутренней организации схем, технологичность материалов, экономичность промышленного освоения, где кремний, особенно в диапазоне нормальных температур ($-30 \div +150^\circ\text{C}$), пока не знает конкурентов.

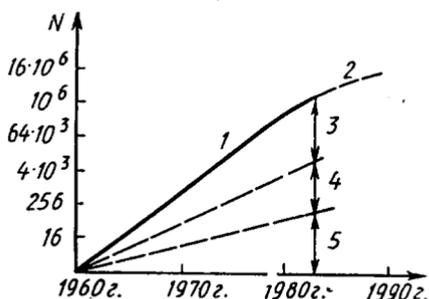


Рис. 1.1. Тенденции роста числа компонентов на кристалле БИС:

1 — двукратное увеличение за год; 2 — двукратное увеличение за два года; 3 — усовершенствование системо- и схемотехнических решений; 4 — увеличение размеров кристалла; 5 — уменьшение литографических размеров

В научно-технической литературе неоднократно приводилась любопытная аналогия: не случайно кремний — ближайший сосед и аналог углерода; углерод оказался основным элементом биологических систем в ходе эволюции природы и жизни на Земле в течение миллионов лет, кремний стал основным элементом компьютерной технологии в ходе непрерывной по времени технoэволюции.

Таким образом, даже при условии бурного прогресса и огромных потенциальных возможностей технологий изготовления БИС на

основе полупроводниковых соединений А³В⁵, особенно при низкотемпературных условиях эксплуатации, кремниевая технология останется наиболее распространенной промышленной технологией, достаточно живучей и экономически оправданной. Несмотря на то что кремниевая субмикронная микротехнология близко подошла к пределу своего потенциального быстродействия, в рамках системного рассмотрения фактора «степень интеграции — быстродействие — стоимость, выход годных схем» перспективные кремниевые БИС, СБИС и С³БИС на БТ и МОПТ будут конкурентоспособны БИС, созданным на основе некремниевой технологии, и еще длительное время будут сохранять доминирующее положение в аппаратных средствах высокопроизводительных ЭВМ и систем обработки информации.

1.2. ПЕРСПЕКТИВЫ ПОВЫШЕНИЯ СТЕПЕНИ ИНТЕГРАЦИИ И БЫСТРОДЕЙСТВИЯ КРЕМНИЕВЫХ БИС, ИХ ОГРАНИЧЕНИЯ И ПРЕДЕЛЫ

Переход от ИМС к БИС и СБИС происходил на протяжении последних 20 лет. В качестве параметра, количественно иллюстрирующего этот процесс, используем ежегодное изменение числа элементов N , размещаемых на одном кристалле (что соответствует сте-

пени интеграции) (рис. 1.1). Как показано на рис. 1.1, степень интеграции, например, для запоминающих устройств с достаточно регулярной структурой с 1960 г. практически до 1980 г. возрастала пропорционально времени. При этом до 1975 г. данный процесс характеризовался двукратным увеличением степени интеграции за год. В то же время уже с начала 80-х годов наблюдается тенденция некоторого снижения темпов роста интеграции БИС и СБИС, т. е. повышение степени интеграции в четыре раза за три года или двукратное — за два года.

Как видно из рис. 1.1, процесс повышения степени интеграции (увеличения быстродействия вентилях) БИС был обусловлен действием трех факторов: 1) усовершенствованием системной и схемотехники решений; 2) увеличением размеров кристаллов; 3) микроминиатюризацией топологических решений за счет уменьшения ширины линии рисунка и горизонтальных размеров элементов. Например, с 1960 по 1975 г. степень интеграции возросла более чем в 64 000 раз. Вследствие усовершенствования структур разрабатываемых устройств и разработки новых схемотехнических решений степень интеграции увеличилась в 100 раз, размеры кристаллов — в 20 раз, а плотность размещения элементов благодаря использованию прогрессивных методов микролитографии — в 32 раза.

Если проанализировать усовершенствование структуры ячеек кремниевых БИС (памяти и логических), то можно прийти к заключению, что достигнутые уровни интеграции и быстродействия уже близки к предельным и дальнейшее повышение степени интеграции будет происходить главным образом за счет микроминиатюризации БИС и ее элементов, а также улучшения системной и схемотехники. Увеличение роста кристаллов после превышения некоторой критической площади около 10×10 мм пока сопряжено с ростом их дефектности и резким снижением выхода годных схем. В этом направлении проводятся исследования, включая разработки многокристалльных систем, однако не исключено, что для экономически оправданного промышленного производства БИС разработка кристаллов площадью более 20×20 мм нецелесообразна. В то же время в области повышения степени интеграции в будущем прилагаются усилия по совершенствованию методов изоляции, микролитографии при создании уникальных поверхностных конфигураций, а также новых технологических приемов сверхсамосовмещения, увеличения плотности компоновки элементов.

Изоляция элементов схемы. Фундаментальной проблемой, имеющей непосредственное отношение к совершенствованию технологии изготовления БИС и СБИС, является изоляция их элементов.

До середины 70-х годов повышения степени интеграции и быстродействия биполярных БИС достигали с помощью различных технологических методов межэлементной изоляции за счет перехода от изоляции обратносмещенными *p-n*-переходами к изоляции окислом, что позволило существенно уменьшить размеры БТ и увеличить плотность компоновки БИС. На рис. 1.2 отражена тенденция умень-

шения площади БТ при переходе к изоляции окислом и уменьшении литографических размеров. При этом сокращение размеров БТ позволило существенно уменьшить внутренние барьерные и внешние нагрузочные емкости в БИС, а следовательно, в десятки раз уменьшить потребляемую мощность схем при сохранении или увеличении их быстродействия. В биполярных БИС наиболее широкое применение получила комбинированная изоляция, сочетающая изоляцию диэлектриком и p - n -переходом (рис. 1.2).

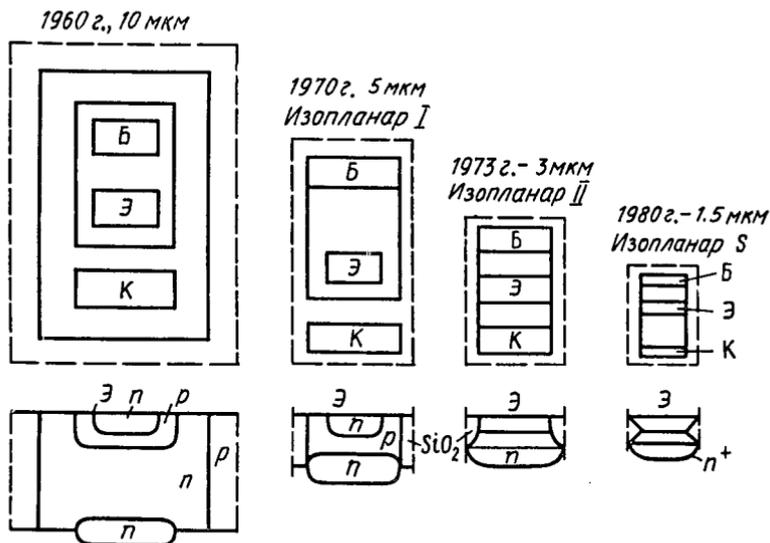


Рис. 1.2. Уменьшение площади БТ при различных способах изоляции и уровнях литографии

Радикальным технологическим средством уменьшения площади эпитаксиальных БТ является их изоляция в БИС глубокими канавками. Такие канавки, имеющие окисные стенки с диэлектрическими (SiO_2) или поликремниевыми наполнителями, изолируют друг от друга скрытые n^+ -слои БТ, обеспечивая требуемые высокие пробивные напряжения БТ и уменьшая площадь БТ примерно в 4—5 раз при одинаковых литографических размерах. На рис. 1.3, а—в показаны поперечные сечения сверхсамосовмещенных интегральных БТ с глубокими изолирующими канавками (вертикального типа n - p - n и горизонтального типа p - n - p), а также резистора. Как видно из рисунка, для изоляции БТ используются канавки глубиной 3—6 мкм, шириной 1 мкм, заполненные, например, окислом, который получают химическим осаждением из газовой фазы. Канавки несколько сдвинуты относительно краев углубленного окисла, проходят сквозь подколлекторный n^+ -слой, изолируя коллекторы БТ. Использованием канавок для перерезания скрытого n^+ -слоя

удаётся не только уменьшить ёмкость подложки и площадь БТ, но и увеличить пробивное напряжение между коллекторами изолированных БТ.

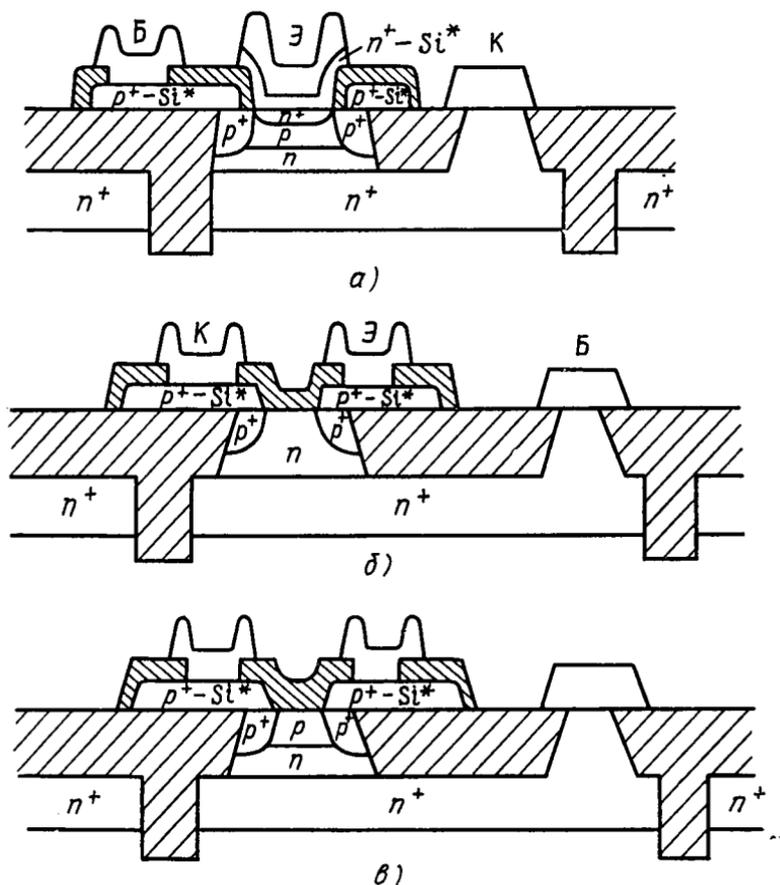


Рис. 1.3. Поперечные сечения интегральных БТ, изготовленных методом сверхсамосовмещения с глубокими изолирующими канавками:

а — вертикального БТ типа *n-p-n*; б — горизонтального БТ типа *p-n-p*; в — резистора

Таким образом, при реализации субмикронной технологии применение диэлектрической изоляции элементов является довольно эффективным средством повышения степени интеграции и быстродействия устройств, выполненных на СБИС и С³БИС.

Сверхсамосовмещение и использование поликремниевых слоев. Технологические приемы самосовмещения и сверхсамосовмещения полупроводниковых областей интегральных структур для миними-

зации пассивных и паразитных областей, уменьшения горизонтальных размеров структур успешно развиваются одновременно в МОП- и биполярной технологии изготовления СБИС и С³БИС. Горизонтальные размеры приборов можно значительно уменьшить, используя свойство поликристаллического кремния образовывать более тонкие, чем при окислении монокристаллического кремния, пленки окисла. На этом основан метод самосовмещения областей поликремниевых затвора, истока и стока МОПТ, а также базовых и эмиттерных областей БТ. Такое применение поликристаллического кремния (поликремния) становится ключевым методом уменьшения размеров и улучшения параметров активных элементов БИС.

Применение поликремния позволяет также уменьшить глубины переходов приборов, не опасаясь дефектов типа прокола (например, в БТ алюминием от металлического контакта), что приводило к уменьшению выхода годных схем. Кроме того, в биполярной технологии характерной особенностью структуры области эмиттера — база (рис. 1.3) является независимость от минимального литографического размера процесса изготовления областей эмиттера и базы, когда площадь эмиттера определяется уже не площадью контакта и даже не соответствующим эмиттерным окном. Очень важно, что из поликремния делают высококачественные резисторы, при этом отпадает необходимость в разделительной области для изоляции между резистором и транзистором и, следовательно, уменьшается паразитная емкость резистора. Диапазон регулирования сопротивления поликремниевых резистора значительно шире, чем диффузионного резистора.

Рассмотренные примеры технологического воплощения сверхсамосовмещенных БТ и аналогичных тенденций в МОП-технологии БИС свидетельствуют об уникальных особенностях данного технологического направления, позволяющего дополнительно повышать степень интеграции и быстродействие фрагментов БИС по сравнению с возможностями улучшения данных параметров только средствами микролитографии — уменьшения геометрических размеров рисунка на поверхности кристалла.

Ограничения и пределы развития кремниевых БИС. Причина, заставляющая уменьшать размеры элементов и повышать степень интеграции, одна — снижение стоимости БИС и всей разрабатываемой аппаратуры на БИС. Однако со снижением стоимости вычислений и (или) хранения информации при росте степени интеграции улучшаются также рабочие характеристики элементов и схем, технико-эксплуатационные параметры собственно БИС. Следовательно, снижается стоимость не просто вычислительных средств, а более качественных вычислительных средств.

Однако если бы темпы микроминиатюризации — двукратное увеличение степени интеграции за два года — сохранились до 2000 г., то размеры БТ или МОПТ уменьшились бы до размеров вируса, число элементов в СБИС превысило бы триллион, а размеры линий литографического рисунка оказались бы в несколько со-

тен раз меньше размеров атома. Нереальность этих выводов по прогрессу уровня интеграции БИС, а также перспектива уменьшения системного быстродействия БИС с ростом интеграции заостряют проблему текущих ограничений и фундаментальных пределов развития БИС.

К ограничениям, преодолеваемым на пути развития сверхинтегрированных быстродействующих БИС, следует отнести три группы взаимосвязанных факторов: 1) сложность и высокую стоимость технологии субмикронных БИС; 2) сложность и значительные сроки проектирования, верификации БИС; 3) экономические проблемы обеспечения выхода годных схем при разработках субмикронных БИС с высокими (оптимальными) технико-эксплуатационными параметрами.

Диалектические противоречия, присущие разработкам перспективных БИС, связанные с постоянным преодолением внутренних ограничений в ходе создания опытных образцов и дальнейшего массового производства БИС, обуславливают «естественный экономический отбор». В конечном итоге экономичны только те технологические процессы, методы, структуры и схемы, физика и функционирование которых допускают надежное массовое производство.

Геометрические пределы. Прогресс микроминиатюризации будет уменьшать геометрические размеры компонентов до тех пор, пока есть возможность активно и целенаправленно регулировать изменения размеров с целью достижения лучших электрических параметров прибора. Однако должен наступить такой критический момент, когда некоторые фундаментальные геометрические размеры прибора не будут поддаваться изменению при любой сколь угодно полезной модификации технологического процесса, направленной на кардинальное уменьшение геометрических размеров структуры. К таким фундаментальным геометрическим размерам следует отнести толщину p - n -перехода, которая зависит от концентрации легирующих примесей и не зависит, например, от литографических методов создания поверхностных конфигураций полупроводниковой структуры. Толщина обедненного слоя p - n -перехода 0.025 — 0.03 мкм будет определять минимальные геометрические размеры сверхмикроминиатюризованных гипотетических приборных структур. Для оценки минимальных площадей приборов требуется учитывать размеры физических областей, необходимых для функционирования приборов при статистическом рассмотрении ансамбля подвижных носителей, пересекающих области полупроводника. В качестве оценки минимальной площади прибора примем величину $S_{\min} = 1 \text{ мкм}^2$; элементарный логический вентиль с коэффициентами объединения и разветвления 4, даже в сверхинтегрированных элементах, требует как минимум двух площадей прибора $S_{\text{в}} \approx 2 \text{ мкм}^2$ (при допущении, что область изоляции приборов друг от друга очень мала, все резисторы — поликремниевые типа и находятся над поверхностью активного элемента). Тогда максимальная плотность компоновки (геометрическая интеграция) $D_{\max} = 1/S_{\text{в}} = 5 \times$

$\times 10^7$ вент/см² $\approx 10^8$ эл/см². Достигнутый уровень интеграции промышленных БИС отстает от уровня геометрической интеграции на полтора два порядка.

Быстродействие и размеры кристалла БИС. Максимальная тактовая частота устройства. Единицей производительности ЭВМ является скорость выполнения арифметических, логических или других операций. Производительность универсальных ЭВМ оценивают по формуле $W_{\text{пр}} \approx 1/(0.7 t_{\text{к}} + 0.3 t_{\text{д}})$, где $t_{\text{к}}$, $t_{\text{д}}$ — времена выполнения «коротких» и «длинных» операций; 0,7, 0,3 — коэффициенты, учитывающие относительную частоту выполнения этих операций при реализации типовых алгоритмов. Времена выполнения различных операций существенно отличаются. Например, для «коротких» операций (сложение или вычитание двоичных чисел) требуется обычно 4—5 циклов (тактов машинного времени), для «длинных» операций (умножение или деление) — несколько десятков циклов. При этом можно оценить, что отношение длительности цикла выполнения операции к времени переключения простейшего логического элемента (ЛЭ) (вентиля) равно примерно 12—20.

Для получения длительности цикла перспективной однопроцессорной ЭВМ около 1 нс нужны ЛЭ с временем переключения ≈ 50 пс. Но увеличение быстродействия элементов ЭВМ не является еще полным решением проблемы. Поскольку за 1 нс сигнал по полосковой линии связи проходит расстояние всего 10 см, для достижения такой малой длительности цикла процессор необходимо разместить в минимально возможном объеме (~ 400 см³), чтобы не ухудшить системное быстродействие. Рассматриваемый пример заостряет проблемы оценки реального быстродействия элементов БИС с учетом размеров кристалла, а также минимизации потребляемой мощности элементов в БИС высокой степени интеграции.

Найдем связь предельной степени интеграции БИС и тактовой частоты вычислительного устройства. Скорость передачи сигналов по внутренним линиям связи в БИС $v_c \approx (5 \div 10) \cdot 10^9$ см/с. Самое длительное время распространения сигнала в системе определяет ее тактовую частоту f_c . На кристалле БИС максимальное расстояние L_{max} — половина его периметра; следовательно, полная площадь $A_{\text{кр}} = L_{\text{м}}^2/4$, а максимальный уровень интеграции (максимальное число вентилях или компонентов, которое можно разместить на площади кристалла) $I_{\text{м}} = A_{\text{кр}} D_{\text{м}}$. Для размеров кристалла 1 см² $v_c = 10^{10}$ м/с; $f_c = 5 \cdot 10^8$ Гц; для увеличения частотного диапазона до 1 ГГц необходимо уменьшить размеры кристалла до 0,25 см² с размещением 10^8 вент/крист; для $f_c = 10$ ГГц $A_{\text{кр}} = 0.25$ мм², $I_{\text{м}} = 6.25 \cdot 10^4$ вент/крист. В последних двух случаях времена переключения БТ или МОПТ должны быть меньше 200 и 20 пс соответственно.

Предельное соотношение быстродействие — потребляемая мощность. Уменьшая размеры элементов БИС, разработчики борются не только за повышение степени интеграции, но и за их высокое быстродействие, и, следовательно, не могут пропорционально раз-

мерам снижать мощность, поскольку при этом увеличивается системное время задержки вентиля (уменьшается рабочая частота вычислительных устройств). Когда плотность компоновки БИС становится большой, кристалл не справляется с отводом теплоты. Чем выше требуемое быстродействие элементов БИС, тем выше рассеиваемая мощность, ниже плотность компоновки и степень

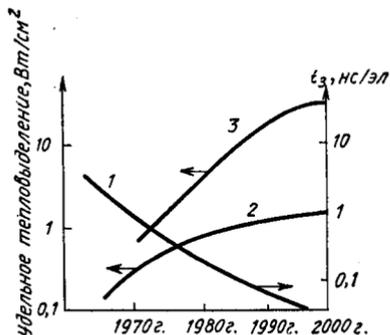


Рис. 1.4. Прогноз эволюции системного быстродействия (1) и удельного тепловыделения при воздушном (2) и жидкостном (3) охлаждении

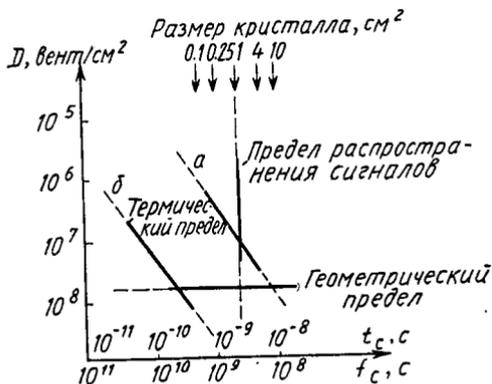


Рис. 1.5. Диаграмма быстродействие — потребляемая мощность элементов (уровень интеграции БИС) при наличии трех пределов разработки БИС (геометрического, термического и распространения сигналов) для двух технологий с отводом теплоты с единицы площади кристалла: $Q=20$ Вт/см² (кривая а), $Q=1000$ Вт/см² (кривая б)

интеграции. На рис. 1.4 представлены эволюционные характеристики системного быстродействия вентилях БИС и удельного тепловыделения при воздушном (1—2 Вт/см²) и жидкостном (10—20 Вт/см²) охлаждении. Недавно была предложена уникальная конструкция охлаждения, использующая поток жидкости непосредственно вдоль каналов, вытравленных на обратной стороне кристалла. Данная схема охлаждения позволяет эффективно отводить с кристалла теплоту до 1000 Вт/см² при обеспечении нормального функционирования С³БИС.

Для оценки предельного соотношения мощность — быстродействие сверхинтегрированных БИС воспользуемся тем фактом, что энергия в элементе рассеивается преимущественно от зарядки или разрядки эффективной внутренней емкости, присущей прибору в БИС (это справедливо для достаточно малых токов переключения). Выражение для эквивалентной емкости $C = \epsilon \epsilon_0 L^2/d$, где ϵ , ϵ_0 — относительная и абсолютная диэлектрические проницаемости; $S = L^2$ — площадь; d — толщина диэлектрика при пропорциональной миниатюризации, может быть преобразовано в равенство $C \approx \Lambda L$, где Λ — константа пропорциональности, связанная с величиной

ее; обычно для инженерных оценок $\Lambda \approx 10^{-10}$ Ф/см. Для времени цикла t_c и логического перепада $U_{\text{л}}$ мощность, рассеиваемая элементом на внутренней емкости,

$$p \approx \Lambda U_{\text{л}}^2 L / t_c. \quad (1.1)$$

При ограничении на удельное тепловыделение с возможностью отвода теплоты с единицы площади кристалла Q должно выполняться соотношение

$$p / L^2 < Q. \quad (1.2)$$

Из (1.1) и (1.2) следует неравенство

$$p t_c^2 > (\Lambda U_{\text{л}}^2)^2 / Q, \quad (1.3)$$

позволяющее для сверхвысокой плотности компоновки элементов в БИС связать основные эксплуатационные параметры p , t_c , Q , $U_{\text{л}}$. Термическое ограничение (1.3) можно трактовать следующим образом. Энергия, приложенная к схеме, используется для зарядки или разрядки эквивалентной внутренней емкости, причем емкость пропорциональна линейному размеру L . В процессе микроминиатюризации энергия на переключение может быть уменьшена пропорционально линейному размеру L . В то же время количество вентиля (элементов) увеличивается обратно пропорционально L^2 и мощность, рассеиваемая на единице площади, обратно пропорциональна L при неизменной длительности цикла. В конечном итоге плотность рассеиваемой мощности становится больше скорости отвода теплоты, что отражено термическим пределом $p t_c^2 = \text{const}$ (рис. 1.5). При этом константа $(\Lambda U_{\text{л}}^2)^2 / Q$ определяет качество элементов БИС (для улучшения разработки БИС следует уменьшать $U_{\text{л}}$ и увеличивать Q).

Взаимодействие трех типов пределов разработки БИС — геометрического, термического и распространения сигналов — иллюстрируется рис. 1.5 для двух уровней технологии: современного ($Q = 20$ Вт/см²) и перспективного ($Q = 1000$ Вт/см²). Увеличение максимальной (термически ограниченной) плотности интеграции $D_{\text{т}} = Q/p$ с ростом Q позволяет добиваться получения одновременно больших тактовых частот и уровня интеграции разрабатываемых БИС, причем для небольших размеров кристалла системное быстроедействие вентиля определяется быстрымдействием активных компонентов, а для больших размеров кристалла становятся доминирующими задержки в соединениях. Из рассматриваемых предельных характеристик БИС видно, что для развития технологии, системо- и схемотехники имеется значительный диапазон прогресса. Однако получение максимального быстрогодействия и степени интеграции БИС, в первую очередь логических СЗБИС, будет несколько сдерживаться влиянием межэлементных линий связи на кристалле, а также получением экономических и эффективных конструкций отвода теплоты от кристалла.

Кремниевые С³БИС и нейронные системы мозга. Из сравнительных предельных зависимостей рис. 1.5 видно, что ресурсы быстрого действия и интеграции в кремниевых БИС далеко не исчерпаны. Можно утверждать, что сложившаяся ветвь развития кремниевых БИС, СБИС и С³БИС будет интенсивно развиваться в XX и даже в XXI вв. Однако рассмотренные пределы развития кремниевых БИС свидетельствуют не только о грядущем диапазоне прогресса, но и об удивительном системном проигрыше венца техноэволюции — кремниевой компьютерной технологии — венцу живой природы —

Таблица 1.1

Сравнительные характеристики кремниевых логических элементов БИС и нейронов

Наименование параметра	БИС			Нейрон
	1986 г.	1990 г.	2000 г.	
Задержка распространения сигнала на элемент (нейрон), с	$100 \cdot 10^{-12}$	$30 \cdot 10^{-12}$	$10 \cdot 10^{-12}$	$2 \cdot 10^{-4}$
Рассеиваемая мощность, Вт	10^{-3}	10^{-4}	10^{-4}	10^{-9}
Энергия переключения, пДж	10^{-13}	$3 \cdot 10^{-15}$	10^{-15}	$2 \cdot 10^{-13}$
Линейный размер, м	10^{-5}	$3 \cdot 10^{-6}$	10^{-6}	10^{-5}
Интенсивность (число) отказов, отказ/ч	$2 \cdot 10^{-9}$		10^{-9}	$5 \cdot 10^{-9}$
Плотность упаковки, м ³				10^{13}
Число атомов, см ⁻³		$10^{17} - 10^{20}$		$2 \cdot 10^{13}$
Режим обработки информации	Параллельно-последовательная			Высокая степень параллелизма и избыточности

человеческому мозгу. В табл. 1.1 приведены сравнительные характеристики кремниевых логических элементов БИС и нейронов мозга по ряду основных параметров. Видно, что природный живой элемент значительно превосходит элементы кремниевых С³БИС по таким параметрам, как рассеиваемая мощность и плотность упаковки. Элементы кремниевых С³БИС на шесть порядков превышают нейроны по быстродействию и за счет этого по энергии переключения.

Парадокс сравнения состоит в том, что нейронные сети, проигрывая на шесть порядков в быстродействии кремниевым С³БИС (нейроны переключаются в миллион раз медленнее современных БТ и МОПТ, и во столько же раз медленнее электромагнитных сигналов в С³БИС распространяются сигналы по нервным волокнам), превосходят по всем системным параметрам логические элементы кремниевых С³БИС за счет фантастического совершенства внутренней организации живой материи. Возможно, главные тай-

ны работы мозга еще не раскрыты, однако адаптивная сущность нейрона как порогового логического элемента, связанного нервными волокнами с сотнями и тысячами своих соседей, позволяет за счет высочайшей степени параллелизма и избыточности, оперирования информацией в цифровой и аналоговой формах, при относительно медленной передаче и обработке информации каждым нейроном получать эффективные быстрые решения сложнейших проблем. Коэффициент заполнения нейронного логического элемента много меньше 10—20% и, следовательно, время, в течение которого нейронная система мозга потребляет энергию, составляет ничтожную часть. Вследствие удивительной экономии расхода энергии мозга, содержащего $\sim 10^{10}$ нейронов (10—20 Вт)*, удается получить высочайшую эффективность и производительность «решающего устройства».

Очевидно, идеальная цель интегральной и молекулярной электроники будущего, симбиоза кремниевых С³БИС с максимальным приближением к предельным параметрам и биосистем — это создание систем, сочетающих быстрдействие и экономичность кремниевой технологии с совершенством организации живого мозга.

1.3. ПРОБЛЕМЫ МНОГОУРОВНЕВОГО МОДЕЛИРОВАНИЯ И ОСОБЕННОСТИ ФИЗИКО-ТЕХНОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ КРЕМНИЕВЫХ БИС

Необходимость внедрения комплексных гибких систем автоматизированного проектирования (САПР) в разработку и производство С³БИС вызвана огромной сложностью, значительными сроками и стоимостью проектирования, а также спецификой субнаносекундного и пикосекундного диапазонов, не характерной для ИМС и БИС невысокого быстрдействия.

Развитие современных комплексных САПР кремниевых БИС и С³БИС, включающих на верхних уровнях проектирования мощные программные средства верификации, логического синтеза, логиковременного расчета, автоматизированного размещения, трассировки, а также схемотехнического анализа, немислимо без соответствующего развития математического обеспечения физико-технологического моделирования основных активных элементов БИС—БТ и МОПТ. Структурная схема сквозного многоуровневого моделирования кремниевых БИС приведена на рис. 1.6, а, б. При многоуровневом моделировании ИМС и БИС последовательно производят расчет характеристик и параметров на следующих уровнях: технологическом (Т), физико-топологическом (ФТ), электрическом и фрагментном. В общем случае предусмотрен программный переход от более низких уровней к высоким с помощью программ оптимиза-

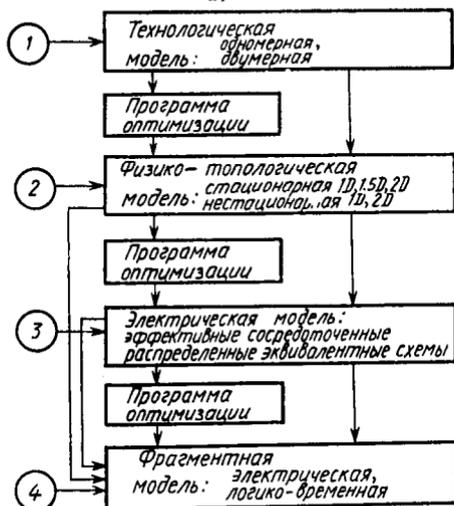
* При расценке 1 кВт·ч=2 коп. решение задач колоссальной сложности интеллектуального исследователя за 8 ч рабочего времени стоит всего $10 \cdot 10^{-3} \times 8 \cdot 2 = 0.16$ коп.

ции по среднеквадратическому отклонению модельных характеристик соответствующих уровней.

С помощью Т-моделей сначала рассчитывают результирующий профиль примеси в активных и пассивных компонентах ИМС по технологическим параметрам (имплантации, диффузии и др.). Этот



а)



б)

Рис. 1.6. Взаимодействие программ многоуровневого моделирования в рамках «супер-САПР» БИС:

а — с последовательной передачей выходной информации; б — с использованием программ оптимизации для определения параметров модели последующего уровня расчета БИС

профиль вместе с основными электрофизическими параметрами кремния и топологией прибора является входной информацией для ФТ-моделирования.

С помощью ФТ-моделей рассчитывают основные электрические характеристики БТ и МОПТ: ВАХ, емкости, сопротивления, пробив-

ное напряжение коллекторного перехода и др. На этом уровне с помощью специальных методик идентифицируют модели приборов или макромоделей ЛЭ.

С помощью электрических моделей БТ и МОПТ на следующем уровне рассчитывают цепи ЛЭ из нескольких десятков—сотен элементов, с помощью физических макромоделей ЛЭ—сети БИС из нескольких сотен ЛЭ, а с помощью простейших функциональных макромоделей—из тысяч ЛЭ.

В результате прогресса технологических методов изготовления С³БИС, уменьшения минимальных литографических размеров до субмикронных резко усилилась взаимосвязь между электрическими параметрами схемы, поведением и характеристиками приборных структур и технологическими режимами изготовления С³БИС. При отладке современных и разработке перспективных технологий БИС ведущая роль отводится машинному физико-технологическому моделированию с последующим схмотехническим расчетом БИС.

Актуальность проблем, решаемых на этапах физико-технологического и схмотехнического расчетов кремниевых БИС, обуславливается многими факторами проектирования разработки элементно-технологической и конструктивной базы ЭВМ. Приведем эти факторы.

1. Многообразие и противоречивость схмотехнических и технологических принципов одновременного повышения степени интеграции и быстродействия устройств ЭВМ заостряет проблему выбора и точного расчета новых схмотехнических решений ЛЭ и технологий БИС.

2. Увеличение сроков и стоимости проектирования и изготовления БИС, невозможность внесения исправлений в готовую БИС и неоправданные затраты на совершенствование компонентов ЛЭ и фрагментов БИС с помощью многочисленных «тестовых итераций» делают эффективной и экономически оправданной методологию детального сквозного многоуровневого моделирования: технологического, физико-топологического, электрического и функционально-логического.

3. Отсутствие у разработчиков БИС тестовых образцов и ЛЭ, полученных экспериментально параметров моделей и макромоделей МОПТ и ЛЭ сдерживает процесс качественного проектирования БИС или вынуждает их использовать плохо обусловленные модели и параметры, что обесценивает зачастую результаты расчета и оптимизации в рамках автоматизированного компонентного или функционально-логического моделирования, делая крайне актуальными методы машинной идентификации параметров МОПТ и ЛЭ БИС.

4. Переход к тонкослойным элементам с микротопологией субнаносекундных БИС и С³БИС, неадекватность и значительная погрешность аналитических и упрощенных численных моделей формирования профилей реальным распределениям примесей делают не-

обходимым многоэтапный расчет профиля примесей на основе достаточно точных моделей технологических операций в одно- и двумерном приближениях.

5. Увеличение степени легирования полупроводниковых областей требует учета в ФТ-моделях зависимости основных электрофизических параметров кремния от концентрации и совокупности эффектов высокого легирования: сужения и асимметрии сужения запрещенной зоны, Оже-рекомбинации (наряду с ШРХ — механизмом Шокли — Рида — Холла), рассеяния на подвижных носителях заряда и др. Вследствие микротопологии и высоких плотностей рабочих токов нельзя использовать аналитические и численные модели БТ, основанные на разбиении БТ на квазинейтральные области и области объемного заряда.

6. Требование повышения степени интеграции и быстродействия БИС приводит к возрастанию роли совокупности эффектов высокого уровня инжекции и топологических особенностей БТ и МОПТ ЛЭ, что эффективно может быть отражено в точных численных ФТ-моделях и далее в электрических моделях.

7. Тенденция уменьшения горизонтальных и вертикальных размеров БТ БИС затрудняет экспериментальную идентификацию инерционных параметров БТ: времен задержек эмиттер — коллектор и коллектор — эмиттер, барьерных и диффузионных емкостей. Кроме машинных методов идентификации данных параметров актуальны экспериментальное тестирование и контроль на промежуточных и заключительных этапах разработки БИС.

8. Высокое качество проектирования БИС (БТ и МОПТ) возможно только при глубоком машинном исследовании корреляции технологических и электрических параметров БТ и ЛЭ, оптимизации профиля примеси и топологии БТ, МОПТ и фрагментов БИС с учетом разводки.

9. Требования к конструкции компонентов без учета электрического режима БИС, влияния реальных конструкций и связей в БИС приводит к недоиспользованию технологических возможностей, неоптимальности БТ и МОПТ в подсистеме БИС и даже к браку.

Относительно небольшие временные и стоимостные затраты на комплексное физико-технологическое моделирование элементов БИС, создание электрических моделей и эквивалентных схем элементов для последующего сквозного схемотехнического расчета больших фрагментов С³БИС делают этап физико-технологического моделирования основным средством проектирования. Это связано с тем, что за время, затрачиваемое на создание первой работающей приборной структуры, параметры которой можно исследовать экспериментально, имеется возможность детально проанализировать множество вариантов технологического процесса и различных топологий приборных структур БИС матричного или заказного типа. Ценной с точки зрения оптимизации создаваемых конструктивных решений приборных структур и предварительного планирования проведения необходимых экспериментальных исследований в ряде

случаев является возможность генерирования различных, иногда самых неожиданных для инженерной мысли конкурентоспособных вариантов проектных решений.

Процесс физико-технологического моделирования эффективен еще потому, что измерение электрических и электрофизических параметров полупроводниковых структур является довольно сложным, трудоемким и зачастую характеризуется существенной погрешностью. Вследствие этого использование методов физико-технологического моделирования для идентификации электрических параметров приборных структур может оказаться неоценимым в установлении обратной связи (характеристики прибора — технология его изготовления), перекрестной проверке результатов натурного и машинного экспериментов при разработке и изготовлении.

У машинного сквозного физико-технологического моделирования элементов БИС есть четыре характерные особенности, которые и в настоящее время, и в перспективе будут обеспечивать постоянный успех применения моделирования технологии и приборных структур при изготовлении БИС: точность, скорость, системная интегрируемость и огромные сервисные возможности. Точное определение параметров профиля примесей и электрических параметров приборных структур является важнейшим фактором и приобретает доминирующее значение по мере уменьшения горизонтальных и вертикальных размеров сверхминиатюрных элементов БИС.

Основная предпосылка успешного применения методов физико-технологического моделирования субмикронных элементов БИС — достаточно точное предсказание физических ограничений и пределов, критических функциональных характеристик и параметров для конечной цели проектирования и (что еще важнее) реального изготовления структур.

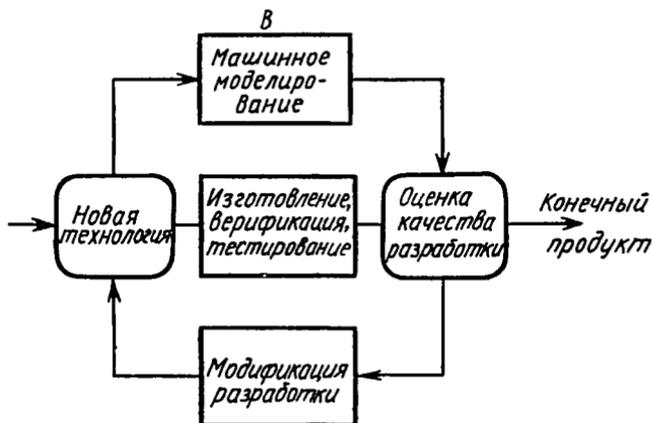
Фактор скорости методов сквозного физико-технологического моделирования имеет несколько важных аспектов. Во-первых, технологи и разработчики приборных структур и схем должны иметь реальную и оперативную «обратную связь», которая и позволила бы им удешевить использование технологического оборудования. Физико-технологическое моделирование не заменяет изготовления тестовых БТ и МОПТ — оно только уменьшает количество тестов для получения и модернизации стабильного технологического процесса. Более того, физико-технологическое моделирование позволяет исследовать всевозможные допустимые отклонения в технологии и выявить доминирующие факторы, которые могут влиять на процесс изготовления БИС, выход годных схем. Во-вторых, имеется растущая возможность использования технологических моделей в качестве сложных эффективных контроллеров технологического оборудования. Хотя этот аспект физико-технологического моделирования только входит в практику, необходимое условие его применения — возможность проведения моделирования в реальном масштабе времени средствами вычислительной техники, используемы при управлении технологическим процессом.

Входной язык для пользователя средств машинного физико-технологического моделирования — важнейший фактор широкого применения этих средств при разработке и изготовлении БИС. Взаимное слияние сеток для моделирования технологий и приборных структур заметно облегчит работу пользователя, сделает ее более эффективной без заметного ухудшения эффективности численных методов на различных этапах моделирования. Несомненно, что разработка автоматических алгоритмов взаимодействия технологических и физико-топологических программ, повышающих эффективность численных методов и одновременно улучшающих входной язык сквозного физико-технологического моделирования, является наиболее предпочтительным направлением развития.

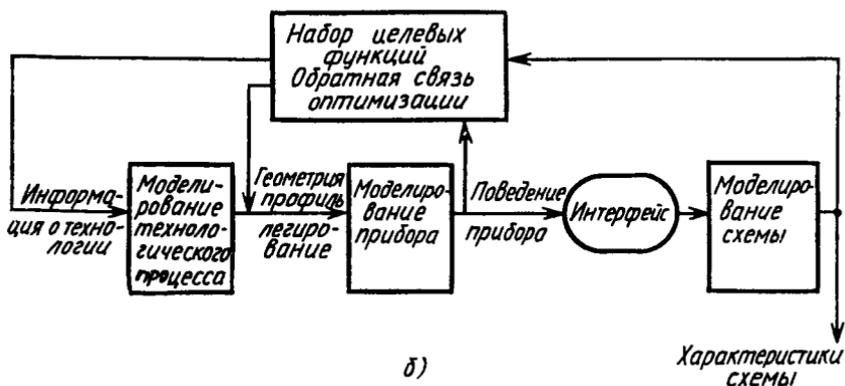
Последняя существенная черта методов машинного проектирования технологий и приборных структур — системная интегрируемость — связана с выработкой требований к технике взаимодействия, моделям, алгоритмам, расчетным характеристикам всех уровней моделирования для осуществления с минимальными вычислительными затратами подхода «от нажатия кнопки до конечного результата» при проектировании, изготовлении, верификации и тестировании элементов БИС. Подобная интеграция методов физико-технологического проектирования вызывается актуальной необходимостью все более жесткого контроля субмикронных, тонко-слоистых приборных структур. На системном уровне растущая необходимость управления процессом производства требует высокой степени интеграции различных программ моделирования технологии и приборных структур, представляет собой многообещающую перспективную альтернативу традиционному методу. Сказанное проиллюстрировано на рис. 1.7, *а* контуром *В*.

Обычно стоимость изготовления одной партии (12 пластин) по усовершенствованному технологическому процессу колеблется от 10 до 50 тыс. долларов, причем на изготовление такой партии может потребоваться от нескольких недель до нескольких месяцев. В то же время при наличии соответствующих САПР с программным обеспечением физико-технологического моделирования элементов численные эксперименты по технологии изготовления БИС могут быть выполнены за несколько часов со значительно меньшими затратами. Разработка технологии и проектирование схем, изготавливаемых по этой технологии, могут вестись параллельно, причем результаты, полученные при машинном моделировании, позволяют глубже понять физику технологического процесса, а также работу приборов и схем, представляющих неоценимую информацию о субмикронных БИС, экспериментальное исследование которых сопряжено с огромными сложностями и затратами.

Взаимодействие уровней моделирования микротехнологий изготовления приборов и схем проиллюстрировано на рис. 1.6, *а*, *б*, 1.7, *а*, *б*, где схематически представлены различные программные средства, используемые для разработки технологии производства С³БИС и связи между отдельными уровнями.



А
а)



б)

Рис. 1.7. Взаимосвязь моделирования и оптимизации технологии изготовления, приборных структур и БИС:

а — альтернативные пути разработки технологии и САПР; контур А представляет экспериментальный подход, контур В — машинное моделирование технологий приборов, схем; б — пути оптимизации характеристик приборов, схем посредством многовариантных расчетов с целенаправленной модификацией параметров технологии, топологии, электрических режимов и других факторов

Использование программ многоуровневого моделирования БИС приходит на смену традиционным методам имитационного и экспериментального макетирования, поскольку такой подход позволяет точно прогнозировать характеристики БИС, дает информацию, которую нельзя получить путем лабораторных измерений. Открывается перспектива эффективного машинного исследования корреляций технологических и электрических параметров элементов и схем, оптимального проектирования С³БИС, получения их предельных степени интеграции и быстродействия для достигнутого уровня технологии, схемотехники и системотехники.

МОДЕЛИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ИЗГОТОВЛЕНИЯ БИС

2.1. ПОКОЛЕНИЯ ПРОГРАММ ТЕХНОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ БИС

В течение последних лет, прошедших с момента сообщения (конец 70-х годов) о первой высокоэффективной инженерной программе одномерного моделирования технологических процессов изготовления БИС SUPREM II, данный уровень моделирования БИС стремительно развивался. Сегодня уже можно говорить о нескольких поколениях программ технологического моделирования. К первому поколению относятся упоминавшаяся программа SUPREM II, а также большое количество зарубежных и отечественных программ, так или иначе использующих и развивающих основные идеи и модели, заложенные в программе SUPREM II. Основное внимание в этих программах уделялось моделированию процессов ионного легирования, диффузии, окисления и эпитаксии, ответственных за распределение примесей в полупроводниковых структурах, как правило, в одномерном приближении. Остальные операции, в частности создание поверхностных конфигураций структур, моделировались крайне упрощенно или не моделировались вообще. Возможности технологических программ первого поколения отражали состояние технологии БИС конца 70-х — начала 80-х годов, для которых параметры приборов определялись в основном распределением в глубь полупроводниковых структур, глубины залегания переходов были порядка 0.5—1.5 мкм, толщины термически выращенных окислов составляли 1.0—0.5 мкм, а процессы формирования поверхностных конфигураций являлись не столь критическими. На этой стадии была показана несомненная польза от применения указанных программ и получена приемлемая точность расчетов по данным моделям.

Стремительный прогресс в кремниевой технологии в последние 5—10 лет инициировал дальнейшее развитие технологического моделирования. Многие из появившихся недавно программ обладают признаками второго поколения, для которых характерны следующие возможности:

- 1) моделирование процессов обработки многослойных структур: кремния (Si), поликремния (Si*), SiO₂, Si₃O₄, силицидов, металлов и т. д.;

- 2) двумерное моделирование процессов окисления и диффузии в прямоугольной области;
- 3) более точное моделирование процесса ионного легирования;
- 4) сокращение времени расчетов за счет применения сложных численных методов;
- 5) двумерное моделирование практически всех операций полного технологического цикла изготовления ИМС: ионного легирования, диффузии, окисления, эпитаксии, литографии, травления и осаждения пленок;
- 6) богатые сервисные возможности, модульный принцип построения; удобный входной язык, рассчитанный на неподготовленного пользователя, возможность использования на ЭВМ различных типов.

Все эти требования к программам второго поколения вызваны современным состоянием технологии БИС, которая характеризуется в первую очередь предельным уменьшением размеров приборных структур за счет перехода к низкотемпературному отжигу, использования многослойных структур, рентгеновской и электронной литографии и других перспективных процессов. При этом резко возрастает роль двумерности распределения профиля примеси, а также процессов формирования поверхностных конфигураций, которые стали влиять на характеристики приборов. Сложность задач двумерного многоэтапного моделирования заставляет акцентировать внимание на вычислительных аспектах, а необходимость постоянного использования программ разработчиками новых приборов — на сервисных возможностях; коммуникабельности программного обеспечения (легкости общения пользователя с ЭВМ).

В ближайшее время можно ожидать появления первых программ с некоторыми признаками третьего поколения, которые будут отражать дальнейшие тенденции в развитии технологии БИС. Несомненно, что кроме аспектов, отраженных в п. 1—6, они будут характеризоваться усовершенствованием кинетических механизмов в модифицированных неравновесных моделях для расчета распределения примесей при быстропотекающих термических и лазерных отжигах, тонким учетом точечных дефектов как основного кинетического механизма миграции примесей, еще более точным и сложным моделированием процессов литографии, кинетики осаждения и травления пленок исходя из «первых принципов».

2.2. МОДЕЛИ ПРОЦЕССОВ ИОННОГО ЛЕГИРОВАНИЯ

Легирование полупроводника примесями проводится с целью создания различных приборных структур за счет изменения его электрофизических свойств: типа электропроводности, удельного сопротивления и других характеристик.

Реализованные и потенциальные преимущества ионного легирования позволяют: осуществлять процесс с высокой производительностью; создавать практически любые профили распределения за

счет ступенчатого легирования; совмещать процесс легирования с другими технологическими процессами поверхностной обработки кристалла; получать прецизионное формирование профиля полупроводниковых приборных структур. С другой стороны, ионное легирование имеет недостатки и ограничения. Есть определенные трудности в проведении процесса легирования, связанные с нарушениями, созданными ионной бомбардировкой, и окончательным местоположением внедренных ионов. Как правило, необходимо устранить эти нарушения в виде смещенных из узлов кристаллической решетки атомов полупроводниковой мишени и в то же время сделать внедренные атомы примеси электрически активными. Обычно это достигается частичным или полным отжигом кристаллов сразу же после ионного легирования при умеренных температурах. К другим ограничениям следует отнести трудность создания и воспроизведения глубоких легированных областей, сложность обработки больших полупроводниковых пластин из-за расфокусировки при существенных отклонениях ионных пучков.

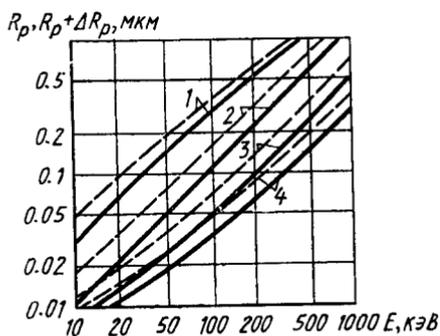


Рис. 2.1. Зависимости средней проекции пробега R_p (сплошные линии) и истинной проекции $R_p + \Delta R_p$ с учетом ее случайного разброса (пунктирные линии) при имплантации в кремниевую подложку ионов бора (кривые 1), фосфора (кривые 2), мышьяка (кривые 3) и сурьмы (кривые 4) от их энергий

Большое число регулирующих параметров процесса ионного легирования (доза, тип, энергия ионов, температура и среда отжига и др.) позволяют в широких пределах изменять свойства легированных слоев, но наряду с этим требуют глубокого физического понимания процессов внедрения ионов, их поведения в кристаллической решетке, кинетики образования и устранения радиационных дефектов, что необходимо для высококачественного технологического моделирования в конечном итоге эффективной реализации приборных структур и схем в интегральном исполнении.

Физические основы ионного легирования. Процесс ионного легирования заключается в ионизации и ускорении до больших скоростей атомов примеси. Эффективная масса иона в 10^3 — 10^5 больше массы электрона, поэтому при заданной энергии ион имеет импульс, в 10^2 — 10^4 раз превышающий импульс электрона. Ускоренные атомы примеси внедряются в кристаллическую решетку полупроводника под воздействием приобретенного импульса. Проникая в кристаллическую решетку, ионизированный атом примеси постепенно теряет кинетическую энергию за счет взаимодействия с электронами и упругих столкновений с атомами полупроводника и примеси, т. е. в результате электронного и ядерного торможения.

Траекторий атома примеси до полной остановки представляет собой ломаную линию, характеризует его пробег R , а проекция этого пути на направление падения первичного ионного пучка — проекцию пробега R_p . Профиль легирования обычно определяется проекцией пробега и ее случайного разброса ΔR_p . На рис. 2.1 приведены зависимости R_p и $R_p + \Delta R_p$ от энергии E внедренных в кремний ионов бора, фосфора, мышьяка и сурьмы. В предположении, что полупроводник ориентирован относительно направления ионного пучка, значительно отличающегося от направлений главных кристаллографических осей, результирующие профили легирования примесей близки к распределению Гаусса. При этом для изотропной или аморфной полубесконечной мишени в упрощенной классической модели ЛШШ (Линхарда — Шарфа — Шитта) концентрация легирующей примеси зависит от расстояния от поверхности полупроводника:

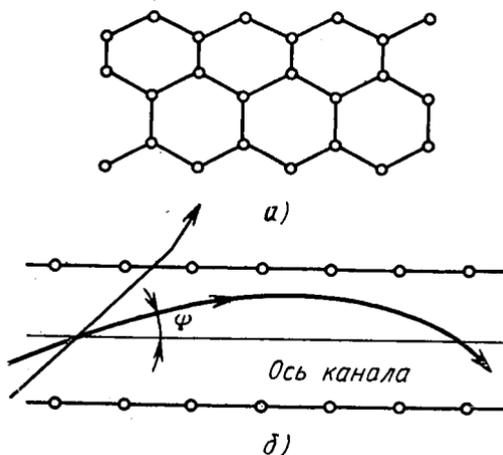


Рис. 2.2. Иллюстрация эффекта каналирования в кремниевых структурах:

a — проекция структуры на плоскость $\langle 110 \rangle$; *б* — возможные траектории ионов в зависимости от направления входа в канал

$$C(x) = C_{\max} \exp \left[-\frac{(x - R_p)^2}{\Delta R_p^2} \right], \quad (2.1)$$

где $C_{\max} = C_D / (2\pi)^{1/2} \Delta R_p^2$; C_D — число имплантированных атомов на единицу площади, см^{-2} .

Однако в данной модели не рассматриваются зависимости пробега внедренных ионов от различных факторов в многослойной мишени и эффекты каналирования.

При облучении мишени в направлениях, вдоль которых имеются периодически расположенные атомы и свободные от атомов каналы, наблюдаются аномально большие пробеги ионов. В кремнии эффект легирования, когда большая часть примесных атомов проникает в глубь решетки по каналам со слабым торможением, наблюдается в плоскостях $\langle 110 \rangle$, $\langle 100 \rangle$, $\langle 111 \rangle$. Наименьшая плотность атомов имеется в плоскостях $\langle 110 \rangle$ (рис. 2.2, *a, б*), наибольшая — в плоскостях $\langle 111 \rangle$, при этом средняя длина пробега ионов в плоскостях $\langle 110 \rangle$ вдвое больше, чем в плоскостях $\langle 111 \rangle$.

При каналировании потери кинетической энергии ионов происходят практически за счет электронного торможения. При проникновении иона в кристалл кремния параллельно главной кристалло-

графической оси или плоскости глубина проникновения увеличивается за счет эффекта каналирования, что приводит к профилю распределения, показанному на рис. 2.3. По мере облучения мишени плотность дефектов в приповерхностном слое возрастает, так как каналы перекрываются атомами, смещенными в глубь канала из-за столкновений ионов с атомами решетки, и, следовательно, эффект каналирования исчезает. Если отклонения пучка атомов примеси от основных кристаллографических осей невелики, то эффект каналирования по-прежнему сохраняется. При отклонении пучка на угол, больший критического, движение ионов соответствует случаю разориентированной мишени (см. рис. 2.2, б). Критические углы каналирования для монокристаллического кремния и большинства используемых в практике ионов лежат в диапазоне $3-6^\circ$. Критический угол возрастает с уменьшением энергии ионов, а также с увеличением атомного номера внедряемых ионов и полупроводника мишени. В промышленных микротехнологиях для полного исключения эффекта каналирования ионное легирование часто осуществляют при

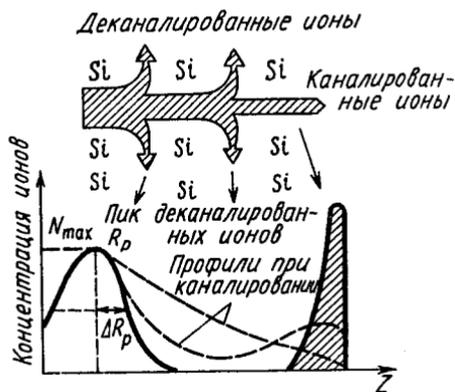


Рис. 2.3. Каналирование ионов в кристалле вдоль главной кристаллографической оси или плоскости

углах разориентации пластин, превышающих критический угол, $\psi > \psi_k$, $\psi_k = 7-8^\circ$. Однако в случае уменьшения энергии ионов эффект легирования не устраняется при разориентации пластин на $4-7^\circ$ и не может быть учтен простыми модификациями функции распределения.

Поскольку ускоренные ионы взаимодействуют с полупроводником — мишенью, в кристаллической решетке образуется значительное количество дефектов. Тяжелые ионы мышьяка и сурьмы (As, Sb) тормозятся в результате столкновений с ядрами, поэтому они вызывают большее количество дефектов, чем легкие ионы бора и фосфора (B, P), торможение которых происходит практически за счет взаимодействия с электронами. В результате столкновения тяжелых ионов с атомом, расположенным в узле кристаллической решетки, ион примеси смещает этот атом в междоузлия, который в свою очередь вызывает каскад последующих смещений, создавая разупорядочение дефектных зон. При больших дозах легирования дефектные зоны с нарушенной структурой могут смыкаться, образуя обширные аморфные слои в поверхностной области. Минимальная доза легирования, при которой возникает аморфный слой, называется дозой аморфизации, которая зависит от атомного номе-

ра легирующего элемента и возрастает с повышением температуры мишени.

Большинство дефектов, возникающих при ионном легировании, может быть исправлено путем кратковременного отжига полупроводниковых кристаллов при температуре 700—1000°C. Отжиг приводит к тому, что смещенные атомы возвращаются в свои равновесные положения в решетке, восстанавливая тем самым монокристаллическую структуру, причем температура и продолжительность (15—20 мин) этого процесса значительно меньше, чем для типичных диффузионных процессов.

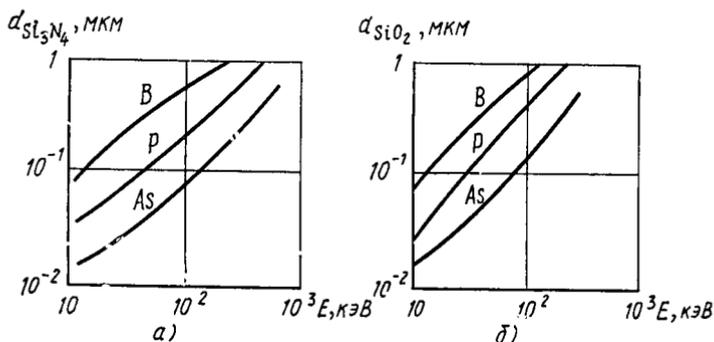


Рис. 2.4. Зависимости минимальной толщины маскирующего диэлектрика от энергии ионов для SiO_2 (а) и Si_3N_4 (б)

Ионы сохраняют большинство свойств исходных атомов. После внедрения в кристаллическую решетку они могут нейтрализоваться или сохранить заряд. При небольших дозах ионы занимают места в кристаллической решетке, становясь примесями замещения, обладающими электрической активностью. При больших дозах большинство ионов останавливается в междоузлиях и становится электрически нейтральными. В частности, при ионном легировании бора имеется большой процент междоузельных ионов, не обладающих электрической активностью.

Селективное легирование выполняют с помощью маскирования, которое достигается использованием экрана соответствующей толщины. Для проведения локального ионного внедрения служат маскирующие экраны из материалов, в которых пробег ионов существенно меньше, чем в кремнии. К материалам масок относятся диэлектрики (SiO_2 , Si_3N_4 , Al_2O_3) и металлы (Al и др.). Пробег иона зависит от энергии, следовательно, от энергии ионов зависит и необходимая толщина маски, что показано на рис. 2.4, а, б. Используя данные рис. 2.4, можно сравнить толщины слоев из разных материалов, поглощающих практически весь пучок (99,9% ионов). Оки-

сел SiO_2 обладает хорошими маскирующими свойствами, у нитрида кремния и алюминия маскирующие характеристики лучше, чем у SiO_2 .

Аналитические аппроксимативные модели кривых распределения внедренных примесей. Вследствие того что профили при ионном легировании, получаемые в результате расчета, являются начальными условиями в уравнениях диффузии, точный профиль распределе-

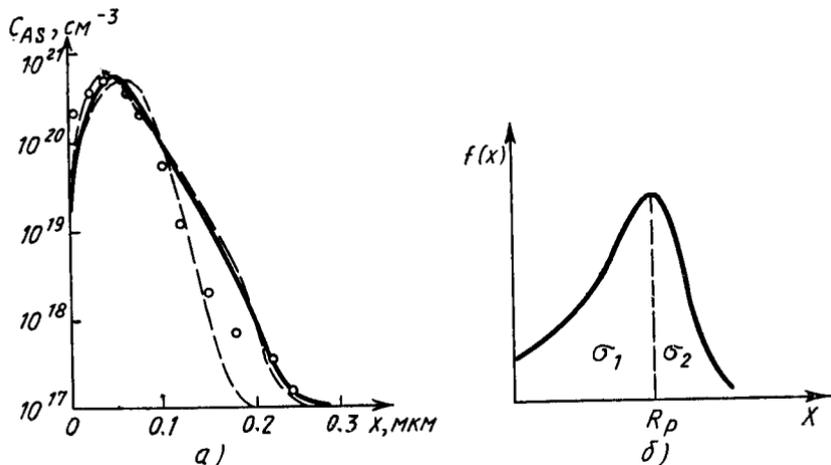


Рис. 2.5. Сравнение моделей ионного легирования мышьяка с энергией 150 кэВ, дозой $3 \cdot 10^{15} \text{ см}^{-2}$ через маску SiO_2 толщиной 50 нм (а); распределение типа «двойная гауссиана» (б):

○ — эксперимент; — — — распределение Пирсона; - - - - - распределение Гаусса

ния примесей зависит от качества моделирования начального этапа его формирования. Наиболее распространенным методом моделирования процесса ионного легирования является аналитическое приближение реальным, экспериментально определяемым формам распределения внедренных в полупроводник примесей. Суть его заключается в том, что, задавшись априори формой кривой распределения, можно экспериментально определить или рассчитать коэффициенты этого распределения. Преимуществом данного метода является простота и наглядность полученного распределения, вычисляемого по аналитической формуле. Однако при данном подходе невозможно рассчитать профили распределения внедренных ионов и радиационных дефектов. В то же время для многих вариантов расчета профилей справедливо сделанное предположение, что начальное распределение дефектов слабо влияет на последующую диффузию внедренной примеси. Это означает, что для большинства стандартных технологических процессов равновесное распределение активной примеси достигается за время, много меньшее времени диффузии. Для ряда технологических процессов, включая перспективные, связанные с созданием тонкослойных и сверхтон-

кослойных структур С³БИС и использованием быстрых термических отжигов или низких температур диффузии, нарушение сделанного допущения требует более сложных моделей ионного легирования.

Наиболее старым и испытанным описанием профиля внедренной примеси, согласно классической теории ЛШШ, является симметричная гауссиана (2.1). Однако уже ранние экспериментальные исследования показали, что данное описание существенно не адекватно экспериментально снятым профилям для большинства ионов как в кремнии, так и в других полупроводниках. Было найдено, что профили большинства внедренных ионов асимметричны в аморфных и кристаллических мишенях, причем асимметрию можно отразить введением третьего центрального момента распределения σ_2 к двум прежним σ_1 , R_p (рис. 2.5, а, б). Большинство экспериментальных данных по распределениям ионов фосфора, мышьяка, сурьмы хорошо аппроксимируется теоретическими расчетами с использованием третьего центрального момента в распределении типа «сдвоенная гауссиана» — гауссианы с различными дисперсиями на левом и правом плечах распределения (рис. 2.5, б).

Таким образом, в рассматриваемом методе моделирования профиля внедренной примеси задается аналитически с помощью формулы

$$C(x) = \frac{Qf(x)}{\int_0^{\infty} f(x) dx} \quad (2.2)$$

Здесь Q — доза имплантированной примеси, функция распределения $f(x)$ для фосфора, мышьяка, сурьмы имеет вид сдвоенной полу-гауссианы (рис. 2.5):

$$f(x) = \begin{cases} \exp[-(x - R_p)^2/(2\sigma_1)], & x \leq R_p, \\ \exp[-(x - R_p)^2/(2\sigma_2)], & x > R_p, \end{cases} \quad (2.3)$$

а коэффициенты R_p , σ_1 , σ_2 приведены в табл. 2.1.

Распределение ионов бора, а также некоторых других ионов в Si, SiO₂, Si₃N₄ лучше всего описывается распределением PEARSON IV. Более того, с помощью семейства распределения PEARSON имеется потенциальная возможность в некотором приближении описать эффекты каналирования, когда зависимость избыточного количества примеси при каналировании от координаты может быть полуэмпирически включена в соответствующие моменты распределения PEARSON.

В общем виде распределение PEARSON определяется дифференциальным уравнением

$$\frac{d f(x)}{d x} = \frac{(x - a_0) f(x)}{b_0 + b_1 x + b_2 x^2}, \quad (2.4)$$

где функция распределения $f(x)$ удовлетворяет соотношению

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (2.5)$$

Четыре константы a_0, b_0, b_1, b_2 в уравнении (2.4) могут быть выражены четырьмя центральными моментами функции распределения (2.5). Первый момент μ_1 характеризует среднюю проекцию пробега ионов:

$$\mu_1 = R_p \int_{-\infty}^{\infty} x f(x) dx. \quad (2.6)$$

Т а б л и ц а 2.1

Параметры распределения имплантированных ионов фосфора, мышьяка, сурьмы

$E, \text{кэВ}$		20	40	60	80	100	200	400	600	800	1000
$R_p, \text{нм}$	P	21.8	44.34	68.94	94.95	122	263.6	545.7	810	1060	1295
	As	19.00	31.60	44.32	35.25	72.8	134.7	253.2	344.7	484.6	609.2
	Sb	12.36	20.67	27.84	—	41.45	174.1	140.5	210.7	284.3	360.3
$\sigma_1, \text{нм}$	P	7.343	15.66	24.7	34	43.35	87.54	161	220	270	316.5
	As	9.11	14.44	19.864	25.484	31.50	55.08	86.96	96.56	133.5	157.3
	Sb	2.07	4.21	5.622	—	8.53	15.55	29.83	45.33	61.43	77.53
$\sigma_2, \text{нм}$	P	15.6	25.8	34.35	41.63	47.8	68.97	81.39	50	13.5	15.6
	As	7.957	13.04	17.17	20.17	22.55	39.16	81.36	135.7	161.3	192.6
	Sb	5.925	9.543	12.78	—	18.65	31.83	55.85	77.25	96.84	115

Три других высших момента задаются соотношением

$$\mu_i = \int_{-\infty}^{\infty} (x - R_p)^i f(x) dx, \quad i = 2, 3, 4. \quad (2.7)$$

Удобно использовать стандартное отклонение ΔR_p , которое определяется как квадратный корень из второго момента μ_2 , и соотношения, связывающие высшие моменты (2.7) с соответствующими параметрами распределения:

$$\Delta R_p = \mu_2^{1/2}, \quad \gamma = \mu_3 / \Delta R_p^3, \quad \beta = \mu_4 / \Delta R_p^4. \quad (2.8)$$

С учетом соотношений (2.5) — (2.8) для четырех констант в выражении (2.4) получаются следующие выражения:

$$a_0 = - \frac{\Delta R_p \gamma (\beta + 3)}{A},$$

$$b_0 = - \frac{\Delta R_p (4\beta - 3\gamma^2)}{A}, \quad (2.9)$$

$$b_1 = a_0, \quad b_2 = -\frac{2\beta - 3\gamma^2 - \Delta R_p}{A},$$

$$A = 10\beta - 12\gamma^2 - 18.$$

Таким образом, профиль распределения внедренной примеси задается выражением (2.2), где аналитическое распределение PEARSON IV имеет вид

$$f(x) = k [b_2(x - R_p)^2 + b_1(x - R_p) + b_0]^{1/2b_2} \times \\ \times \exp \left[-\frac{(b_1/b_2) + 2a}{(4b_2b_0 - b_1^2)^{1/2}} \operatorname{arctg} \frac{2b_2(x - R_p) + b_1}{(4b_2b_0 - b_1^2)^{1/2}} \right]. \quad (2.10)$$

Здесь a_0, b_0, b_1, b_2 — коэффициенты, определяемые согласно (2.9); k — константа нормализации распределения; $R_p, \Delta R_p, \beta, \gamma$ — коэффициенты, значения которых приведены в табл. 2.1 и 2.2.

Т а б л и ц а 2.2

Параметры распределения имплантированных ионов бора

$E, \text{кэВ}$	10	30	50	70	100	200	300	400	600	800
$R_p, \text{нм}$	40	101.5	164.5	226	310	556	735.5	900	1237.5	1515
$\Delta R_p, \text{нм}$	28	42.5	57	70	84,5	110	116	127	130	146.5
γ	-0,3	-0.550	-1.0	-1.0	-1.3	-2.15	-2.25	-2.75	-2.6	-3.3
β	2.9	3.6	5.5	5.5	7.5	17.5	19	32	28	60

Различные распределения типа PEARSON классифицируются поведением корней уравнения $b_0 + b_1x + b_2x^2 = 0$. Для случая $b_1 = b_2 = 0$, соответствующего $\gamma = 0, \beta = 3$, решение уравнения (2.4) является функцией гауссова распределения, что представляет предельный вариант для всех распределений типа PEARSON. Возможны следующие распределения:

$$\gamma \neq 0, \beta < 3 + 1.5\gamma^2 \quad \text{PEARSON I}$$

$$\gamma = 0, \beta < 3 \quad \text{PEARSON II}$$

$$\gamma = 0, \beta = 3 + 1.5\gamma^2 \text{ для } b_2 = 0, b_1 = 0 \quad \text{PEARSON III}$$

$$0 < \gamma^2 < 32, \beta > \frac{39\gamma^2 + 48 + 6(\gamma^2 + 4)^{3/2}}{32 - \gamma^2} \quad \text{PEARSON IV}$$

$$0 < \gamma^2 < 32, \beta = \frac{39\gamma^2 + 48 + 6(\gamma^2 + 4)^{3/2}}{32 - \gamma^2} \quad \text{PEARSON V}$$

$$\gamma \neq 0, 3 + 1.5\gamma^2 < \beta < \frac{39\gamma^2 + 48 + 6(\gamma^2 + 4)^{3/2}}{32 - \gamma^2} \quad \text{PEARSON VI}$$

$$\gamma = 0, \beta > 3 \text{ для } b_0 > 0, b_2 > 0, b_1 = 0 \quad \text{PEARSON VII.}$$

В большинстве случаев одномерные распределения профилей бора хорошо настраивались с помощью аналитических зависимостей PEARSON IV. В то же время имеются литературные данные, что для высокоэнергетичного ионного легирования бора в кремний ($E > 120$ кэВ) более подходит распределение PEARSON VI, а для низкоэнергетичного легирования ($E < 70$ кэВ) — PEARSON I.

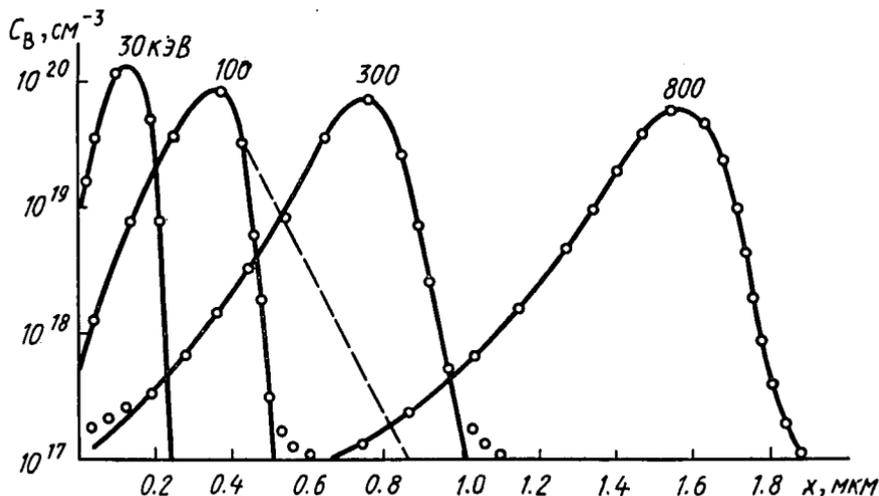


Рис. 2.6. Распределение примеси бора при ионном легировании в кремнии для дозы 10^{15} см $^{-2}$:

— распределение PEARSON IV; --- изменение распределения в случае каналирования ионов (показано без перенормировки распределения); ○ — эксперимент

Для примеси бора распределение PEARSON IV позволяет получить хорошее согласование с экспериментом при $\beta = 1.5\beta_{\min}$, $\beta_{\min} = \frac{48 + 39\gamma^2 + 6(\gamma^2 + 4)^{3/2}}{32 - \gamma^2}$, в то время как для примеси мышьяка настройка распределения PEARSON IV эффективна при $\beta = \beta_{\min}$.

На рис. 2.5, 2.6 представлены сравнительные зависимости распределений примесей мышьяка и бора для различных моделей ионного легирования и эксперимента. Изменение распределения в случае каналирования ионов, показанное на рис. 2.6, без перенормировки распределения дает наиболее серьезный вклад в погрешность аналитической модели PEARSON IV. В рамках модели

PEARSON IV имеется возможность описать так называемые экспоненциальные хвосты модифицированной функцией $f(x) = f(x) + A \exp(-\lambda(x-x_0^0))$, где константы A , λ , x_0 определяются дополнительной настройкой соответствующих экспериментальных данных с учетом каналирования ионов в ходе конкретного технологического маршрута изготовления БИС с использованием ионного легирования.

К аналогичному способу учета эффекта каналирования можно отнести «прикрепление» к правому плечу распределения экспоненциально спадающего «хвоста» с фиксированным параметром $\lambda \approx 0.03 \div 0.06$ мкм. Известно, что каналирования всегда стараются избежать разориентацией кристалла на некоторый угол или (и) ионным легированием через тонкий слой SiO_2 . В последнем случае профиль распределения в кремнии считается неизменным, но сдвигается к границе $\text{Si}-\text{SiO}_2$ на величину $d(W) = W_{0x} \frac{R_p(\text{Si})}{R_p(\text{SiO}_2)}$, где W_{0x} — толщина слоя SiO_2 ; $R_p(\text{Si})$, $R_p(\text{SiO}_2)$ — расчетные или экспериментальные значения математического ожидания распределений ионов в Si и SiO_2 , причем

$$C(x) = \begin{cases} C_{\text{SiO}_2}(x), & x < d, \\ C_{\text{Si}}(x-d), & x > d. \end{cases}$$

Другим более эффективным способом моделирования ионного легирования через маски типа SiO_2 является использование соотношения

$$C(x) = \begin{cases} \frac{(R_p \text{Si})}{R_p(\text{SiO}_2)} C_s \left(\frac{R_p(\text{Si})}{R_p(\text{SiO}_2)} x \right), & x < d, \\ C(\text{Si})(x-d) \left(1 - \frac{R_p(\text{Si})}{R_p(\text{SiO}_2)} \right), & x \geq d. \end{cases}$$

Для двумерного моделирования необходимо знать боковое распределение внедренных примесей с учетом расположения маски и ее края (рис. 2.7, а) при ионном легировании. Двумерный профиль внедренной примеси описывается распределением Гаусса для координаты y с учетом дополнительных параметров маски $d(y)$ и бокового стандартного отклонения σ_y :

$$C(x, y) = \frac{1}{(2\pi\sigma_y^2)^{1/2}} \int_{-\infty}^{\infty} C(x, d(y)) \exp\left(-\frac{(y-y_0)^2}{2\sigma_y^2}\right) dy, \quad (2.11)$$

где $f(x, d(y))$ — соответствующее одномерное распределение с учетом сдвига, зависящего от координаты y ; y_0 — координата края маски.

Значения дисперсии распределения σ_y для различных примесей приведены на рис. 2.7, б.

В ряде программ принята такая зависимость $d(y)$:

$$d(y) = \begin{cases} 0, & 0 \leq y \leq W, \\ y - W, & W \leq y \leq W + h, \\ h, & y > h, \end{cases} \quad (2.12)$$

где W — полуширина окна, в которое производится ионное легирование; h — толщина окисного слоя.

Метод интегрирования кинетического уравнения Больцмана.

Использование результатов классической теории ЛШШ и аналитических аппроксимативных моделей распределения внедренных примесей ранее оказывалось в большинстве случаев достаточно удовлетворительным, так как при продолжительном цикле загонки погрешности начального распределения после ионного легирования были незначительными. В современных технологических процессах изготовления нашли применение чрезвычайно короткие циклы термического отжига, а также ионное легирование через один или несколько слоев непосредственно в подложку. Наличие многослойных структур приводит к тому, что на поверхностях раздела слоев профили распределения примесей становятся разрывными. Эффект каналирования, который для кристаллических мишеней не устраняется при разориентации пластины для малоэнергичных ионов, практически нельзя учесть стандартными функциями распределения. Модифицированные функции распределения с экспоненциальными хвостами учитывают «хвостовые» эффекты только после апостериорной обработки экспериментальных результатов и верно описывают процесс ионного легирования в ограниченном диапазоне доз и энергий для конкретных технологических процессов. Кроме того, на электрические свойства и характеристики изготавливаемых полупроводниковых приборов оказывает существенное влияние вторичное «осколочное» ионное легирование атомов, «выбиваемых» из маскирующего поверхностного слоя в результате столкновений с частицами первичного пучка. Это «осколочное» ионное легирование также нельзя промоделировать с помощью теории ЛШШ и описать аппроксимативными аналитическими выражениями.

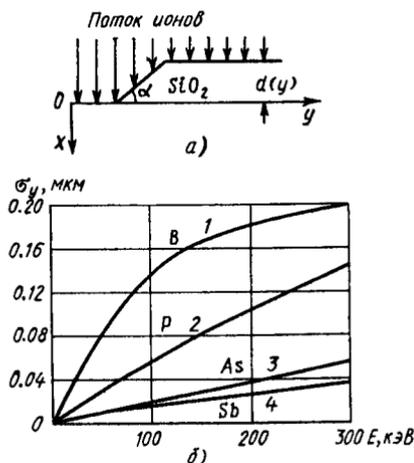


Рис. 2.7. Модель края маски при ионном легировании (а); зависимость боковой дисперсии распределения примеси от энергии ионов для различных примесей (б): бор (кривая 1), фосфор (кривая 2), мышьяк (кривая 3), сурьма (кривая 4)

Эффект каналирования, который для кристаллических мишеней не устраняется при разориентации пластины для малоэнергичных ионов, практически нельзя учесть стандартными функциями распределения. Модифицированные функции распределения с экспоненциальными хвостами учитывают «хвостовые» эффекты только после апостериорной обработки экспериментальных результатов и верно описывают процесс ионного легирования в ограниченном диапазоне доз и энергий для конкретных технологических процессов. Кроме того, на электрические свойства и характеристики изготавливаемых полупроводниковых приборов оказывает существенное влияние вторичное «осколочное» ионное легирование атомов, «выбиваемых» из маскирующего поверхностного слоя в результате столкновений с частицами первичного пучка. Это «осколочное» ионное легирование также нельзя промоделировать с помощью теории ЛШШ и описать аппроксимативными аналитическими выражениями.

Суть применения кинетического уравнения Больцмана (КУБ) к проблеме моделирования ионного легирования заключается в следующем. Процесс ионного легирования может быть рассмотрен как процесс движения ионов в течение своего торможения с изменением кинетической энергии (импульса, скорости) до нулевой энергии (импульса, скорости). Ансамбль частиц, движущихся со скоростью v (импульсом p) и координатой Z , может быть описан некоторой функцией распределения $F(v, Z)$ или $F(p, Z)$. Количество частиц в ансамбле рассчитывают по выражению

$$dN = F(v, Z) d^3v d^3Z, \quad (2.13)$$

а вероятность рассеяния частиц со скоростью v после столкновения и перехода в новый диапазон скоростей $[v^1, v^1 + d^3v]$ в течение времени dt описывается соотношением

$$k(v \rightarrow v^1) d^3v dt = N_s |v| d\sigma(v \rightarrow v^1) dt, \quad (2.14)$$

где $k(v \rightarrow v^1)$ — скорость переноса; N_s — плотность центров рассеяния; $d\sigma$ — дифференциальное сечение рассеяния.

При этом среднее количество ионов, рассеянных внутри дифференциального диапазона фазовых скоростей, описывается КУБ для функции распределения:

$$\begin{aligned} & \frac{\partial F(v, Z)}{\partial t} + v \nabla F(v, Z) = Q(v, Z) + \\ & + N_s \int [d\sigma(v' \rightarrow v) |v'| F(v', Z') - d\sigma(v \rightarrow v') |v| F(v, Z)], \end{aligned} \quad (2.15)$$

где $Q(v, Z)$ — генерационный член, характеризующий возрастание частиц в состоянии покоя.

Уравнение (2.15) может быть переписано для распределения импульсов внедренных ионов по глубине мишени Z , заданных функцией $F(p, Z)$ для ионов, движущихся под углом θ_p :

$$\frac{\partial F(p, Z)}{\partial Z} = \int \left\{ \frac{F(p', Z) d\sigma(p' \rightarrow p)}{\cos \theta_{p'}} - \frac{F(p, Z) d\sigma(p \rightarrow p')}{\cos \theta_p} \right\}, \quad (2.16)$$

где верхний индекс (штрих) обозначает параметры частиц после столкновения.

Интегрирование КУБ (2.15), (2.16) производится от $Z=0$ (поверхность образца) до $Z>0$ при граничном условии

$$F(v, 0) = C_D \delta(v \rightarrow v_0), \quad (2.17)$$

где C_D — начальная доза; v_0 — начальная скорость потока ионов.

Интегрируя КУБ (2.16) при известном граничном условии (2.17), получаем распределение ионов по глубине Z . При этом каждое состояние определяется энергией E_i ($0 \leq E_i \leq E_0$) и углом θ_i , $0 \leq \theta_i \leq \pi/2$. Таким образом, при численном решении функция распределения $F(p, Z)$ представляется в виде матрицы $\{F_{i,j}\}$, опи-

сывающей количество ионов, движущихся с энергией E_i ($0 \leq E_i \leq E_0$) под углом θ_i ($0 \leq \theta_i \leq \pi/2$). Обычно матрица имеет размеры 10×10 , 20×10 (рис. 2.8, а, б). Вычисления начинают с оценки рассеяния на малом интервале dZ , используя КУБ (2.16). После каждого шага ионы с энергией, меньшей некоторого порога E_{\min} , считаются остановившимися и выбывают из рассмотрения. Интегрирование продолжается до тех пор, пока все ионы не остановятся.

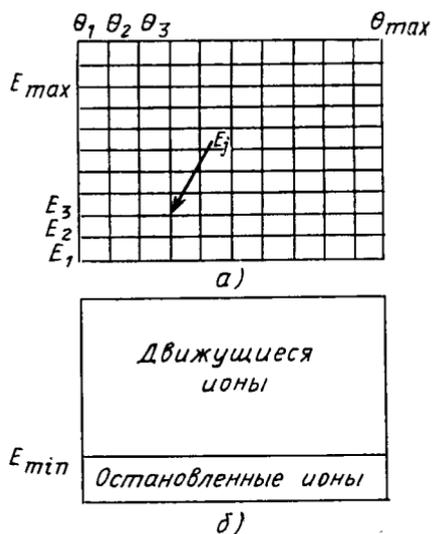


Рис. 2.8. Матрица распределения, учитывающая столкновительные процессы (а); матрица распределения после каждого шага интегрирования (б)

Распределение остановленных ионов по глубине и есть искоемое распределение внедренных ионов (рис. 2.9). Приведенный алгоритм был рассмотрен на примере аморфной подложки. Для кристаллической под-

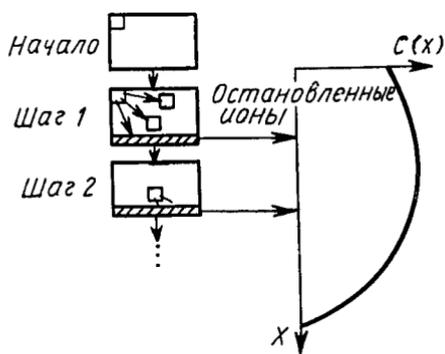


Рис. 2.9. Структурная схема расчета профиля примеси ионного легирования по методу КУБ для аморфной мишени

ложки эффекты каналирования добавляются в вычислениях следующим образом. На каждом шаге интегрирования распределение $F(p, Z)$ проводят, чтобы определить среднее количество ионов, рассеянных внутри каждого канала за предыдущий интервал dZ . Эти ионы перемещаются от распределения $F(p, Z)$ и рассматриваются отдельно. Электронное торможение внутри канала используется затем для определения распространения конечного распределения ионов так, чтобы ионы могли быть перемещены в соответствующие позиции концентрационного профиля внедренных примесей (рис. 2.10). На рис. 2.11 приведенные расчетные зависимости концентрационных профилей внедренных примесей по методу интегрирования КУБ характеризуются приемлемой адекватностью при существенном влиянии эффектов каналирования.

Отметим, что метод интегрирования КУБ позволяет также эффективно рассчитывать распределение дефектов после ионного

легирования. Существенное влияние дефектов на моделирование последующих после ионного легирования операций в настоящее время пока не установлено, но следует ожидать возрастания их влияния по мере уменьшения времени и температуры отжига (в частности, лампового и лазерного).

Метод интегрирования КУБ, модифицированный с учетом эффектов каналирования указанным способом, достаточно просто распространяется на двумерное моделирование ионного легирования без значительного увеличения вычислительных затрат.

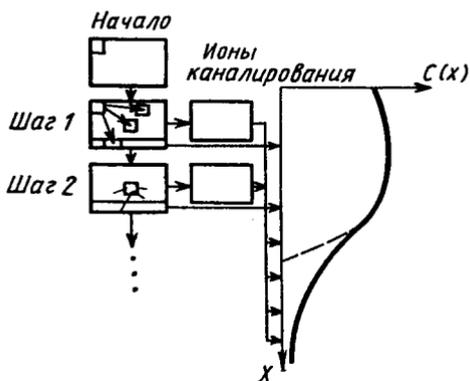


Рис. 2.10. Структурная схема расчета профиля примеси ионного легирования с учетом каналирования в кристаллической мишени

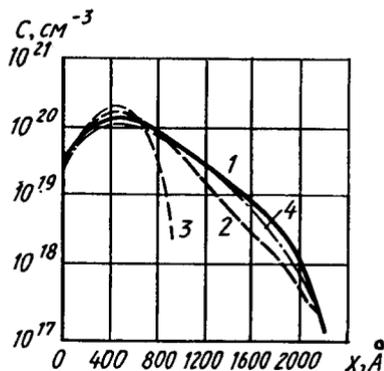


Рис. 2.11. Зависимости профилей распределения примеси бора при ионном легировании в кремний $\langle 100 \rangle$:

1 — эксперимент без угла разориентации; 2 — эксперимент с углом разориентации 7° ; 3 — расчет для аморфной мишени; 4 — расчет для кристаллической мишени

Моделирование ионного легирования в многослойных структурах с помощью метода подбора доз. Простейшим аналитическим аппроксимативным методом практически невозможно построить точный разрывный профиль распределения примесей при ионном легировании в многослойную структуру типа $\text{Si}_3\text{N}_4\text{—SiO}_2\text{—Si}$ или $\text{Si}^*\text{—SiO}_2\text{—Si}$, за исключением очень тонких слоев SiO_2 (5—30 Å). Поскольку данная операция часто используется в современных технологических процессах изготовления БИС, это вынуждает для улучшения адекватности применять либо метод Монте-Карло, либо метод интегрирования КУБ для построения требуемого распределения. Метод интегрирования КУБ является более быстрым: если расчет одного одномерного профиля данным методом требует t условных единиц машинного времени, то по методу Монте-Карло (40—60) t . С целью сравнения отметим, что метод аналитических аппроксимативных функций в трехслойной структуре $\text{Si—SiO}_2\text{—Si}$ для тонкого слоя $\text{SiO}_2 \leq 30 \text{ Å}$ при моделировании ионного легирования

требует порядка $10^{-2} t$ единиц машинного времени. Однако разрывный профиль распределения примесей при ионном легировании в многослойную структуру с достаточно толстыми промежуточными слоями требует существенной модификации алгоритмов использования методов аналитических аппроксимативных функций.

Высокой эффективностью при моделировании ионного легирования в многослойных структурах обладает метод подбора доз, с помощью которого можно получать приемлемую адекватность профиля распределения примеси, характерную для метода интегрирования КУБ, но с минимальными вычислительными затратами, например $10^{-1} t$ условных единиц машинного времени.

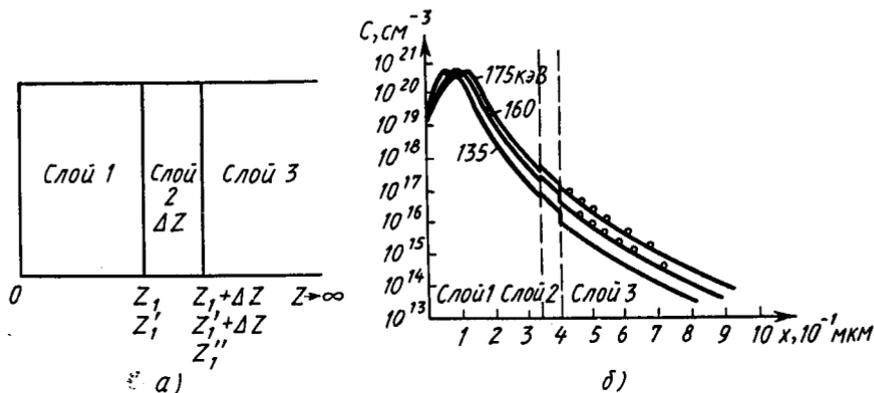


Рис. 2.12. Трехслойная полупроводниковая структура (а); расчетные и экспериментальные зависимости для ионного легирования мышьяка с различными энергиями в многослойную мишень (б)

В данном методе, основанном на статистическом распределении и численном интегрировании доз в каждом слое, требуются следующие шаги для моделирования ионного легирования с общей дозой D и энергией E .

Шаг 1. Распределение внедренной примеси в слое 1 ($0—Z_1$) на рис. 2.12, а определяется для дозы D и энергии E как $p_1 f_1(Z)$, где p_1 — пик концентрации; $f_1(Z)$ — функция статистического распределения от глубины Z . Количество внедренных ионов в слое 1 равно d_1 , где

$$d_1 = \int_0^{Z_1} p_1 f_1(Z) dZ. \quad (2.18)$$

Шаг 2. Предполагая, что ионное легирование (D, E) было непосредственно проведено в слой 2, на глубине Z_1 — границе двух слоев — содержится $d_1 = d_2$ внедренных ионов:

$$d_1 = d_2 = \int_0^{Z_1} p_2 f_2(Z) dZ, \quad (2.19)$$

что эквивалентно перекрытию слоя 1 слоем 2 толщиной Z_1' .

Шаг 3. Распределение внедренной примеси из $Z_1' \rightarrow Z_1 + \Delta Z$, взятое с дозой D и энергией E , переносится в слой 2 из $Z_1' \rightarrow Z_1 + \Delta Z$. При этом количество ионов в слое 2 толщиной ΔZ равно d_3 , где

$$d_3 = \int_{Z_1}^{Z_1 + \Delta Z} p_2 f_2(Z) dZ. \quad (2.20)$$

Шаг 4. Шаг 2 повторяется для слоя 3, чтобы получить глубину Z_1'' , для которой количество внедренных примесей определяется выражением

$$d_1 + d_3 = d_4 = \int_0^{Z_1'} p_3 f_3(Z) dZ. \quad (2.21)$$

Шаг 5. Распределение примеси из $Z + \Delta Z \rightarrow a$ берется таким же, как из слоя 3 (из $Z_1'' \rightarrow a$).

Использование соотношений (2.18) — (2.21) в заданной последовательности шагов позволяет просто и эффективно рассчитывать профили распределения примесей.

На рис. 2.12, б представлены расчетные и экспериментальные зависимости для ионного легирования мышьяка в многослойную подложку Si—SiO₂—Si, подтверждающие приемлемую адекватность приведенного метода для статистического распределения.

Наличие в программах технологического моделирования БИС моделей ионного легирования различного уровня сложности — Монте-Карло, интегрирования КУБ, подбора доз для заданных статистических распределений — позволяет в зависимости от типа и конфигурации многослойных мишеней применять наиболее подходящие модели с целью минимизации вычислительных затрат.

2.3. МОДЕЛИ ДИФFUЗИОННЫХ ПРОЦЕССОВ

Диффузией называется перенос атомов вещества, обусловленный их хаотическим тепловым движением, в направлении уменьшения концентрации. Диффузия атомов в кристаллической решетке осуществляется отдельными скачками из одного положения равновесия в другое, причем длины таких элементарных перемещений имеют порядок межатомных расстояний и составляют несколько десятых долей нанометра. За счет элементарных скачков атомы могут перемещаться на большие расстояния.

Возможны три механизма атомных скачков: взаимный обмен местами, движение по вакансиям и перемещение по междоузлиям (рис. 2.13). Обмен местами двух соседних атомов является простейшим актом диффузии (рис. 2.13, а). Однако при этом в плотноупакованной структуре атом должен преодолевать большой потенциальный барьер, что обусловлено необходимостью смещения соседних атомов. При кольцевом обмене (рис. 2.13, б) несколько

атомов согласованно перемещаются на одно межатомное расстояние. При подобном перемещении потенциальный барьер, преодолеваемый атомом, меньше, чем в первом случае. Однако вероятность осуществления такого диффузионного механизма резко уменьшается с ростом числа атомов в кольце, так как при этом возрастает суммарная энергия элементарного акта перемещения.

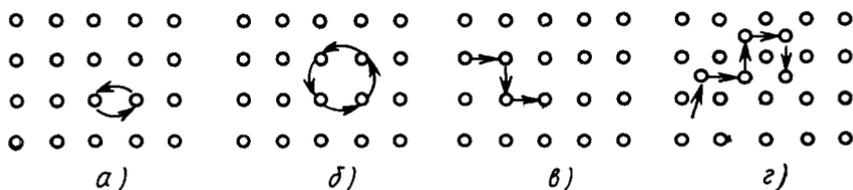


Рис. 2.13. Механизм диффузионных перемещений атомов:

a — обмен местами соседних атомов; *b* — кольцевой обмен; *в*, *г* — перемещение по вакансиям и междоузлиям соответственно

Механизм диффузии по вакансиям происходит следующим образом: вначале в кристаллической решетке образуются вакансии (рис. 2.13, *в*), затем атом примеси перескакивает на последнее место решетки, далее атомы последовательно перемещаются по кристаллу.

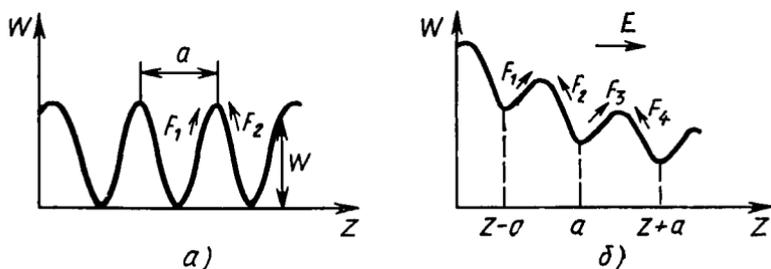


Рис. 2.14. Термостимулированная диффузия (*a*); диффузия, стимулированная полем (*б*)

Механизм замещения или диффузии атомов по междоузлиям (рис. 2.13, *г*) заключается в том, что происходит перескок атома из узла решетки в междоузье и дальнейшее перемещение атома по междоузлиям в ближайшее вакантное место решетки.

Подытоживая результаты большого количества исследований кремния, для достигнутого современного технологического уровня можно сделать заключение, что в примесной диффузии доминирующую роль играют механизмы по вакансиям и междоузлиям. По данным исследований последних лет, вакансионный механизм диффузии доминирует при низких температурах (800—950°C), междоузельный — при высоких температурах (1100—1200°C).

Диффузионные скачки атомов примесей в решетке происходят в трех измерениях, а суммарный поток определяется статистическим усреднением за определенный промежуток времени. Атомы в кристалле образуют последовательность потенциальных барьеров с высотой около 0.1 эВ (рис. 2.14, а). При приложении постоянного электрического поля распределение потенциала по координате Z имеет наклон, как показано на рис. 2.14, б. Это облегчает перемещение положительно заряженных частиц вправо и затрудняет их движение влево. Поток F в точке Z равен среднему значению потоков в точках $Z-a/2$ и $Z+a/2$, которые равны соответственно F_1-F_2 и F_3-F_4 (рис. 2.14, б).

Составляющая потока F_1 равна произведению: 1) плотности примесей на единицу площади в плоскости потенциальной ямы при $Z-a$, $aC(Z-a)$; 2) вероятности перескока любой из этих примесей в яму с координатой Z , $\exp\left[-\frac{q}{kT}\left(W-\frac{aE}{2}\right)\right]$; 3) частоты скачков ν :

$$F_1 = aC(Z-a) \exp\left[-\frac{q}{kT}\left(W-\frac{1}{2}aE\right)\right] \nu. \quad (2.22)$$

Здесь $aC(Z-a)$ — плотность частиц на единицу площади, расположенных в яме с координатой $Z-a$; W — высота барьера; E — напряженность электрического поля.

Высота барьера уменьшается за счет действия электрического поля и соответственно увеличивается вероятность перескока частиц из ямы $Z-a$ в яму Z .

Формулы, аналогичные (2.22), могут быть записаны для потоков F_2 , F_3 , F_4 . Объединяя соотношения для вычисления результирующего потока F в точке Z в предположении, что $C(Z+a) = C(Z) \pm a \partial C / \partial Z$, получим

$$F(Z) = -(\nu a^2 e^{-qW/kT}) \frac{\partial C}{\partial Z} \operatorname{ch}\left(\frac{qaF}{2kT}\right) + (2a\nu e^{-qW/kT}) C \operatorname{sh}\left(\frac{qaF}{2kT}\right). \quad (2.23)$$

В предельном случае, имеющем важное практическое значение, для относительно слабого электрического поля ($E \ll kT/qa$), используя разложение в ряд гиперболических функций $\operatorname{ch} Z \approx 1$, $\operatorname{sh} Z \approx Z$ при $Z \rightarrow 0$, получим выражение для предельного потока положительно заряженных частиц

$$F(Z) = -D \frac{\partial C}{\partial Z} + \mu EC, \quad (2.24)$$

где поток представлен суммой диффузионного и дрейфового членов; D — коэффициент диффузии:

$$D = \nu a^2 \exp(-qW/kT); \quad (2.25)$$

μ — подвижность:

$$\mu = \frac{va^2 \exp(-q\psi/kT)}{kT/q}. \quad (2.26)$$

Как следует из формул (2.25), (2.26), коэффициент диффузии D и подвижность μ связаны между собой известным соотношением Эйнштейна

$$D = \frac{kT}{q} \mu, \quad D = \varphi_T \mu, \quad (2.27)$$

где $\varphi_T = kT/q$ — температурный потенциал.

Коэффициент диффузии зависит от энергии связи атомов примеси в решетке, плотности вакансий в кристалле, постоянной кристаллической решетки и других параметров. С повышением температуры коэффициент диффузии резко возрастает, изменяясь по экспоненциальному закону (2.25): $D = D_0 \exp(-E_{\text{акт}}/kT)$, где D_0 — константа диффузии, $\text{м}^2/\text{с}$; $E_{\text{акт}}$ — энергия активации диффузионного процесса, т. е. энергия, необходимая для элементарного скачка диффузионного атома примеси, Дж. Например, для мышьяка $D_0 = 0.32 \cdot 10^{-4} \text{ м}^2/\text{с}$, $E_{\text{акт}} = 5.76 \cdot 10^{-19} \text{ Дж}$.

На основании формулы (2.24) получается обобщенное уравнение Фика, описывающее процесс диффузий:

$$\begin{aligned} \frac{\partial C}{\partial t} &= -\text{div } F, \\ \frac{\partial C}{\partial t} &= D \frac{\partial^2 C}{\partial z^2} - \mu_+ E \frac{\partial C}{\partial z}, \\ \frac{\partial C}{\partial t} &= D \frac{\partial^2 C}{\partial z^2} + \mu_- E \frac{\partial C}{\partial z}, \end{aligned} \quad (2.28)$$

соответственно для положительных, отрицательных зарядов атомов примесей.

В более общем виде математической формулировкой процесса диффузии служит уравнение непрерывности в дифференциальной форме

$$\frac{\partial C}{\partial t} + \text{div } F = g - l, \quad (2.29)$$

где g , l — скорости образования и уничтожения примеси.

При $g = l$ уравнение (2.29) переходит в уравнение Фика (2.28). В случае неподвижных границ объема, где происходит процесс диффузии, уравнение (2.28) может быть переписано с учетом электростатического потенциала ψ :

$$\frac{\partial C}{\partial t} = \text{div}(D \text{ grad } C) \pm \text{div } DC \text{ grad } \frac{\psi}{\varphi_T}. \quad (2.30)$$

В общем случае для определения Ψ вместе с уравнением (2.30) необходимо решать уравнения Пуассона. Однако во всех практически важных случаях потенциал Ψ можно рассчитать из условий квазинейтральности и квазиравновесия $p-n=C_A-C_D$, $p-n=n_i^2$, $\Psi \approx \varphi_T \ln \frac{n}{n_i}$, где n — концентрация электронов; n_i — собственная концентрация электронов в полупроводнике при температуре диффузии; C_D , C_A — соответственно концентрации ионизированной донорной и акцепторной примесей, Θ — знак заряда примеси;

$$\Psi = \varphi_T \operatorname{arcsch} \left(-\frac{C\Theta}{2n_i} \right) \approx \varphi_T \ln \frac{n}{n_i}. \quad (2.31)$$

Все основные физические эффекты, влияющие на процесс диффузии, можно разделить на два типа: а) «дальнодействующие», обусловленные такими неравновесными процессами, как окисление, например, на границе Si—SiO₂, генерация вакансий; б) «локальные», обусловленные концентрацией примесей (или электронов) в данной точке.

Важнейшими из «локальных» эффектов являются:

1. Кулоновское взаимодействие с заряженными вакансиями.
2. Возникновение в кристаллах полупроводника внутреннего электростатического поля за счет различной скорости диффузии ионов и свободных носителей заряда; возникающее в кристалле поле приводит к эффективному увеличению скорости диффузии.
3. Взаимодействие между заряженными ионами, приводящее к образованию и распаду ионных комплексов — кластеров. В случае донорно-акцепторного взаимодействия, сопровождающегося образованием спаренных донорно-акцепторных кластеров, происходит уменьшение эффективной скорости диффузии. При взаимном влиянии двух донорных (или акцепторных) примесей взаимодействие между ними носит характер отталкивания, что ускоряет диффузионный процесс.
4. Влияние генерации концентрации заряженных вакансий. Из-за различия размеров атомов (ионов) растворителя и примеси возникают деформации решетки, локальные нарушения вокруг каждого атома примеси. При высоких концентрациях примеси накопление локальных деформаций приводит к появлению макроскопических деформаций структуры. Взаимодействие локальных и макроскопических дислокаций может приводить к увеличению концентрации вакансий в кристалле, росту коэффициентов диффузии примесей.

Рассмотрим суть возникновения и влияния внутреннего электростатического поля в кристалле.

В случае, например, диффузии в полупроводник донорной примеси свободные электрические заряды, возникающие в результате ионизации примесных атомов, диффундируют в глубь монокристалла значительно быстрее, чем соответствующие примесные ионы. В результате происходит разделение зарядов и возникает электри-

ческое поле напряженностью $E = \frac{kT}{gC_D} \frac{\partial C_D}{\partial Z}$, где C_D и $\frac{\partial C_D}{\partial Z}$ — концентрация и градиент концентрации донорной примеси.

Данное внутреннее поле, возникающее вследствие отличного от нуля градиента примесей, влияет на диффузионный поток и позволяет частично отразить факт взаимодействия двух примесей при их совместной диффузии (например, замедление диффузии базовой примеси при совместной диффузии с эмиттерной). С учетом выражения (2.31) градиент электростатического потенциала

$$\text{grad } \Psi \approx -\frac{kT}{q} \left[1 + \left(\frac{C}{2n_i} \right)^2 \right]^{-1/2} \frac{\theta}{2n_i} \text{grad } C. \quad (2.32)$$

Подставляя выражение (2.32) в уравнение диффузии (2.30), после алгебраических преобразований получим новое уравнение

$$\frac{\partial C}{\partial t} = \text{div} (Dh \text{grad } C), \quad (2.33)$$

где h — фактор усиления поля:

$$h = 1 + \left[1 + 4 \left(\frac{n_i}{C} \right)^2 \right]^{-1/2} = 1 + \left[1 + 4 \frac{n_i}{n} \right]^{-1/2}; \quad (2.34)$$

$$n_i^2 = 1,5 \cdot 10^{33} T^3 \exp \left[\left(1,21 + 7,1 \cdot 10^{-10} \sqrt{\frac{n_i}{T}} \right) / kT \right]. \quad (2.35)$$

Суммарный диффузионный поток для донорной примеси C_D (аналогично для акцепторной примеси C_A) записывается в виде

$$F = -D(h \text{grad } C_D - (h-1) \text{grad } C_A). \quad (2.36)$$

Согласно выражениям (2.33)–(2.34), внутреннее поле может увеличивать эффективный коэффициент диффузии максимально в два раза при $C \gg n_i$. По полученным соотношениям (2.33), (2.34), (2.36) необходимо сделать несколько замечаний. При больших концентрациях примеси в общем случае при вырождении полупроводника нельзя пользоваться статистикой Больцмана, как при выводе формулы (2.31) для Ψ , и соотношением Эйнштейна $D = \mu kT/q$. Однако сравнительный расчет эффекта внутреннего поля для использования статистик Ферми — Дирака и Больцмана дает минимальное отличие друг от друга.

При постоянном коэффициенте диффузии, не зависящем от концентрации примесей, легко получаются аналитические решения как для одномерного, так и для двумерного расчетов. Обе основные стадии диффузионного процесса, описываемого уравнением диффузии (2.33) при соответствующих граничных условиях, представлены на рис. 2.15, а на рис. 2.16, а, б приведены температурные изоконцентрационные зависимости коэффициентов диффузии различных примесей.

Однако на сам процесс диффузии и его коэффициенты существенное влияние оказывает эффективная концентрация заряженных

вакансий, зависящая от положения уровня Ферми. Для доминирующего в кремнии вакансионного механизма диффузии при произвольных концентрациях примеси вводят понятие эффективного коэффициента диффузии, равного сумме коэффициентов диффузии по различным зарядовым состояниям вакансий:

$$D(n_n) = D^{\times} + D^{-}n_n + D^{=}n_n^2 + D^{+}|n_n|, \quad (2.37)$$

Стадия диффузионного процесса	Граничные условия	Решение одномерного уравнения диффузии	Модель процесса
Загонка	$C = C_0$ при $Z = 0$ $C = 0$ при $Z = \infty$	$C(Z, t) = C_0 \operatorname{erfc} \frac{Z}{2(Dt)^{1/2}}$ $Q(t) = \int_0^{\infty} C(Z, t) dZ = \frac{2}{\sqrt{\pi}} (Dt)^{1/2} C_0$	
Разгонка	$\frac{\partial C}{\partial Z} \Big _{0,t} = 0$ $C(\infty, t) = 0$	$C(Z, t) = \frac{Q}{(\pi Dt)^{1/2}} \exp\left(-\frac{Z^2}{4Dt}\right)$	

Рис. 2.15. Основные стадии и граничные условия диффузионного процесса атомов примеси

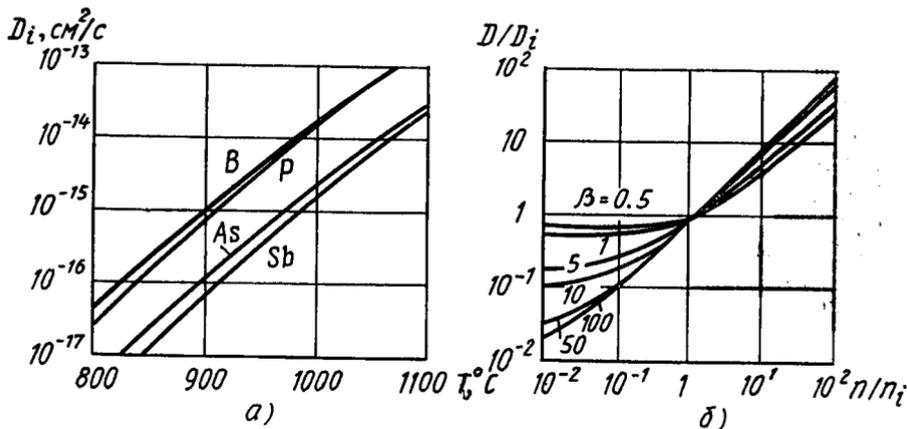


Рис. 2.16. Зависимости собственного коэффициента диффузии бора, фосфора, мышьяка и сурьмы от температуры (а); зависимость нормированного коэффициента диффузии от нормированной концентрации для различных значений параметра β (б)

где знаки \times , $-$, $=$, $+$ в верхних индексах означают соответственно нейтральную, отрицательную, дважды отрицательную и положительно заряженную вакансии; $n_n = n/n_i$ — нормированная концентрация электронов (для статистики Больцмана), эффективное значение, определяющее замену концентрации заряженных вакансий концентрацией атомов примеси.

При низкой концентрации примеси $n_n \ll 1$, по определению, получаем собственный коэффициент диффузии:

$$D_i = D^\times + D^- + D^= + D^+ \quad (2.38)$$

При определении $D(n_n)$ (2.37) было найдено, что для каждой легирующей примеси доминирующими могут быть разные заряженные вакантные состояния. Акцепторы, заряженные отрицательно, диффундируют в кристаллическую решетку по нейтральным положительно заряженным вакансиям. Доноры, будучи заряженными положительно, диффундируют по нейтральным и отрицательно заряженным вакансиям.

Удобной для представления в численных технологических расчетах моделей коэффициентов формулы (2.37) является следующая форма записи:

$$D = D_i \frac{1 + \beta_i^-(n/n_i) + \beta_i^=(n/n_i)^2 + \beta_i^+(p/n_i)}{1 + \beta_i^- + \beta_i^= + \beta_i^+}, \quad (2.39)$$

где D_i — собственный коэффициент диффузии (2.38); β_i^- , $\beta_i^=$, β_i^+ — коэффициенты, описывающие эффективность заряженных вакансий относительно нейтральных; $\beta_i^- = D^-/D^\times$, $\beta_i^= = D^=/D^\times$, $\beta_i^+ = D^+/D^\times$.

Температурные зависимости собственного коэффициента диффузии D_i при $n=p=n_i$ для бора, фосфора, мышьяка и сурьмы представлены на рис. 2.16, а. С учетом температурной зависимости для собственной концентрации (2.34) на рис. 2.16, б приведены температурно-концентрационные зависимости нормированного коэффициента диффузии D/D_i согласно (2.39), отражающие усиление влияния вакансионного механизма на коэффициент D с увеличением концентрации примесей и снижением температуры отжига.

Конкретизация вида формул (2.37) или (2.39) с учетом специфических особенностей процесса диффузии для каждой из используемых примесей составляет суть кинетических моделей температурно-концентрационных моделей коэффициентов диффузии.

Диффузия бора. Диффузия бора происходит по нейтральным и положительно заряженным вакансиям. В большинстве случаев (для низких и средних концентраций) при численных расчетах коэффициент диффузии бора удобно задавать в виде

$$D_B(n_n) = D_i^B \frac{1 + \beta_B(T)/n_n}{1 + \beta_B(T)} = D^\times + D^+(p/n_i), \quad (2.40)$$

где $\beta_B(T) \approx 0.024 \exp(0.512/kT)$ — экспериментально определенная для бора эффективность заряженных вакансий.

Зависимость коэффициента диффузии бора от температуры и концентрации представлена на рис. 2.17. Отклонения от канонической зависимости коэффициента диффузии (2.40) наблюдаются экспериментально при низких температурах диффузии $T < 900^\circ\text{C}$, при этом возникает немонотонность зависимости $D_B(C_B)$ при $C_B \approx \approx 3 \cdot 10^{19} \text{ см}^{-3}$. Это означает, что при высококонцентрационной диффузии бора (а в дальнейшем и мышьяка) необходимо учитывать эффект кластерообразования, т. е. возникновения комплексов — кластеров. Достаточно эффективным и наиболее распространенным способом моделирования эффекта кластерообразования является равновесный, когда предполагается, что концентрации активной примеси C_A и кластеров C_C приходят в равновесие за время, много меньшее времени диффузии. Другой способ моделирования эффекта кластерообразования — неравновесный — будет рассмотрен на примере диффузии мышьяка, а для бора более актуален равновесный способ. Концентрации полной C_T , активной C_A примесей и кластеров C_C бора при условии равенства скоростей генерации и уничтожения

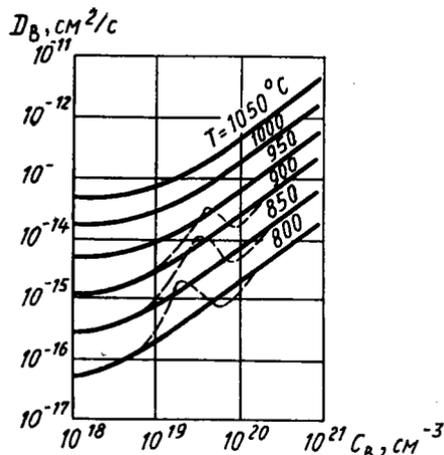


Рис. 2.17. Зависимость коэффициента диффузии бора от концентрации и температуры (пунктиром показано влияние эффекта кластерообразования на коэффициент диффузии)

примесей $g=l=0$ удовлетворяют соотношению

$$C = C_T = mC_C + C_A, \quad (2.41)$$

где m — число атомов в кластере (для бора $m=12$).

Рассмотрим эффект уменьшения коэффициента диффузии с ростом концентрации примеси в связи с кластерообразованием.

Предполагая, что в решетке кристалла могут диффундировать только атомы активной примеси, запишем для них уравнение диффузии:

$$J = -D \frac{\partial C_A}{\partial X}. \quad (2.42)$$

Концентрации C_A и C_C связаны соотношением

$$C_A/C^* = C_C/C^*, \quad (2.43)$$

где C^* — равновесная константа кластерообразования, характеризующая значение концентрации, когда начинается сильное влияние эффекта кластерообразования.

Выражение (2.41) с учетом (2.43) записывается в виде

$$C/C^* = C_A/C^* + m(C_A/C^*)^m. \quad (2.44)$$

Из соотношения (2.44) следует, что при образовании кластеров бора необходимо учитывать отличие полной концентрации бора $C_T = C$ от концентрации электрически активных атомов C_A :

$$C = C_T = C_A [1 + m(C_A/C^*)^{m-1}], \quad (2.45)$$

$$C^* = 4.9 \cdot 10^{23} \exp(-0.9/kT).$$

Уравнение диффузии (2.42) с учетом соотношения (2.44) преобразуется к виду

$$J = -D \frac{\partial C_A}{\partial x} = -\left(D \frac{\partial C_A}{\partial C}\right) \frac{\partial C}{\partial x} = \frac{D_B}{1 + m^2(C_A/C^*)^{m-1}} \frac{\partial C}{\partial x}, \quad (2.46)$$

где

$$D_B = \begin{cases} \frac{D_B^0}{1 + m^2(C_A/C^*)^{m-1}}, & C = C_A + mC_C, \\ D_B^0, & C = C_A. \end{cases} \quad (2.47)$$

Из выражений (2.46), (2.47) видно, что эффективный коэффициент диффузии при кластерообразовании уменьшается с ростом концентрации, так как градиент концентрации активной примеси бора заметно падает. Обобщенное выражение для коэффициента диффузии бора, принятое в ряде зарубежных и отечественных программ, задается как

$$D_B = 2 \cdot 10^7 \exp\left(-\frac{3.46}{kT}\right) \left(1 + \frac{20}{n_n}\right) / \left[1 + 144 \left(\frac{C_A}{C^*}\right)^{11}\right]. \quad (2.48)$$

Таким образом, аномальное поведение зависимости $D_B(C_B, T)$ при $T < 900^\circ\text{C}$ и $C_B \approx 3 \cdot 10^{19} \text{ см}^{-3}$ (рис. 2.17) может быть объяснено формулами для D_B (2.47), (2.48), а часть зависимости $D_B(C_B, T)$ представляет собой не характерные для современной технологии режимы: $T = 900^\circ\text{C}$, $C_B \approx 8 \cdot 10^{19} \text{ см}^{-3}$, $T = 800^\circ\text{C}$, $C_B > 5 \cdot 10^{19} \text{ см}^{-3}$. Отметим также, что вклад диффузии атомов бора по отрицательно заряженным вакансиям незначителен и нет необходимости модифицировать модель коэффициента диффузии (2.40), (2.47), (2.48).

Диффузия мышьяка. Диффузия мышьяка происходит по нейтральным и отрицательно заряженным вакансиям. Для невысоких концентраций коэффициент диффузии мышьяка в численных расчетах удобно задавать формулой

$$D_{As} = D_i^{As} \frac{1 + \beta_{As}(T) n_n}{1 + \beta_{As}} = D^+ + D^- \left(\frac{n}{n_i}\right), \quad (2.49)$$

где $\beta_{As} \approx 100$, $D_i^{As} = 23 \exp(-4.09/kT)$.

В то же время характерной особенностью высококонцентрационной диффузии мышьяка, так же как и бора, является образование m -атомных комплексов — кластеров. В первом приближении влияние кластерообразования на диффузию было выяснено на примере бора. Но поскольку концентрация мышьяка в современных технологических процессах изготовления БИС, как правило, значительно выше концентрации бора, кластерообразование играет доминирующую роль в диффузии мышьяка и его надо рассматривать более подробно, учитывая неравновесную кинетику кластерообразования. Суть моделирования диффузии мышьяка сводится к следующему.

Кластеры считаются неподвижными и уравнение (2.29) решается для подвижной, активной компоненты мышьяка, а расход и образование кластеров во время диффузии учитываются за счет членов генерации и уничтожения примесей g, l . Для описания кинетики кластерообразования введем два коэффициента: коэффициент распада кластеров k_d и коэффициент образования кластеров k_c . Тогда для концентраций кластеров мышьяка C_c и электрически активной примеси C_A уравнение кинетики кластерообразования запишется в виде

$$\frac{\partial C_c}{\partial t} = k_c C_A^m - k_d C_c = l - g. \quad (2.50)$$

Вводя равновесный коэффициент кластерообразования

$$k_l = (k_c/k_d)^{1/m}, \quad (2.51)$$

уравнение (2.50) можно переписать в виде

$$\frac{\partial C_c}{\partial t} = k_d [m (k_e C_A)^m - C_c] = l - g. \quad (2.52)$$

Уравнение (2.52) совместно с уравнениями (2.49), (2.29) описывают диффузию мышьяка. При численно-разностном решении, когда $\frac{\partial C_c}{\partial t} \rightarrow \frac{C_c^i - C_c^{i-1}}{\Delta t}$, где i — номер слоя по времени, левая часть равенства (2.52) используется для исключения C_c^i из уравнения (2.29). Таким образом, при моделировании диффузии мышьяка можно решать только одно уравнение, но дополнительно вводить рабочий массив, длина которого составляет 0.2—0.5 длин основных массивов.

В случае равновесия $\frac{\partial C_c}{\partial t} = 0$ и из уравнения (2.51) следует соотношение для концентраций полной и электрически активной примесей мышьяка

$$C_T = C_A + m k_e^{m-1} C_A^{m-1}, \quad (2.53)$$

что составляет суть второй равновесной модели кластерообразования при диффузии мышьяка ($m \approx 4$).

Время диффузионного цикла t_e , необходимое для распада 95% кластеров, связано экспериментально снятой зависимостью с коэффициентом k_d (мин^{-1}):

$$k_d \approx 3/t_e \approx 1.8 \cdot 10^3 \exp(-1.8/kT). \quad (2.54)$$

На первый взгляд, «равновесный» способ моделирования диффузии должен давать значительные погрешности при временах диффузии, сравнимых с t_e в (2.54): t_e (1000°C) = 19 мин, t_e (1100°C) = 6 мин. Равновесная концентрация достигается за

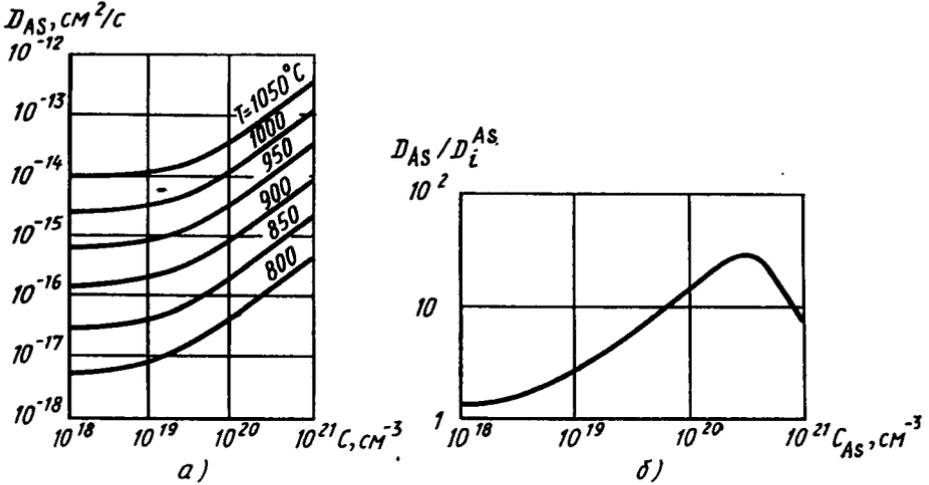


Рис. 2.18. Зависимость коэффициента диффузии мышьяка от концентрации и температуры (а); уточненная зависимость нормированного коэффициента диффузии для больших концентраций при интенсивном кластерообразовании (б)

60—80 мин при температуре 800°C . Очевидно, если время отжига значительно больше указанных времен, то можно пользоваться «равновесным» способом моделирования; в противном случае необходимо учитывать неравновесную кинетику кластерообразования. Однако в действительности погрешности моделирования не столь значительны, так как уменьшение коэффициента диффузии мышьяка D_{As} для больших C_{As} (рис. 2.18, а, б) в процессе кластерообразования ослабляет влияние этой части распределения на остальной профиль.

В первой модели более точно моделируются эффекты кластерообразования и сама динамика данного процесса, однако реализация модели в технологической программе требует несколько больших вычислительных затрат. Результаты расчета профилей с использованием двух моделей практически совпадают, если время диффузии мышьяка больше времени распада кластеров.

Диффузия сурьмы. Сурьма используется, как правило, для формирования скрытого n^+ -слоя коллектора при изготовлении высоко-

качественных интегральных транзисторных *n-p-n*-структур. Так как знание точного профиля распределения примеси слоя до последнего времени было не столь существенным, экспериментальных и теоретических работ по исследованию диффузии сурьмы значи-

тельно меньше, чем для других распространенных примесей. Лишь при переходе на транзисторные структуры БИС с тонкими эпитаксиальными слоями для работы с максимальным быстродействием в режиме высоких и сверхвысоких плотностей токов понадобилось точно определять профили распределения примеси в низколегированной коллекторной области.

Диффузия сурьмы происходит по нейтральным и отрицательно заряженным вакансиям, причем коэффициент диффузии сурьмы зависит не только от локальной электронной концентрации, но также и от поверхностной концентрации n_s и типа атмосферы. В численных

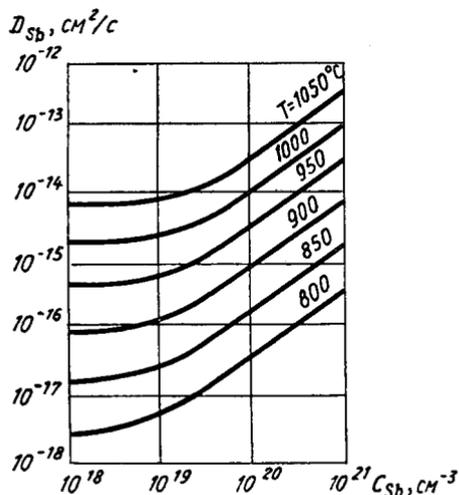


Рис. 2.19. Зависимость коэффициента диффузии сурьмы от концентрации и температуры

расчетах коэффициент диффузии Sb удобно задавать в виде

$$D_{Sb} = D_i^{Sb} \left(n_s \right)^{\frac{1 + \beta_{Sb}(T) n_n}{1 + \beta_{Sb}(T)}} = D \times + D^{-} \frac{n}{n_i}, \quad (2.55)$$

где $\beta_{Sb} (0.1 \leq \beta_{Sb} \leq 1)$ — подстроечный параметр, получаемый настройкой формулы (2.55) на экспериментальные кривые $D_{Sb}(C_{Sb})$ (рис. 2.19). Причем зависимости от поверхностной концентрации электронов собственного коэффициента диффузии определяются следующим образом:

$$D_i^{Sb} = \begin{cases} 9.7 \cdot 10^3 \exp \left[-\frac{4.87 \pm 0.13}{kT} \right], & n_s = 4.5 \cdot 10^{19} \text{ см}^{-3}, \\ 365 \exp \left[-\frac{4.44 \pm 0.13}{kT} \right], & n_s = 10^{19} \text{ см}^{-3}, \\ 143 \exp \left[-\frac{4.3 \pm 0.1}{kT} \right], & n_s = 5 \cdot 10^{18} \text{ см}^{-3}. \end{cases} \quad (2.56)$$

Отметим, что для сурьмы концентрации полной и электрически активной примеси практически полностью совпадают. Кроме того, для большинства диапазонов рабочих технологических температур и получаемых концентраций примеси удобна другая аппроксимация коэффициента диффузии

$$D_{sb} \approx 0.214 \exp\left(-\frac{3.65}{kT}\right) + 15 \frac{n}{n_i} \exp\left(-\frac{4.08}{kT}\right), \quad (2.57)$$

хорошо настраиваемая на большинство экспериментальных кривых (рис. 2.19).

Диффузия фосфора. Наиболее сложной с точки зрения моделирования является высококонцентрационная диффузия фосфора. Она происходит по нейтральным, отрицательным и дважды отрицательно заряженным вакансиям. В численных расчетах коэффициент диффузии фосфора удобно задавать, согласно выражению (2.57), в виде

$$D = D^{\times} + D^{-}n_n + D^{-}n_n^2 \quad (2.58)$$

для идеализированного профиля распределения примесей и качественного описания модели генерации вакансий, представленных на рис. 2.20. Для выявления сути модели диффузии фосфора необходимо получить зависимость коэффициента диффузии от концентрации электронов, связанной с полной концентрацией фосфора по формуле

$$C_T = n + 2.04 \cdot 10^{-41} n^3. \quad (2.59)$$

Формула (2.59) была синтезирована из обработки экспериментально полученных профилей методом Больцмана — Матано. Далее для каждого из диапазонов концентрации n фосфора были выявлены специфические температурно-концентрационные зависимости коэффициента диффузии

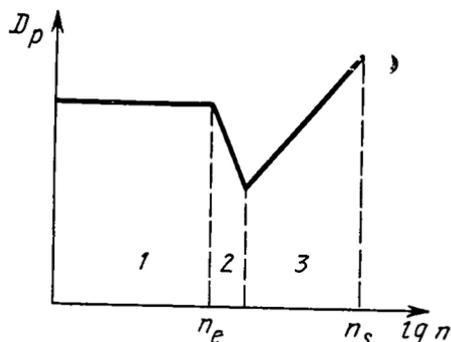


Рис. 2.21. Типовой ход зависимости коэффициента диффузии фосфора от концентрации электронов

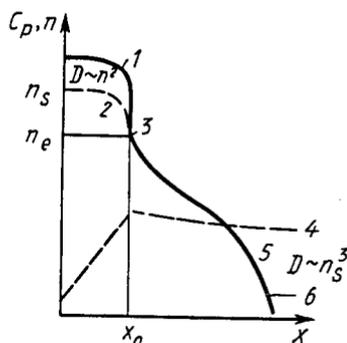


Рис. 2.20. Использование модели диффузии фосфора для описания профиля распределения примеси фосфора:

1 — полная концентрация фосфора; 2 — концентрация электрически активного фосфора; 3 — излом, распад пар pV при концентрации электронов $n = n_e$; 4 — появление избыточной концентрации вакансий; 5 — хвостовая часть; 6 — эффект выталкивания базы

физическим (рис. 2.20) с согласованием их значений на границах диапазонов концентрации n . Распад пар pV при концентрации $n = n_e$ приводит к появлению избыточной концентрации вакансий, показанной на рис. 2.20. Типичный ход зависимости $D_p(n)$ с выделением трех характерных участков: поверхностного, промежуточного, хвостового, а также температурно-концентрационные зависимости коэффициента диффузии фосфора в достаточно широком диапазоне темпе-

ратур и концентраций приведены на рис. 2.21. При высоких концентрациях доминирующая компонента коэффициента диффузии в (2.58) связана с образованием пар pV^- и соответственно с диффузией по дважды отрицательно заряженным вакансиям из-за большого значения концентрации n . Для концентраций $n > 2 \times 10^{20} \text{ см}^3$ коэффициент диффузии необходимо скорректировать из-за эффектов сужения запрещенной зоны кремния (что будет

рассмотрено выше), т. е. фактически несколько уменьшив при высоких концентрациях примесь.

При распаде пар pV^- происходит потеря электрона, в результате чего образуется пара pV^- , а затем происходит дальнейший распад; при этом облегчается создание свободных отрицательно заряженных вакансий V^- , как показано схематически пунктирной линией на рис. 2.20. Вакансии V^- диффундируют в подложку на большие расстояния и увеличивают локальный коэффициент диффузии любых имеющихся примесей.

В модели коэффициента диффузии фосфора согласно рассмотренным физическим механизмам для каждого из характерных участков: поверхностного, промежуточного, хвостового (рис. 2.20—

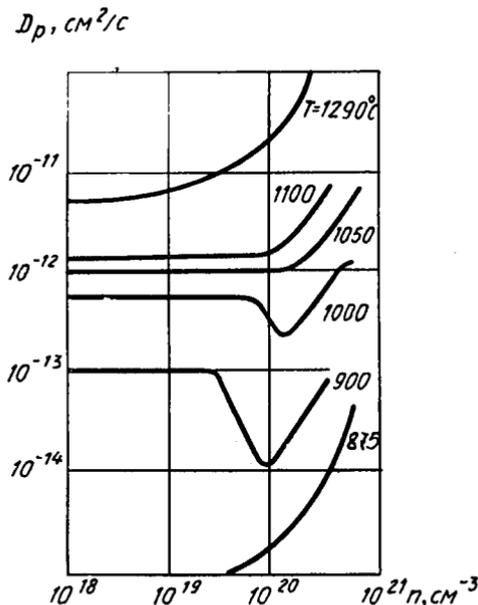


Рис. 2.22. Зависимость коэффициента диффузии фосфора от концентрации электронов и температуры

2.22) — найдены следующие зависимости:

$$D_{\text{пов}} = D^{\times} + D^{-} n_n^2, \quad (2.60)$$

$$D_{\text{хв}} = D^{\times} + D^{-} \frac{[V^-]}{[V^-]_l} =$$

$$= D^{\times} + D^{-} \left\{ \frac{n_s^2}{n_e^2 n_l} \exp\left(\frac{3\Delta E_{gm}}{kT}\right) \left[1 + \exp\left(\frac{0.3}{kT}\right) \right] \exp\left[\frac{-(x-x_0)}{L_v}\right] \right\}. \quad (2.61)$$

Здесь D^{\times} , D^{-} , D^{-} — соответствующие коэффициенты диффузии фосфора (2.37); n_s — максимальная поверхностная составляющая электронов, n_e — характерная концентрация электронов, у которых энергия уровня Ферми отлична от энергии дна зоны проводимости ($E_{Fn} = E_c - 0.12 \text{ эВ}$): $n_e \approx 9.65 \cdot 10^{21} \exp(-0.39/kT)$; ΔE_{gm} — суже-

ние запрещенной зоны кремния, вызванное главным образом механическими напряжениями кристаллической решетки; x_e — глубина, на которой достигается концентрация n_e ; L_v — длина, на которую диффундируют вакансии в подложке.

Длина L_v в общем случае зависит от ряда параметров, участвующих в технологическом процессе изготовления интегральных полупроводниковых приборов, включая концентрацию кислорода в кремниевой пластине, так как она отражает рекомбинационные процессы для избыточных вакансий. Зависимость L_v от рассмотренных параметров описана пока недостаточно хорошо. Целесообразно принять $L_v \approx 20 \div 30$ мкм, что, по существу, означает, что все области прибора, расположенные ниже области фосфора, будут увеличивать коэффициент диффузии.

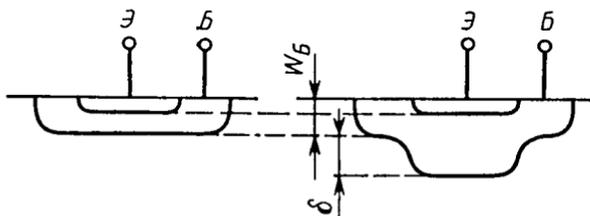


Рис. 2.23. Иллюстрация эффекта выталкивания базы из-за усиления диффузии под фосфорным эмиттером

Необходимо отметить, что на основании уравнения (2.61) это предполагает чрезвычайно сильную чувствительность коэффициента диффузии к поверхностной концентрации фосфора.

Для промежуточного участка (см. рис. 2.21) коэффициент диффузии определяется из условия согласования $D_{пр} = n^2$ и $D_{пр}(n_e) = D_{пов}(n_e)$.

Для моделирования эффекта эмиттерного выталкивания (dir—effect) в БТ (рис. 2.23) вводят соответствующий коэффициент, отражающий факт перенасыщения вакансиями участка кремния под сильнолегированным эмиттером:

$$f_{er} = D_{хн}(p)/D^*(p). \quad (2.62)$$

Согласно рассматриваемой модели диффузии фосфора вакансии V , образующиеся в результате распада пар pV^- при концентрации $n \approx n_e$, диффундируют в глубь кремния (см. рис. 2.21) и увеличивают по сравнению с равновесной концентрацию вакансий в f_{er} раз, поэтому коэффициенты диффузии остальных примесей, например бора (2.40) и сурьмы (2.55), в этой области необходимо увеличить в f_{er} раз. Диффузионная длина генерированных вакансий при температуре 900—1100°C примерно равна 20—30 мкм, что значительно меньше глубин диффузии примесей в современных и перспективных приборных структурах, поэтому спад концентрации вакансий и уменьшение f_{er} по глубине можно не учитывать.

Собственный коэффициент диффузии D_{in}^B в (2.44) в условиях генерации вакансий под областью сильнолегированного эмиттера

$$D_{in}^B = D_{i0}^B \frac{D_{xb}(p)}{D^{\times}(p)} = D_{in}^B f_{er}, \quad (2.63)$$

где D_{i0}^B — собственный коэффициент диффузии бора в отсутствие диффузии фосфора; $D_{xb}(p)$ — коэффициент диффузии фосфора, определяемый формулой (2.61).

Рассмотренная модель диффузии фосфора правильно отражает основные закономерности диффузии и вполне пригодна для инженерных численных технологических расчетов. С другой стороны, неравновесные процессы генерации вакансий, характерные для диффузии фосфора, возможно, учитываются в ней слишком упрощенно, что приводит к некоторым погрешностям моделирования. Весьма перспективно более «физичное» направление моделирования диффузии фосфора, учитывающее динамику образования и распада вакансий и ассоциативных пар за счет совместного решения системы нескольких нелинейных дифференциальных уравнений.

Однако отметим, что наблюдающаяся в последние годы тенденция использования мышьяка в качестве донорной примеси для формирования тонкослойных полупроводниковых приборов БИС несколько снижает актуальность сверхточной модели диффузии фосфора, например для класса БТ БИС.

Быстрый термический (ламповый) отжиг. Большой интерес у технологов тонкослойных стабильных полупроводниковых приборов БИС вызывает быстрый термический отжиг (БТО), позволяющий достигать практически полной активации примеси без существенного размытия профиля после имплантации. При этом для всех исследованных примесей (фосфора, бора, мышьяка и сурьмы) обнаружен аномально большой коэффициент диффузии, во много раз превышающий равновесный коэффициент диффузии (2.37) при температуре БТО. Модель БТО, предложенная в первую очередь для бора и мышьяка, сводится к следующему. Считается, что имплантация бора с дозой более 10^{15} см⁻² приводит к образованию нарушенного слоя вблизи максимума распределения концентрации, который стабилизируется за 30 с при $T = 1000 \div 1150^\circ\text{C}$. При БТО происходит интенсивная генерация точечных дефектов (вакансий) внутри этого слоя. Однако концентрация вакансий существенно превышает равновесную лишь вне нарушенного слоя, а внутри него вакансии очень быстро рекомбинируют. Поэтому внутри слоя дислокаций коэффициент диффузии бора равен равновесному значению, а вне его, в хвосте распределения, он сильно увеличивается:

$$D_B = D^V [1 + k(T) \exp(-t/\tau)], \quad (2.64)$$

где $\tau = 4.4$ с; D^V — равновесный коэффициент (2.37).

Отношение D_B/D^V в начальный момент времени достигает 100—400, но затем очень быстро уменьшается. При $t > 35 \div 40$ с $D_B \approx D^V$.

Для БТО мышьяка предлагается несколько иная модель. Коэффициент диффузии задается формулой (2.64) во всей области моделирования, а $k(T)$ пропорционален энергии имплантации E . Эта модель справедлива для энергии $50 \leq E \leq 200$ кэВ, доз мышьяка, больших $2 \cdot 10^{14}$ см⁻², и температур отжига $1000 \leq T \leq 1150^\circ\text{C}$. Влияние кластеров мышьяка учитывается независимо. При совместном БТО бора и мышьяка быстрой диффузии бора не происходит, так как генерируются вакансии только того типа, который ускоряет диффузию мышьяка. Как правило, модель дает очень хорошее совпадение с экспериментальными данными, за исключением низкоконцентрационной хвостовой области (что, возможно, связано с ошибками эксперимента) (рис. 2.24).

2.4. МОДЕЛИ ТЕРМИЧЕСКОГО ОКИСЛЕНИЯ

Термическое окисление кремния — один из наиболее важных технологических процессов для изготовления современных и перспективных полупроводниковых приборов БИС. Слой двуокиси кремния формируется на подложке за счет химического соединения в полупроводнике атомов кремния с кислородом при окислении в атмосфере либо сухого кислорода, либо влажного кислорода, либо в парах воды. Окислитель подается к поверхности кремниевой подложки, нагретой в технической печи до высокой температуры (900—1200°C). Окисление происходит гораздо быстрее в атмосфере влажного кислорода, поэтому влажное окисление используют для образования более толстых защитных пленок двуокиси кремния. Для получения тонких стабильных пленок двуокиси кремния применяют окисление в атмосфере сухого кислорода.

Физика термического окисления описывается с помощью простой модели трех потоков в статическом состоянии (рис. 2.25). Процесс окисления происходит на границе Si—SiO₂, поэтому молекулы окислителя диффундируют через все предварительно сформированные слои окисла и лишь затем вступают в реакцию на границе с окислом. Согласно закону Генри, равновесная концентрация твердой фазы пропорциональна парциальному давлению газа P :

$$C_M = H p, \quad (2.65)$$

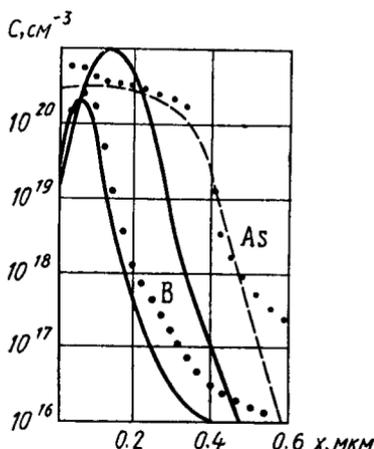


Рис. 2.24. Моделирование быстро-го термического отжига (БТО) бора и мышьяка:

— расчет по равновесной модели;
 - - - расчет по модели БТО; ● — эксперимент

где C_M — максимальная концентрация окислителя в газе для данного значения p ; H — постоянный коэффициент Генри. В неравновесном состоянии концентрация окислителя на поверхности твердой фазы меньше C_M . Поэтому поток F_1 окислителя из объема газоносителя к поверхности газ — окисел определяется разностью концентраций окислителя:

$$F_1 = h(C_M - C_0), \quad (2.66)$$

где h — коэффициент массопереноса газовой фазы; C_0 — поверхностная концентрация окислителя.

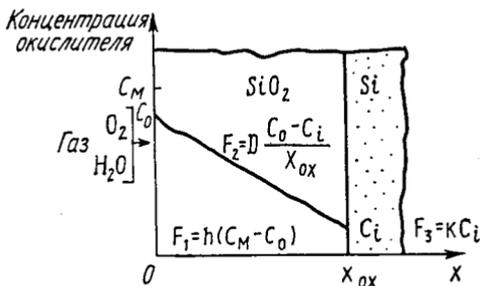


Рис. 2.25. Образование потоков в системе Si—SiO₂ при термическом окислении

Концентрация окислителя C_0 на поверхности окисла зависит в первую очередь от температуры, скорости газового потока и растворимости окислителя в окисле. Для определения скорости роста окисла вместе с потоком F_1 необходимо рассмотреть потоки окислителя в окисле F_2 и на границе окисла с кремнием F_3 (рис. 2.25).

Согласно закону Фика, поток окислителя через окисел определяется градиентом концентрации окислителя:

$$F_2 = -D \frac{\partial C}{\partial x} \approx D \frac{C_0 - C_i}{x_{ox}}, \quad (2.67)$$

где D — коэффициент диффузии при данной температуре; C_i — концентрация окислителя в окисле на поверхности окисел — кремний при $x = x_{ox}$, мол/см³; x_{ox} — толщина окисла, мкм.

Поток F_3 , соответствующий реакции окисления на границе окисел — кремний, зависит от постоянной скорости поверхностной химической реакции на поверхности k_S и определяется выражением

$$F_3 = k_S C_i. \quad (2.68)$$

В статическом состоянии все три потока должны быть равны, так что

$$F_1 = F_2 = F_3 = F. \quad (2.69)$$

Согласно (2.69), можно приравнять соотношения (2.67), (2.68) и выразить величины C_i и C_0 через C_M :

$$C_i = C_M / (1 + k_S/h + kx_{ox}/D), \quad C_0 = (1 + kx_{ox}/D) C_i. \quad (2.70)$$

Поток F_3 (2.68) с учетом соотношений (2.70) можно записать в виде

$$F_3 = F = \frac{k_S}{1 + k_S/h + kx_{ox}/D} C_M. \quad (2.71)$$

Поток окислителя, достигающий поверхности окисел — кремний, описывается обыкновенным дифференциальным уравнением

$$N_i \frac{dx_{ox}}{dt} = F_3 = F, \quad (2.72)$$

где N_i — количество молекул окислителя, необходимое для образования окисла в единичном объеме.

Заметим, что, поскольку концентрация молекул SiO_2 в окисле равна $2.2 \cdot 10^{22} \text{ см}^{-3}$, для получения двуокиси кремния требуется концентрация молекул кислорода $2.2 \cdot 10^{22} \text{ см}^{-3}$ или концентрация молекул воды $4.4 \cdot 10^{22} \text{ см}^{-3}$.

Соотношение между x_{ox} и t определяется интегрированием уравнения (2.72):

$$N_i \int_{-x_{ox}(0)}^{x_{ox}} \left(1 + \frac{k_s}{h} + \frac{kx_{ox}}{D} \right) dx = k_s C_M \int_0^t dt. \quad (2.73)$$

Решение уравнения (2.72) получается из (2.73) и записывается в виде

$$x_{ox}(t) = [A/2 + x_{ox}(0)]^2 + Bt - A/2]^{1/2}, \quad (2.74)$$

где

$$A = 2D(1/k_s + 1/h), \quad B = 2DC_M/N_i. \quad (2.75)$$

Соотношение (2.74) обычно записывается в канонической форме

$$x_{ox}^2(t) + Ax_{ox}(t) = B(t + \tau), \quad (2.76)$$

где

$$\tau = \frac{x_{ox}^2(0) + Ax_{ox}(0)}{B}. \quad (2.77)$$

Коэффициент B является параболическим коэффициентом скорости роста окисла, поскольку для больших времен окисления $t \gg A^2/4B$

$$x_{ox} \approx (Bt)^{1/2} \approx \left(t^2 \frac{2DC_M}{N_i} \right)^{1/2}. \quad (2.78)$$

Для небольших времен окисления $t \ll A^2/4B - \tau$ выполняется линейное соотношение

$$x_{ox}(t) \approx \frac{B}{A} (t + \tau) \approx \frac{k_s C_M}{N_i} (t + \tau). \quad (2.79)$$

Здесь $x_{ox}(0)$ — начальное значение толщины окисла при $t=0$, причем $x_{ox}(0) \approx 20 \text{ нм}$ для сухого кислорода и $x_{ox}(0) \approx 0$ для влажного кислорода; отношение B/A — линейный коэффициент скорости роста окисла.

Для малых времен окисления, согласно (2.79), толщина окисла определяется постоянной скорости поверхностной реакции k_S и пропорциональна времени окисления. Для больших времен окисления, согласно (2.72), рост окисла пропорционален корню квадратному из времени окисления и коэффициента диффузии,

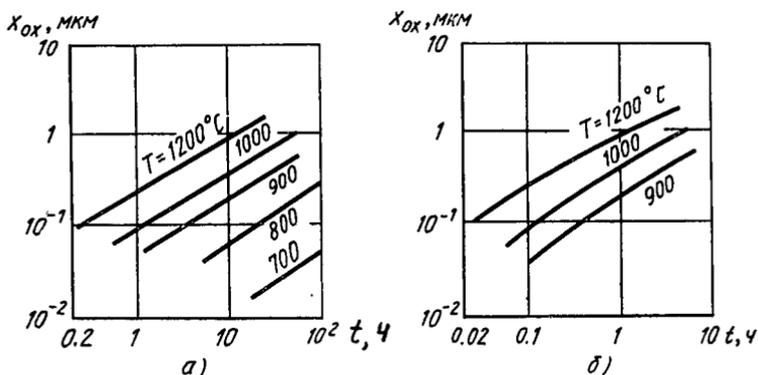


Рис. 2.26. Зависимость толщины окисла от времени и температуры термического окисления в сухом (а) и влажном (б) кислороде

Модель роста толщины окисла, представленная соотношением (2.76), в целом удовлетворительно описывает процесс термического окисления во влажном и сухом кислороде и согласуется с экспериментальными зависимостями толщины окисла x_{ox} от времени окисления в потоке сухого и влажного кислорода в достаточно широком диапазоне температур и времен (рис. 2.26).

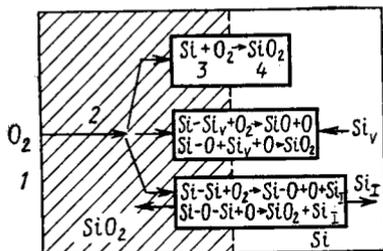


Рис. 2.27. Механизм реакции на атомном уровне на границе раздела Si—SiO₂ во время термического окисления:

1 — газ; 2 — молекулярная диффузия; 3 — дислокации; 4 — деформации

будет наблюдаться существенное несоответствие, в результате чего появятся значительные деформации (верхняя часть на рис. 2.27). На средней и нижней частях рис. 2.27 показаны реакции вакансионного и междоузельного типов, которые обосновывают роль точечных дефектов в подложке. Место соединения кислорода с атомами кремния в решетке обеспечивается либо за счет вакансий,

либо за счет внедренных атомов. Чтобы объяснить количественное увеличение скорости окисления, наблюдаемое в сильнолегированных подложках, была использована вакансионная реакция Si_V . Большое увеличение общей концентрации Si_V приводит к появлению n^+ - и p^+ -слоев, следствием чего является увеличение скорости окисления на границе раздела $Si-SiO_2$.

Для объяснения увеличения коэффициента диффузии, наблюдаемого во время термического окисления даже на расстояниях более 10 мкм от границы раздела $Si-SiO_2$, в модели используется междоузельная реакция Si_I . Реакция Si_I используется также для моделирования роста и уменьшения дефектов кристаллической структуры за счет окисления. В обоих случаях сделано допущение, что только небольшое количество ($\ll 1\%$) атомов кремния типа Si_I с границы окисления перемещается в подложку. Важным этапом моделирования кинетики термического окисления является установление относительных реакций, показанных на рис. 2.27, хотя при соответствующих условиях, в которых выполняется процесс окисления, все три механизма могут играть главные роли.

В современных программах многоэтапного моделирования технологических процессов изготовления БИС используется дифференциальная форма линейно-параболического уравнения роста окисных пленок (2.76). Для каждого последующего приращения интервала времени рассчитывают коэффициенты B , B/A с учетом характеристик подложки и окружающих условий, существующих в течение каждого временного шага.

В случае окисления влажным или сухим кислородом коэффициент B определяется выражением

$$B = B^i \left[\frac{1}{2} (p_i + p_{i-1}) \right] [1 + \delta C^{\beta}] \varepsilon, \quad (2.80)$$

где B^i — собственно параболический коэффициент; другие параметры моделируют давление, введение примесей и зависимость B от паров HCl .

Аналогично, линейный коэффициент скорости при окислении водой представляется в виде

$$\frac{B}{A} = \left(\frac{B}{A} \right)^i \left[\frac{1}{2} (p_i + p_{i-1}) \right] [1 + \gamma [V_n^i - 1] \eta [\theta]], \quad (2.81)$$

где (B/A) — собственно линейный коэффициент; другие параметры моделируют давление, введение примесей, HCl и зависимость B/A от ориентации кристаллов.

Экспериментально было найдено, что при окислении сухим кислородом коэффициент B/A имеет другую зависимость от давления и происходит «аномально» быстрое начальное окисление. В этом случае

$$\begin{aligned} \frac{B}{A} = & \left(\frac{B}{A} \right)^i \left[\frac{1}{2} (p_i^{0.75} + p_{i-1}^{0.75}) [1 + \gamma (V_n^i - 1)] \times \right. \\ & \left. \times [h(\theta)] \left[1 + k \exp\left(-\frac{x}{L}\right) \right] \right]. \end{aligned} \quad (2.82)$$

Здесь последний член моделирует режим получения тонкого окисного слоя.

При окислении сухим кислородом коэффициенты V^i и $(V/A)^i$ хорошо описываются отдельными процессами активации:

$$V^i = C_1 \exp(-E_1/kT), \quad \left(\frac{V}{A}\right)^i = C_2 \exp(-E_2/kT), \quad (2.83)$$

где $C_1 = 7.72 \cdot 10^2$ мкм²/ч; $E_1 = 1.23$ эВ; $C_2 = 6.23 \cdot 10^6$ мкм/ч; $E_2 = 2.0$ эВ.

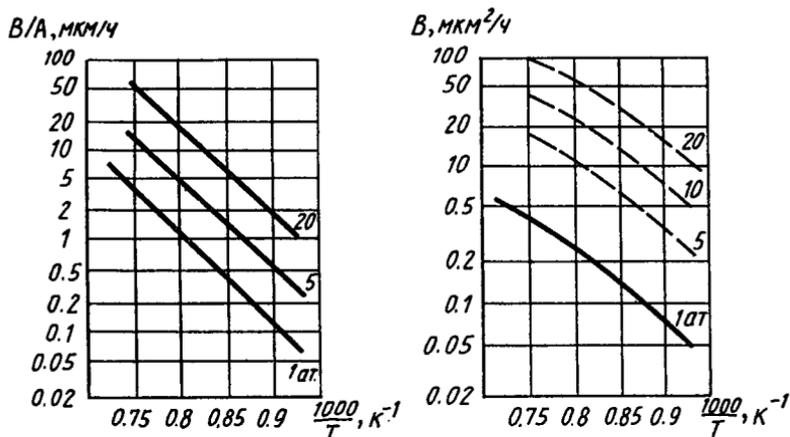


Рис. 2.28. Зависимости линейного V/A и параболического V коэффициентов скорости окисления от температуры для кремниевых кристаллов с ориентацией $\langle 100 \rangle$ и $\langle 111 \rangle$ при окислении пироженным влажным кислородом и давлении 1—20 ат

Физический процесс, определяемый коэффициентом V , представляет перенос окислителя через слой SiO_2 , а энергия активации 1.23 эВ соответствует диффузии O_2 в SiO_2 . Энергия активации 2 эВ для коэффициента V/A определяется разрывом связи Si—Si на границе раздела Si—SiO_2 , хотя, как показано на рис. 2.27, может равняться значительно больше связей.

При окислении влажным кислородом формулы для коэффициентов V^i и $(V/A)^i$ имеют вид, аналогичный (2.83), но энергия активации каждого коэффициента изменяется при температуре 900—950°C, как показано на рис. 2.28, причем для коэффициента V^i $C_1 = 4.2 \cdot 10^2$ мкм²/ч; $E_1 = 0.78$ эВ при $T > 950^\circ\text{C}$; $C_1 = 1.7 \times 10^4$ мкм²/ч, $E_1 = 1.17$ эВ при $T < 950^\circ\text{C}$; для коэффициента $(V/A)^i$ $C_2 = 1.77 \cdot 10^8$ мкм/ч, $E_2 = 2$ эВ при $T > 900^\circ\text{C}$; $C_2 = 2.07 \cdot 10^6$ мкм/ч, $E_2 = 1.6$ эВ при $T < 900^\circ\text{C}$.

Экспериментально определено, что в пироженном влажном кислороде (соответствует воде под давлением 0.92 ат) линейный и параболический коэффициенты линейно пропорциональны давлению воды, как показано на рис. 2.28. При окислении сухим

кислородом у коэффициентов В, В/А также наблюдается линейная и квазилинейная зависимость от давления.

Увеличение коэффициентов В и В/А за счет НСl описывается с помощью подстрочных коэффициентов η , ε в (2.80), (2.81), являющихся функциями температуры и концентрации.

Наращивание очень тонких окисных слоев ($<500\text{\AA}$) в атмосфере сухого кислорода необходимо описывать с помощью модифицированного физического механизма, отличного от механизма при создании более толстых слоев. Экспериментально установлено, что скорость роста окисла увеличивается примерно на порядок для толщин окисных пленок менее 200\AA (рис. 2.29). На сегодня удовлетворительного объяснения быстрой кинетики роста окисных пленок, других их электрофизических и оптических свойств не существует. Однако большинство экспериментальных данных можно согласовать с адекватными эмпирическими данными, если скорость окисления на первоначальной стадии описывать соотношением

$$\frac{dx_{ox}}{dt} = \frac{B}{2x_{ox} + A} + k \exp(-x_{ox}/L);$$

$$L = 70\text{\AA}, \quad k = 7 \cdot 10^7 \exp(-2.35/kT). \quad (2.84)$$

Отметим, что при окислении во влажном кислороде аномальное возрастание скорости роста окисной пленки либо не наблюдалось, либо происходило при $x_{ox} < 25\text{\AA}$. Поскольку на первоначальной стадии роста окисной пленки доминирующую роль играет линейный коэффициент В/А, этот процесс, отраженный соотношением (2.84), эквивалентен возрастанию коэффициента В/А. Характерная длина затухания L в (2.84) практически не зависит от температуры. Параметр k в (2.84) является функцией температуры при энергии активации 2.35 эВ для кремния с ориентацией $\langle 111 \rangle$ и $\langle 100 \rangle$ и энергии активации 1.8 эВ для кремния с ориентацией $\langle 110 \rangle$.

С практической точки зрения для модификации коэффициента В/А можно использовать одноэкспоненциальную зависимость от пространственной, а не от временной переменной, что дает очень хорошее приближение к экспериментальным данным для толщин

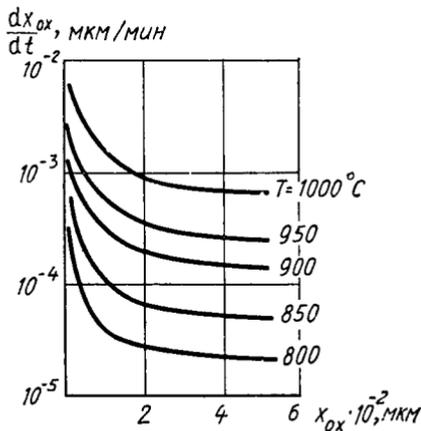


Рис. 2.29. Зависимость скорости окисления от толщины окисла и температуры для тонких слоев окисла кремния с ориентацией $\langle 111 \rangle$

окисла более 50 Å. Комбинация эффектов окисла и сильного легирования дает следующее соотношение для коэффициента В/А:

$$\frac{B}{A} = \left(\frac{B}{A}\right)^i \left\{ [1 + \gamma(c^n - 1)] [1 + ke^{-x/L}] \prod_{l=1}^n \right\},$$

где включены члены, описывающие элементы тонкого окисла и сильного легирования. Член, произвольно обозначенный как произведение по i -му эффекту, предполагает необходимость учета парциального давления, ориентации, влияния НСІ.

Модели процессов диффузии под окисляющейся поверхностью. Окисляющаяся граница раздела Si—SiO₂ инжектирует в кремниевую подложку избыточные атомы кремния (Si_I). Это перенасыщение внедренных атомов вызывает рост дефектов кристаллической структуры и изменяет коэффициент диффузии.

Если влияние механизма внедренных атомов значительно, то полный коэффициент диффузии примесей можно записать в виде суммы эффективных коэффициентов диффузии, соответствующих механизмам вакансий и внедренных атомов, с учетом их относительного вклада в полную диффузию:

$$D = D^V + D^I, \quad (2.85)$$

где D^V , D^I — коэффициенты диффузии за счет вакансий и внедренных атомов соответственно.

При тепловом равновесии, отсутствии окисления для низколегированного кремния можно считать, что диффузия примеси за счет внедренных атомов составляет некоторую часть f_I от диффузии за счет вакансий:

$$D_i = D_i^V + f_I D_i^V, \quad (2.86)$$

где оба выражения фактически входили в ранее приводимые уравнения для D_i [см. (2.38)]. При более общих условиях коэффициенты D^V будут возрастать при увеличении концентрации вакансий Si_V (например, при легировании примесью), Si_I. Таким образом, имеем

$$\left\{ \begin{aligned} D = D^V + f_I D_i^V \frac{[Si_I]}{[Si_I]_i}, \end{aligned} \right. \quad (2.87)$$

где $[Si_I]_i$ — концентрация внедренных атомов кремния при отсутствии окисления, а $[Si_I]$ — концентрация во время окисления.

В дополнение к независимым эффектам, определяющим концентрацию Si_V и Si_I в процессе легирования или окисления, вероятно также, что эти две группы взаимодействуют между собой через реакцию



где Si_S представляет атом решетки кремния, а коэффициент K зависит от температуры T .

Применение уравнений типа (2.87) к данным окислительного усиления диффузии показало, что избыточная концентрация Si_I ведет себя как $(dx_{ox}/dt)^n$, где dx_{ox}/dt — скорость окисления, а n — эмпирически заданная постоянная. Значения n колеблются от 0.3 до 1.0 с наибольшим группированием в диапазоне 0.4—0.5. В технологических программах используется обычно значение $n \approx 0.5$.

Экспериментально было также установлено, что присутствие во время окисления HCl уменьшает эффект окислительного усиления диффузии. Хлор на границе раздела Si—SiO₂ вступает в реакцию с кремнием, в результате чего генерируются Si_V или расходуются Si_I. Уменьшение эффекта окислительного усиления диффузии пропорционально $(p_{HCl})^{0.5}$, где p_{HCl} — парциальное давление HCl.

Объединив все эти эффекты, получим выражения для коэффициентов диффузии примесей:

$$D = D_V + \Delta D_{ox}, \quad (2.89)$$

$$D_N = D^V + f_I D_I^V \frac{[Si_V]_I}{[Si_V]}, \quad (2.90)$$

$$\Delta D_{ox} = f_I D_I^V \frac{[Si_V]_I}{[Si_V]} \left(\frac{dx_{ox}}{dt} \right)^{0.5} \frac{k_1(T) \exp(-x/L_I)}{\rho [1 + k_2(T) P_{HCl}^{0.5}]}. \quad (2.91)$$

Коэффициент диффузии легко разделить на две компоненты. Первая компонента D_N представляет собой коэффициент диффузии примеси в отсутствие окисления. Она состоит из коэффициента D^V , определяющего роль вакансий в диффузии, и члена, определяющего вклад в диффузию внедренных атомов. В действительности D_N очень хорошо моделируется полностью на основании механизма вакансий. Если значение f_I мало, то вклад внедренных атомов в диффузию сравнительно невелик, особенно в условиях примесного легирования, когда возрастают D^V и $[Si_V]$. Однако в режиме окисления к общему коэффициенту диффузии добавляется компонента ΔD_{ox} , соответствующая вкладу в диффузию механизма внедренных атомов. Это окислительное усиление диффузии в зависимости от условий окисления может даже превосходить D_N . Коэффициент ρ определяет наблюдаемую разницу эффектов окислительного усиления диффузии для кристаллографических направлений $\langle 111 \rangle$, $\langle 100 \rangle$.

2.5. МОДЕЛИРОВАНИЕ ГРАНИЧНЫХ ПРОЦЕССОВ

При моделировании полного технологического процесса изготовления полупроводниковых структур БИС приходится сталкиваться с большим количеством физико-химических явлений, протекающих на самых разнообразных границах раздела Si—SiO₂, SiO₂ — вакуум, Si — диффузانت в газовой фазе и т. д. Механизмы и кинетические особенности этих процессов довольно сложны и детально пока не изучены. В общем случае моделирования многослойных

полупроводниковых структур необходимо рассмотреть перенос или перераспределение (сегрегацию) примесей через границы раздела соответствующих материалов и областей и далее точно спрогнозировать результирующие потоки примесей. Эффекты сегрегации, возникающие в системе Si—SiO₂, пояснены качественными концентрационными зависимостями от координаты для различных типов диффузии и материалов (рис. 2.30), где m_s — коэффициент термической равновесной сегрегации, больший или меньший единицы и зависящий от параметров технологического режима, типа решетки полупроводника и др. В случае легирования подложки

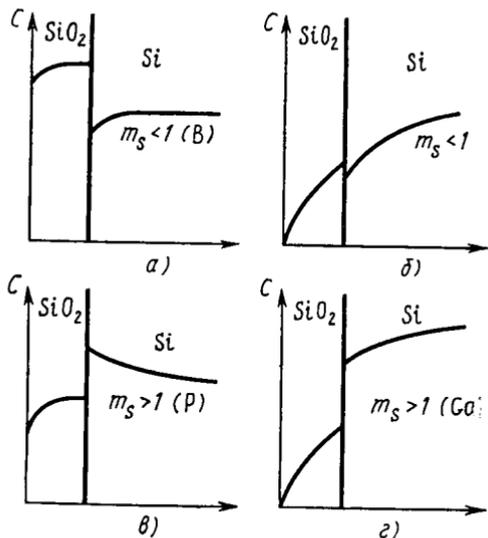


Рис. 2.30. Эффекты перераспределения примесей:

а— $m_s < 1$, медленный диффузانت в SiO₂ (бор); б— $m_s < 1$, быстрый диффузонт в SiO₂; в— $m_s > 1$, медленный диффузонт в SiO₂ (фосфор); г— $m_s > 1$, очень быстрый диффузонт в SiO₂ (галлий)

фосфором, мышьяком и сурьмой концентрация атомов на границе раздела увеличивается, а при легировании бором происходит обеднение приповерхностной области.

Как правило, точное описание физико-химических процессов переноса и сегрегации требует решения сложных нелинейных дифференциальных уравнений в частных производных с большим количеством параметров, значения которых известны лишь приближенно для конкретных граничных процессов. С другой стороны, большинство граничных процессов (за исключением сегрегации) не очень сильно влияют на окончательное распределение примесей в полупроводниковых структурах, поэтому они могут быть

рассчитаны с помощью более простых аналитических моделей. Наиболее простым, но весьма эффективным подходом к численному моделированию граничных процессов, ориентированных на расчет многоэтапных технологических процессов, является описание перемещения границы раздела кинетической моделью первого порядка.

Данный способ основан на использовании потоков, описывающих массоперенос через границу областей 1 и 2 кинетическим уравнением первой степени:

$$F_{1,2} = b(C_1 - C_2)/m_s, \quad (2.92)$$

где b — коэффициент поверхностного массопереноса (в единицах скорости); C_1, C_2 — концентрации диффузанта в областях 1 и 2.

Изменением коэффициентов в уравнении (2.92) можно с удовлетворительной точностью моделировать разнообразные граничные процессы.

1. Испарение из Si или SiO₂ моделируется уравнением (2.92) при $C_1=0$, $m_s=1$ и константой испарения $b=V_{ev}(T)$, определяемой экспериментально. Так как в ряде экспериментов получена примерно одинаковая величина V_{ev} для фосфора и бора, то в технологических программах принята единая константа испарения для всех диффузантов при испарении с поверхности Si и SiO₂: $V_{ev}=6.6 \cdot 10^{-3} (-3.7/kT)$ см/с.

2. Загонка из источника с постоянной поверхностной концентрацией моделируется уравнением (2.92) при $m_s=1$, $b \rightarrow \infty$, при этом концентрация C_1 полагается равной постоянной поверхностной концентрации. Условие $b \rightarrow \infty$ в расчетах заменяется достаточно большим значением параметра b . Использование уравнения (2.92) означает быстрое установление равенства $C_1=C_2$. Однако необходимо иметь в виду, что решение уравнения диффузии (2.29) с граничным условием первого рода ($C_1=C_2$, без установления) сопряжено с трудностями обеспечения сходимости итерационного процесса при низких температурах диффузии.

3. Для неподвижной границы раздела Si—SiO₂ в уравнении (2.92) m_s представляет собой экспериментально определяемый коэффициент сегрегации. На основании экспериментов для фосфора и мышьяка получен коэффициент $m_s=50$, для бора — зависимости $m_s(T)$: при окислении в сухом кислороде $m_s=13.4 \exp(-0.33/kT)$ и во влажном кислороде с различной ориентацией $m_s\langle 111 \rangle = 65.2 \exp(-0.66/kT)$, $m_s\langle 100 \rangle = 104 \exp(0.66/kT)$. Отметим, что процесс окисления для диффундирующего бора можно считать влажным уже при небольших следах паров воды.

На основании экспериментальных исследований получено, что коэффициент b , по крайней мере для бора, значительно больше скорости роста окисной пленки v_{ox} даже в начале окисления. Поэтому для донорной примеси целесообразно выбрать $b=0.1$ м/с, причем выбор достаточно большого значения b слабо влияет на результаты моделирования.

4. Для движущейся границы раздела Si—SiO₂ вместе с потоком (2.92) необходимо дополнительно учитывать еще один поток

$$F_b = -v_{ox}(C_1 - \gamma C_2), \quad (2.93)$$

вызванный увеличением объема SiO₂ по сравнению с объемом кремния Si, затраченного на его образование.

В (2.93) v_{ox} — скорость движения; γ — коэффициент отношения объемов Si и SiO₂: $\gamma = Q_{Si}/Q_{SiO} = 0.44$.

Детали использования потока (2.93) вместе с основным потоком (2.92) в случае движущейся границы раздела Si—SiO₂ будут рассмотрены ниже.

5. В случаях роста эпитаксиальной пленки и перераспределения примеси в скрытом слое собственно эпитаксиальный процесс опи-

сывается простой моделью с двумя параметрами: скоростью роста пленки $v_{\text{эп}}$ и средней концентрацией примеси в пленке. Более важным и актуальным в технологии является моделирование перераспределения примеси в скрытом n^+ -слое в процессе высокотемпературной эпитаксии, так называемое размытие скрытого слоя. Для этого в модели на границе раздела газ — эпитаксиальная пленка используются два потока аналогично (2.92), (2.93), а именно (рис. 2.31): поток испарения примеси

$$F_3 = v_{\text{эп}}(k_s C_1 - C_2), \quad (2.94)$$

где k_s — поверхностный коэффициент сегрегации;

поток, вызванный движением границы раздела газ — эпитаксиальная пленка,

$$F_B = v_{\text{эп}}(C_1 - C_2). \quad (2.95)$$

Численные методы решения в случае движущейся границы раздела с использованием потоков, описываемых выражениями (2.94), (2.95), будут рассмотрены ниже.

2.6. МОДЕЛИРОВАНИЕ ПРОЦЕССА ЭПИТАКСИИ

Как известно, эпитаксиальные слои являются неотъемлемой частью биполярных полупроводниковых структур и приобретают все большее распространение и значимость в МОПТ. Для БИС, в первую очередь для сверхскоростных БИС, наиболее актуальна разработка тонкослойных субмикронных БТ, способных работать с максимальными скоростями переключения в режиме высоких и

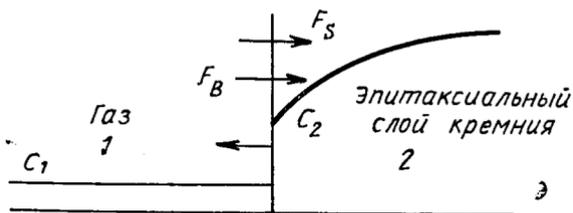


Рис. 2.31. Модель массопереноса через границу раздела двух сред при моделировании процесса эпитаксии

сверхвысоких плотностей тока. Для разработки таких интегральных БТ необходимо получать прецизионные эпитаксиальные слои с заданной стабильной концентрацией. Однако создание эпитаксиальных слоев с равномерным легированием может оказаться чрезвычайно затруднительным из-за неравномерности распределения легирующего вещества на начальной границе эпитаксиальный слой — подложка. Данная ситуация иллюстрируется на рис. 2.32, где качественно представлены нежелательные эффекты «размытия» соответствующих полупроводниковых слоев и нарушения равномерного легирования эпитаксиальных слоев в ходе высокотемпературных технологических операций.

В технологических программах желательно иметь возможность моделирования эпитаксиальных процессов, позволяющих реализо-

вать однородные и неоднородные профили распределения примесей.

Рассмотренный способ моделирования граничных процессов «размытия» скрытого слоя в БТ на основе кинетических уравнений двух потоков не всегда характеризуется удовлетворительной точностью, особенно для полупроводниковых структур с тонкими эпитаксиальными слоями.

В качестве примера, поясняющего модель легирования эпитаксиальных слоев, рассмотрим процесс роста слоев кремния из смеси силана SiH_4 с водородом. Источником легирующей примеси служит арсин AsH_3 . Для расчета термического перераспределения примеси в объеме

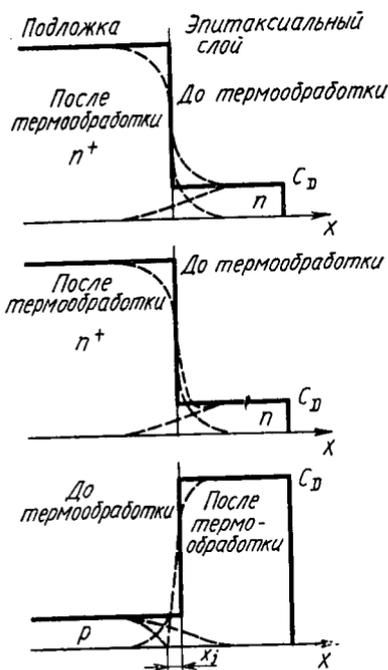


Рис. 232. Влияние эффекта автодиффузии на профиль распределения примеси



Рис. 233. К объяснению процесса роста эпитаксиального слоя

структуры в процессе роста эпитаксиального слоя используется одномерное уравнение диффузии (2.33), решаемое для области $\infty > Z > Z_f$ (рис. 2.33):

$$\frac{\partial c(Z, t)}{\partial t} = \frac{\partial}{\partial Z} \left(D \frac{\partial c}{\partial Z} \right). \quad (2.96)$$

Решение уравнения (2.96) удовлетворяет начальным и граничным условиям:

$$c(Z, 0) = f_1(Z), \quad D \frac{\partial c}{\partial Z} \Big|_{Z \rightarrow \infty} = 0, \quad D \frac{\partial c}{\partial Z} \Big|_{Z=Z_f} = f_2(t), \quad (2.97)$$

где функция $f_1(t)$ описывает диффузионный профиль распределения примеси перед эпитаксиальным наращиванием, граничное условие $D \frac{\partial c}{\partial Z} \Big|_{Z \rightarrow \infty} = 0$ означает равенство нулю потока примесей

ных атомов на большом расстоянии от подложки. Последнее граничное условие (2.97) учитывает, что в ходе эпитаксиального роста диффузионный поток примесных атомов внутрь растущего слоя на границе раздела газ — кристалл является функцией времени. Выражение для функции получается из математического описания механизма захвата примеси решеткой кремния в процессе роста эпитаксиального слоя.

Последовательность процессов, происходящих в газовой фазе эпитаксиального реактора, включает в себя такие основные этапы:

- 1) принудительный массоперенос арсина от входа в камеру к области эпитаксиального наращивания;
- 2) массоперенос молекул арсина из основного газового потока через граничный слой к поверхности, где происходит эпитаксиальное наращивание;
- 3) диссоциацию молекул арсина посредством газофазных химических реакций на мышьяксодержащие частицы;
- 4) адсорбцию мышьяксодержащих частиц на поверхности;
- 5) химическую диссоциацию арсина в адсорбированном слое;
- 6) диффузию на поверхности и встраивание адсорбированных атомов As в наращиваемый эпитаксиальный слой;
- 7) зарастание встроенных в решетку атомов Si за счет последующего подхода атомов Si в процессе наращивания;
- 8) десорбцию водорода с поверхности.

Для последовательности процесса эпитаксиального наращивания граничное условие (2.97) можно записать в виде

$$D \frac{\partial C}{\partial Z} \Big|_{z=z_f} = f_2(t) = -k_{mf} \left[p_{D_0} - \frac{C(Z_f)}{k_p} \right] + gC(Z_f) + k_A \frac{\partial C(Z_f)}{\partial t}. \quad (2.98)$$

Здесь k_{mf} — кинетический коэффициент, связанный с механизмом, определяющим процесс захвата решеткой легирующей примеси; p_{D_0} — парциальное давление на входе; $C(Z_f)$ — концентрация легирующей примеси на границе раздела фаз; k_p — коэффициент сегрегации, связывающий концентрацию легирующей примеси в эпитаксиальном слое с концентрацией атомов легирующей примеси в газовой фазе; k_A — коэффициент, связывающий концентрацию легирующей примеси в эпитаксиальном слое с концентрацией атомов этой примеси в адсорбированном слое.

Первый член в правой части уравнения (2.98) описывает поток легирующих атомов, адсорбирующихся на поверхности (этапы 4—6). Второй член определяет скорость уменьшения концентрации атомов легирующей примеси в адсорбированном слое. Третий член соответствует диффузии атомов легирующей примеси.

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ МОДЕЛИРОВАНИЯ ТЕХНОЛОГИИ ИЗГОТОВЛЕНИЯ БИС

3.1. НАЗНАЧЕНИЕ ТЕХНОЛОГИЧЕСКИХ ПРОГРАММ

Общее назначение программного обеспечения моделирования технологических процессов изготовления БИС состоит в предоставлении разработчику технологии полупроводниковых структур возможности создать точную модель полной кремниевой технологии и рассчитать необходимые профили распределения примеси в структурах. Первые попытки машинной реализации моделей технологических процессов изготовления ИМС и БИС, предпринятые в ряде отечественных и зарубежных лабораторий, относятся к середине 1960-х — началу 1970-х годов. Однако уже первые попытки создания программного обеспечения расчета технологий показали, что чисто математические и вычислительные аспекты, например численного решения задачи диффузии — окисления с постоянными коэффициентами диффузии или другими эмпирическими параметрами, даже для двумерного случая не являются главными проблемами, от решения которых зависит прогресс развития данной ветви многоуровневого моделирования БИС.

Основным фактором, обусловившим развитие моделирования процессов изготовления ИМС и БИС, было отсутствие точных и гибких моделей кинетики процессов и жизнеспособных механизмов взаимодействия пользователя с моделями. Отметим, что реально пользователи программ неизбежно будут использовать технологические модели и параметры, выходящие за рамки заданной библиотеки моделей и кинетических параметров определенного программного обеспечения. Однако даже в условиях расчета новых перспективных технологий вероятность получения неудачных результатов для эмпирических моделей будет намного больше, чем для точных кинетических моделей, основанных на самых современных передовых представлениях с учетом выявленных теорией и экспериментом физических механизмов.

При бурном прогрессе новых технологических методов изготовления БИС трудно предположить, что промышленные лаборатории и предприятия пойдут по пути эмпирической разработки кристаллов БИС с огромными затратами средств и времени на оптимизацию таких технологий. Более притягательной альтернативой является понимание и формулировка точных механизмов и моделей основных физико-химических процессов и их представление в виде

всеобъемлющей технологической программы для ЭВМ. Такая технологическая программа, развитая и реализованная на ЭВМ, позволяет предсказывать профили распределения примесей и другие структурно-технологические параметры полупроводниковых приборов в зависимости от любой предлагаемой последовательности технологических режимов и операций их изготовления.

Численная реализация на ЭВМ технологического моделирования, хотя и не является главной проблемой, сдерживающей развитие моделирования процессов, интересна как самостоятельная задача, связанная с реализацией точных физико-химических кинетических моделей. Численные методы позволяют получить более точное отображение физики процессов в развитых кинетических моделях на различных этапах технологических процессов. Кроме того, они обеспечивают более точное математическое решение проблемы и применимы к гораздо более широкому диапазону технологических условий.

Точность и общность развитых кинетических моделей технологических операций, для которых нужно получить численное решение, достигается за счет увеличения их сложности. Поперечное сечение полупроводниковых приборов в одномерном или двумерном отображении должно быть представлено пространственно-дискретной формой — в виде множества отдельных элементов. Время протекания процесса, например продолжительность этапа окисления, необходимо также представлять в дискретной форме. Такая пространственно-временная дискретизация позволяет сделать важные допущения о том, что в течение шага временного квантования количество примесей, коэффициенты диффузии и другие физические параметры каждого из элементов остаются постоянными.

Входные данные современных технологических программ представляют, по существу, режим обработки, определяющий последовательность времен, температуру, условия окружающей среды (тип атмосферы, парциальное давление и др.), а также основное и предварительное осаждение, дозу и энергию ионного легирования, режимы процесса эпитаксии и т. д. Выходные данные технологической программы после каждого шага в последовательности процессов технологии представляют профили примесей (одномерные и двумерные) в кремниевой подложке и в поверхностном слое SiO_2 , причем полупроводниковая структура может содержать дополнительные слои материала под подложкой и/или SiO_2 , включая поликремний и нитрид кремния. Во время каждого шага технологического процесса программно рассчитывают толщину, профили присадок в каждом слое, их удаление с поверхности, сопотивление. В большинстве современных технологических программ для точного расчета электрических параметров изготовленной полупроводниковой структуры в конце моделирования получают решение уравнения Пуассона.

3.2. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ОДНОМЕРНОГО МОДЕЛИРОВАНИЯ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ИЗГОТОВЛЕНИЯ БИС

В этом параграфе рассмотрим численный метод решения системы уравнений диффузии, особенности и возможности программной реализации технологического моделирования. Все описанные в гл. 2 модели основных технологических операций для наиболее распространенных примесей — бора, фосфора, мышьяка и сурьмы — используются в программе одномерного моделирования технологии изготовления БИС. При этом реализуется один из наиболее эффективных численных методов решения системы уравнений диффузии.

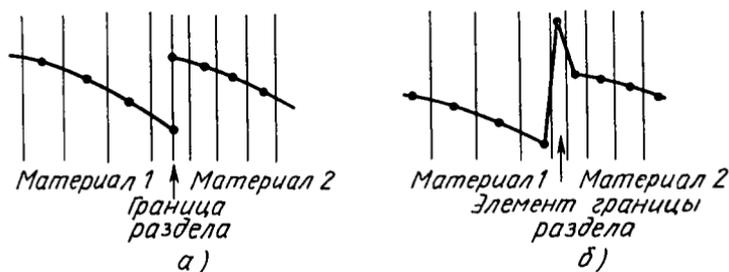


Рис. 3.1. Схематическое представление пространственных полупроводниковых структур:
 а — модель с резкой границей раздела; б — модель с границей раздела конечной толщины

Запись уравнения непрерывности для концентраций (2.29) в интегральной форме

$$\frac{d}{dt} \int_{V(t)} C_1 dV = \int_{V(t)} (g - l) dV - \int_{S(t)} \mathbf{F}(C_1, C_2, C_3) \mathbf{n} dS, \quad (3.1)$$

где C_1, C_2, C_3 — концентрации примесей (например, доноров эмиттера, акцепторов базы и доноров скрытого слоя коллектора); $V(t)$ — объем интегрирования (зависящий в общем случае от времени), ограниченный поверхностью $S(t)$; \mathbf{n} — нормаль к поверхности $S(t)$, позволяет получить абсолютно устойчивую разностную схему. В численном методе реализуется основной принцип, согласно которому движущаяся граница раздела Si—SiO₂ находится внутри области решения, и его большая гибкость заключается в использовании рассмотренных ранее двух потоков $F_S - F_B$, описывающих массоперенос через движущуюся границу Si—SiO₂.

Вся область моделирования от границы раздела Si (SiO₂) с газом на одном конце до края подложки на другом покрывается сеткой с множеством узлов $\{y_i, i=1, 2, \dots, mm\}$, в общем случае неравномерно с шагом $H_i = H_0 g^i, g \approx 1,2$. На рис. 3.1 показаны

примеры разбиения для некоторого момента времени. Одна из реализаций пространственной дискретизации структуры Si—SiO₂ показана на рис. 3.1, а, на котором виден участок границы раздела между двумя слоями материала. Толщина элементов, которыми представлены материалы по обе стороны границы раздела, в конкретном слое одинакова. Различная плотность двух материалов (например, кремния и двуоксида кремния) может привести к изменению шага сетки, в этом случае перемещение подвижной границы раздела создает неравномерность сетки вблизи границы раздела.

Поток примесей через границу раздела определяется потоком при сегрегации. Перемещение границы раздела вызывает дополнительный поток $F = -v_{ox}(\gamma C_{I-1} - C_I)$, который необходим для компенсации относительного расширения слоя окисла вследствие разницы плотностей, когда граница раздела пересекает границу элемента.

Несмотря на правомерность подобного подхода при значительных временах диффузии, эксперименты показали неадекватность используемых ранее моделей границы раздела. Переходная область между двумя материалами имеет участок конечной толщины, который служит ловушкой для примесей, оказавшихся вблизи границы раздела. Для построения модели, которая точно учитывала бы эффект, вносимый границей раздела, предлагается видоизмененная структура сетки (рис. 3.1, б) с промежуточным элементом конечной толщины, заменяющим скачкообразную границу раздела.

Пограничный поток, вызванный перемещением границы раздела, не зависит от такого изменения, но поток при сегрегации через скачкообразную границу раздела необходимо заменить двумя уравнениями, описывающими поток через границы промежуточного элемента:

$$F_m = h [C_I + (C_I + 1)/m_m], \quad (3.2)$$

$$F_n = h (C_{I-1} - C_I/m_n). \quad (3.3)$$

В выражениях (3.2), (3.3) F_m — поток между элементом, заменяющим границу раздела, и нижележащим слоем; F_n — поток между вышележащим слоем и элементом границы раздела; m_m, m_n — соответствующие коэффициенты сегрегации.

Включение в модель промежуточного элемента имеет чрезвычайно большое значение при малой глубине перехода и быстропротекающих процессах диффузии, характерных для большинства современных и перспективных технологий БИС. Тепловое перераспределение примесей между элементами, а не на границе раздела обусловлено диффузионными потоками, которые вычисляются с помощью уравнения (3.1). Границы интегрирования выбирают практически всегда посередине между соответствующими узлами. Это позволяет минимизировать ошибку аппроксимации при замене формулы потока (2.36) разностным аналогом, но существенно усложняет вычисление объемов около движущейся границы раздела Si—SiO₂.

Для любого узла пространственной сетки разностное уравнение записывается в виде

$$\frac{Q_i^l - Q_i^{l-1}}{t^l - t^{l-1}} = -[F_{i+1/2} - F_{i-1/2}] + G_i - L_i. \quad (3.4)$$

Здесь l — номер слоя по времени; i — номер узла пространственной сетки; $F_{i+1/2}$ — поток, текущий из ячейки i в ячейку $i+1$; G_i , L_i — интегралы скоростей образования и уничтожения примесей, причем

$$Q_i^l = C_i^l (y_{i+1}^l - y_{i-1}^l) / 2, \quad i \geq 2, \quad (3.5)$$

$$Q_i^{l-1} = C_i^{l-1} (y_{i+1}^{l-1} - y_{i-1}^{l-1}) / 2, \quad i \geq 2, \quad i \neq i_B + 1, i_B + 2,$$

где i_B — номер узла границы раздела Si—SiO₂.

Для $i \geq 2, i \neq i_B$

$$F_{i+1/2} = -D_{i+1/2} \left(\frac{C_{1,i+1}^l - C_{1,i}^l}{y_{i+1}^l - y_i^l} \right) h_{i+1/2} - \left\{ (h_{i+1/2} - 1) \left[\frac{C_{2,i+1}^l - C_{2,i}^l}{y_{i+1}^l - y_i^l} \right] \right\} \quad (3.6)$$

и $F_{1,5}$ есть разностная аппроксимация одного из граничных потоков (3.2), (3.3) или (2.100).

Как указывалось, для правой границы раздела справедливо граничное условие $C_{mm} = 0$ или $C_{mm} = C_{mm-1}$ (отражающая граница). Для неподвижной границы раздела Si—SiO₂ имеем $F_{i_B+1/2} = b(C_{i_B} - C_{i_B+1}/m_s)$, где обозначения приведены ранее (см. § 2.6).

Рассмотрим теперь движущуюся границу раздела Si—SiO₂ как наиболее важный и сложный случай. Здесь могут возникнуть две ситуации:

1. За интервал времени $[t^{l-1}, t^l]$ граница раздела Si—SiO₂ не пересекает границы ячеек i_B+1 и i_B+2 . При этом поток F_b (2.93), связанный с изменением объемов при превращении Si в SiO₂, находится все время «внутри» ячейки i_B+1 и не вносит вклада в уравнение (3.4). Границы ячеек при этом изменяют так, чтобы они были посередине между новыми узлами сетки, как показано на рис. 3.2. Из рисунка

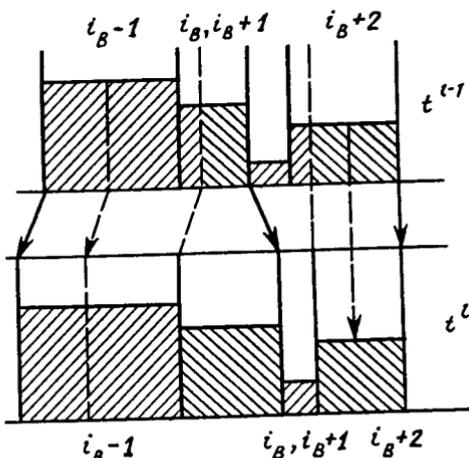


Рис. 3.2. Перераспределение объемов и изменение границ интегрирования при моделировании движущейся границы раздела (граница раздела Si—SiO₂ не пересекает границы ячеек i_{B+1} и i_{B+2})

«внутри» ячейки i_B+1 и не вносит вклада в уравнение (3.4). Границы ячеек при этом изменяют так, чтобы они были посередине между новыми узлами сетки, как показано на рис. 3.2. Из рисунка

видно, что если Q_i^t для уравнения (3.4) определяется по формуле (3.5), то для вычисления Q_i^{t-1} необходимо проследить, где были бы границы соответствующей ячейки в момент времени t^{t-1} , если бы объемы S_i и SiO_2 не изменялись. Эти границы показаны в верхней

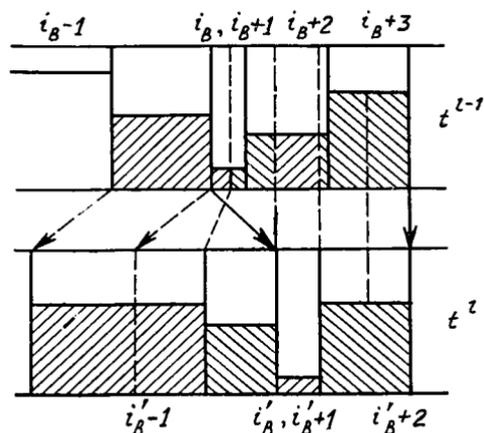


Рис. 3.3. Перераспределение объемов и изменение границ интегрирования при моделировании движущейся границы раздела (граница раздела Si—SiO₂ пересекает границу ячейки i_{B+1})

части рисунка штриховыми линиями, продолжающими новые границы ячеек в нижней части рисунка. Таким образом, в левой части формулы (3.4) должна стоять разность одинаково заштрихованных объемов, находящихся в нижней и верхней частях рисунка.

2. За интервал времени $[t^t, t^{t-1}]$ граница раздела Si—SiO₂ пересекает границу ячейки i_{B+1} (рис. 3.3). Тогда при $t > t_0$ можно считать, что граница находится в узле $i'_B = i_{B+1}$, а ячейка i'_B превращается из «кремневой» в «окисную». Именно тогда и действует поток F_B на границе между ячейками. В данном случае (3.4) транс-

формируется в уравнение

$$\frac{Q_i^t - Q_i^{t-1}}{t^t - t^{t-1}} = -[F_{i+1/2} - F_{i-1/2}] + G_i^- L_i^- \pm \pm \frac{t^t - t_0}{t^t - t^{t-1}} v_{ox} (C_{i'B}^t - \gamma C_{i'B-1}^t), \quad (3.7)$$

причем знак плюс берут для ячейки $i_{B'-1}$, знак минус — для $i_{B'}$.

Численное моделирование процесса эпитаксиального наращивания. На рис. 3.4, а—г показан более эффективный алгоритм использования метода конечных разностей для моделируемой области эпитаксиального наращивания. Приповерхностная область приборной кремниевой структуры разбивается на дискретные элементы, внутри каждого из которых концентрация считается постоянной: $C = C_{i+1} (Z_i + Z_{i+1})/2 \leq Z \leq (Z_{i+1} + Z_{i+2})/2$.

В начальный момент времени $t=0$ профиль легирования определяется условием $C(Z, 0) = f_1(Z)$. Процесс моделирования начинается с добавления нового элемента Z_{i-1} (рис. 3.4, б). Концентрация легирующей примеси C_{i-1} внутри нового элемента определяется из уравнения эпитаксиального наращивания (2.97) путем приравнивания левой части этого уравнения нулю:

$$k_{mf} \left(p_{D0} - \frac{C_{i-1}}{k_p} \right) - g C_{i-1} - k_A \frac{\partial C_{i-1}}{\partial t}. \quad (3.8)$$

В уравнении (3.8) учтены процессы захвата примеси в добавляемый элемент и не учтен процесс перераспределения примеси в кремниевой структуре. Поэтому нужно рассчитывать перераспределение примеси, происходящее в процессе роста элемента Z_{i-1} , путем решения уравнения диффузии. Этот этап показан на рис. 3.4, в и завершает вычисления, происходящие за один временной интервал Δt (при этом предполагается отсутствие потока примесных атомов в кремний в данный момент времени). Далее следует приступить к новому циклу расчетов, положив за начальный момент времени $t=t_0+\Delta t$, причем такая процедура вычислений повторяется до тех пор, пока не будет промоделирован весь временной интервал, в ходе которого происходило эпитаксиальное наращивание.

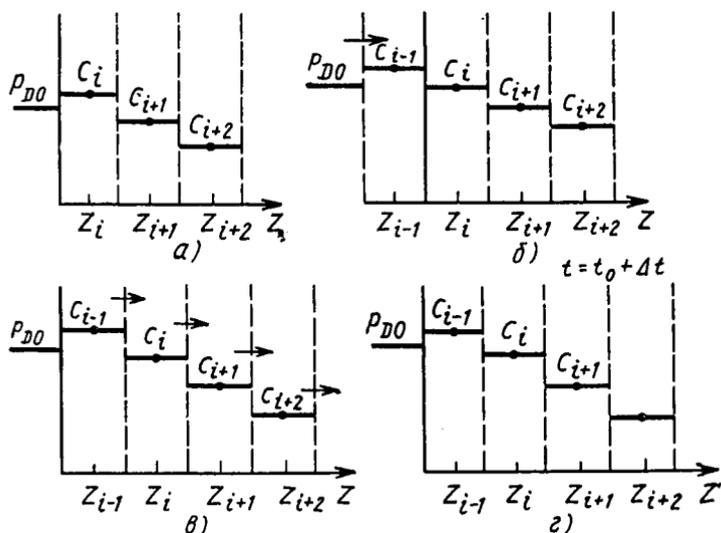


Рис. 3.4. Алгоритм использования численного метода для решения уравнения диффузии с граничными условиями на поверхности в процессе эпитаксиального наращивания

Выбор шага сетки. Важнейшую роль в численной реализации технологического приближения играет правильный выбор шага сетки, на которой моделируется физический процесс. Наиболее очевидным соображением, определяющим этот выбор, является точное представление распределения примесей. Помимо точности физического представления этот выбор влияет и на точность численного решения. В зависимости от погрешности приближения используемого численного метода при уменьшении шага сетки шаг квантования также должен уменьшаться.

Поскольку с увеличением требуемой точности решения резко возрастает число итераций, границы допустимых погрешностей

следует выбирать рационально, не сужая их без необходимости. Разумный выбор соотношения между точностью и временем решения важен потому, что в большинстве распространенных технологий изготовления БИС имеется как минимум один этап, требующий такой точности и максимально допустимого времени решения.

3.3. ДВУМЕРНОЕ МОДЕЛИРОВАНИЕ ТЕХНОЛОГИИ ИЗГОТОВЛЕНИЯ БИС

Структурная схема программы одномерного технологического моделирования. В качестве примера рассмотрим структурную схему и организацию в типовой одномерной технологической программе многостадийного процесса изготовления полупроводниковых ИМС.

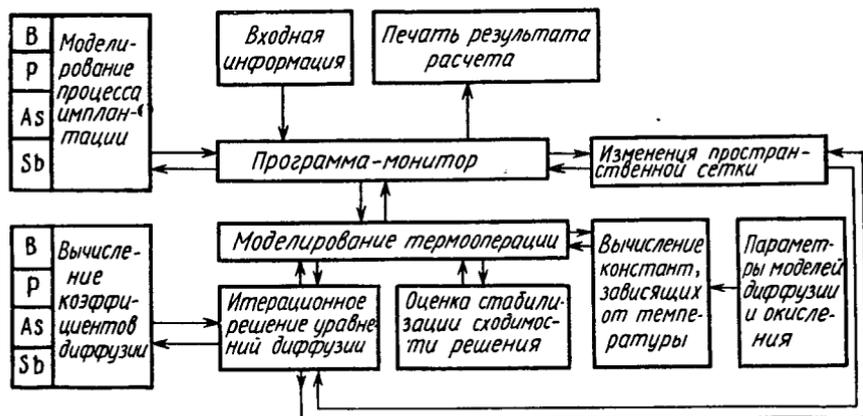


Рис. 3.5. Структурная схема технологической программы

Программа позволяет моделировать:

а) имплантацию бора, мышьяка, сурьмы, фосфора по данным технологического маршрута изготовления БИС (доза, энергия и др.);

б) диффузию этих же примесей в окислительной или инертной атмосфере по данным технологического маршрута — температуре и длительности диффузии, парциальному давлению конкретных окислителей;

в) собственно эпитаксию и размытие примеси скрытого слоя в процессе высокотемпературной эпитаксии;

г) операции травления и низкотемпературного осаждения пленок;

д) произвольную последовательность указанных операций.

Структурная схема одномерной программы технологического моделирования приведена на рис. 3.5. Входными параметрами являются: энергии и дозы ионного легирования; толщина окисла, через которое оно производится; поверхностная концентрация при

моделировании диффузии из бесконечного источника; зависимость температуры диффузии и парциального давления окислителей от времени; уровень легирования подложки или эпитаксиального слоя, скорость роста эпитаксиального слоя и др. Последовательность выполнения технологических операций задается через массивы текстовых констант, по которому программа-монитор вызывает рабочие подпрограммы. Через массивы данных аналогично задается кусочно-линейная зависимость температуры от времени и парциального давления окислителей на каждом временном интервале.

Все теоретические и полуэмпирические коэффициенты для Si и SiO₂ (частично приведенные в гл. 2) собраны в специальные подпрограммы, вызываемые в зависимости от реализуемого типа кристалла. В специальной подпрограмме сосредоточены также некоторые коэффициенты и параметры соответствующих моделей, значения которых целесообразно подстраивать для каждого нового обрабатываемого технологического процесса изготовления ИС.

Двумерное моделирование процесса диффузии в инертной атмосфере. Рассмотрим базовый вариант двумерного моделирования раздельной либо совместной диффузии бора, мышьяка или фосфора в слабоокислительной (инертной) атмосфере. Сначала с помощью программы одномерного моделирования технологии рассчитывают все технологические операции, формирующие профиль примеси в активной области полупроводникового прибора — БТ. Основу двумерного моделирования составляет численно-разностный метод решения системы уравнения, аналогичный (3.1), причем профиль имплантированной примеси в двумерном приближении задавался в расчетах диффузий (2.11).

Область моделирования покрывалась пространственной сеткой с шагами h_x , h_y по координатам x , y и максимальным количеством узлов 70×70 , каждое из уравнений системы (3.1) аппроксимировалось на пятиточечном шаблоне «крест» (рис. 3.6).

Выпишем вид полученных уравнений в узле i, j на разностном шаблоне для двумерной модельной задачи (рис. 3.6):

$$\left[D_{i+1/2,j}^{i+1} \frac{C_{i+1,j}^{i+1} - C_{i,j}^{i+1}}{h_{i+1}} - D_{i+1/2,j} \frac{C_{i,j}^{i+1} - C_{i,j-1}^{i+1}}{h_i} \right] r_j^* +$$

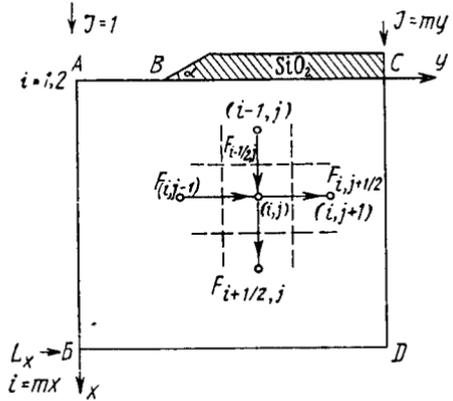


Рис. 3.6. Область двумерного технологического моделирования и разностный шаблон для двумерной модельной задачи

окисного объема определяются из потока атомов кислорода. Движение наружной границы окисла описывается этой скоростью, а также с помощью границ окисла и ряда граничных условий. При высоких температурах ($\geq 960^\circ\text{C}$) окисел ведет себя как вязкая несжимаемая жидкость, описываемая уравнением

$$\nabla \mathbf{v} = 0, \quad (3.13)$$

где \mathbf{v} — скорость элемента окисла.

В предположении отсутствия изменения объема окисла при наличии сжимающих напряжений используется упрощенное уравнение Навье—Стокса

$$\mu \nabla^2 \mathbf{v} = \nabla p, \quad (3.14)$$

где μ , \mathbf{v} , p — вязкость, скорость и давление в области двумерного окисла.

Уравнение (3.14) сводит расчет потока атомов кислорода к более простой задаче объемного расширения на границе раздела фаз.

В упрощенном расчете процесса бокового прокисления можно использовать метод преобразования координат от физической области к координатной системе, в которой движущаяся граница раздела остается неподвижной во времени. Области моделирования для обеих систем координат показаны на рис. 3.7 в предположении, что x , z , t — две пространственные координаты и время в физической системе координат, а ξ , η , τ — соответствующие переменные в преобразованной системе координат. Связь между старыми и новыми координатами записывается таким образом:

$$\xi = z - bf(x, t); \quad \eta = x, \quad \tau = t, \quad (3.15)$$

$$f(x, t) = \frac{d_{\text{ок}}(t)}{2} \operatorname{erfc}(2x/k_l d_{\text{ок}}(t))^{1/2}, \quad (3.16)$$

где f — функция горизонтального положения и времени; k_l — отношение горизонтальной и вертикальной скоростей окисления.

Подстановка соотношений (3.15), (3.16) в уравнение диффузии приводит к преобразованию последнего по отношению к переменным η и ξ , при указанном приеме область решения значительно упрощается.

В обобщенной модели термического окисления сложные граничные условия для C , \mathbf{v} , p учитывают равенство потоков окислителя через все границы раздела, зависимость скорости движения границы раздела от C , упругие свойства изгибаемого в процессе окисления слоя SiO_2 и маскирующего покрытия Si_3N_4 . В результате итерационного решения системы уравнений (3.12) — (3.14) оказывается возможным рассчитать зависимости положения границ раздела SiO_2 — Si_3N_4 , SiO_2 — Si , SiO_2 — газовая фаза, а также распределений скорости, которые, как правило, с удовлетворительной точностью согласуются с экспериментом. При моделировании процесса окисления сильнолегированных подложек необходимо совме-

стно с системой уравнений (3.12)—(3.14) решать соответствующие уравнения диффузии, так как скорость движения границы раздела Si—SiO₂ существенно зависит от концентрации примеси.

Вычислительные аспекты. В зависимости от числа присутствующих примесей в приборной кремниевой структуре нужно решать систему связанных нелинейных дифференциальных уравнений в частных производных в двумерном приближении с соответствующими начальными и граничными условиями. Частные производные в уравнениях аппроксимируются на двумерной сетке методом конечных разностей. Это приводит к замене уравнений в частных производных набором из N обыкновенных нелинейных дифференциальных уравнений, где N — число узлов в сетке разбиения. Для интегрирования уравнений по времени применяется множество различных методов. Для численного решения диффузионных задач приемлемы как явный, так и неявный методы интегрирования. В отношении числа операций, необходимых для расчета каждого шага по времени, неявные методы требуют больших временных затрат по сравнению с явными. Однако неявные методы позволяют использовать большие временные интервалы и обычно являются более стабильными, чем явные методы.

Двумерное моделирование диффузии в слабоокислительной (инертной) атмосфере — весьма трудоемкий процесс с точки зрения машинных ресурсов. Полученная система линейных уравнений (3.9) большой размерности с пятидиагональной матрицей (3.11) может быть решена рядом методов. Для ЭВМ с большим объемом памяти можно рекомендовать запоминание массивов $B_{i,j}$, $E_{i,j}$, $P_{i,j}$, $F_{i,j}$ и использование метода неполного разложения (МНР). Если объем памяти ЭВМ не позволяет выделить дополнительные массивы, то можно воспользоваться строчным методом Зейделя. Он заключается в последовательном применении метода одномерной прогонки к такой системе уравнений для строки j :

$$\begin{aligned} E_{i,j}C_{i-1,j}^{k+1} + B_{i,j}C_{i,j}^{k+1} + Q_{i,j}C_{i+1,j}^{k+1} = \\ = F_{i,j} - P_{i,j}C_{j-1,i}^{k+1} - G_{i,j}C_{j+1,i}^k, \quad i = 2, 3, \dots, mx - 1. \end{aligned} \quad (3.17)$$

Данный метод требует большего числа итераций, чем МНР, но может быть реализован без использования дополнительных массивов.

Существует оригинальный метод построения динамической сетки пространственного разбиения. Он основан на таких критериях: 1) в одномерном приближении последовательные шаги сетки по x и y относятся как $h_{i-1}/h_i = h_i/h_{i+1}$, что позволяет сохранить второй порядок аппроксимаций по h ; 2) в двумерном приближении произведение $(x_{i+1} - x_{i-1})(y_{i+1} - y_{i-1})$ обратно пропорционально второй производной функции распределения в узле (ij) , т. е. чем ближе к линейному виду распределения, тем реже сетка в этой области; 3) в обоих случаях отношение концентраций в соседних узлах не должно превышать заданного числа. Метод позволяет в 10 раз со-

кратить количество узлов по сравнению со стационарной равномерной сеткой. Это сокращение перекрывает вычислительные затраты, связанные с динамическим перестроением сетки и интерполяцией решения в новые узлы. Применение данного метода особенно эффективно в случае моделирования совместной диффузии двух примесей, максимумы распределения которых первоначально сильно разнесены.

В связи с увеличением объема оперативной памяти и быстродействия современных ЭВМ для решения двумерных уравнений диффузии в окислительной атмосфере обычно применяют метод Ньютона, позволяющий быстро получить очень точное решение при достаточно близком начальном приближении — условии, выполняемое в задаче диффузии. Получающиеся при этом системы линейных алгебраических уравнений (СЛАУ) решаются простейшими методами верхней релаксации, Якоби, Гаусса — Зейделя или прямым методом Гаусса.

При дискретизации во временной области вместо простейшего неявного метода Эйлера, используемого в технологических программах первого поколения, применяются сложные формулы дифференцирования «назад» шестого порядка и другие методы, позволяющие оптимизировать последовательность шагов интегрирования по времени.

3.4. МОДЕЛИРОВАНИЕ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ ПОЛИКРЕМНИЯ

В последнее время в интегральной технологии приобрели большое значение и актуальность слои поликристаллического кремния (поликремния — Si*). В МОП-технологии поликремниевые структуры используют главным образом для создания затворов и межсхемных соединений. В биполярной технологии поликремниевые слои применяют в качестве контактов к сильнолегированным областям эмиттера и базы БТ. Поликремниевые слои, находящиеся в непосредственном контакте с кремниевой подложкой, служат в качестве источников диффузии и скрытых контактирующих областей. Благодаря высокому удельному сопротивлению низколегированного поликремния из него выполняют высокоомные резисторы.

Использование слоев поликремния в технологии изготовления современных и перспективных БИС требует модификации моделей отдельных технологических операций для расчета профиля примеси в полупроводниковых приборах. При обработке поликремниевых структур применяют стандартные процессы ионного легирования, диффузии и окисления, однако этим структурам присущи многослойность и, следовательно, наличие нескольких границ раздела, свойства которых могут существенно влиять на характеристики прибора. Для создания технологических программ, способных рассчитывать поликремниевые структуры, в них необходимо ввести характеристики осаждения, окисления, диффузии, сегрегации при-

месей и результирующего сопротивления слоев поликремния. Поликристаллическая структура, содержащая зерна Si^* , требует также создания моделей роста зерна, сегрегации примеси и захвата носителей на границах зерна. Кроме того, некоторые электрофизические параметры слоя Si^* зависят от размера зерен в нем, которые в свою очередь зависят от условий осаждения слоя Si^* и последующей термообработки. Следовательно, для качественного моделирования поликремниевых технологий нужно учитывать взаимосвязь каждого процесса с другими.

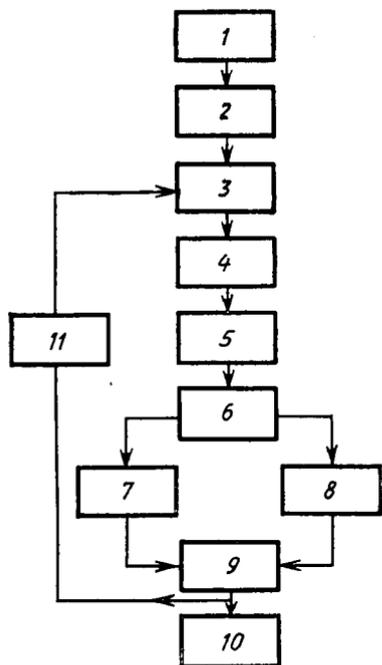


Рис. 3.8. Структурная схема, поясняющая процедуру, используемую в технологической программе расчета электрофизических параметров поликремния:

1 — осаждение; 2 — первоначальные характеристики параметров; 3 — рост зерна; 4 — размер зерна; 5 — сегрегация; 6 — распределение примеси по зерну, активная концентрация примеси; 7 — окисление; 8 — диффузия; 9 — толщина примеси, перераспределенная на границах кремний — поликремний, поликремний — двуокись кремния; 10 — расчет поверхностного сопротивления; 11 — приращения времени

Распределение внедренных ионов мышьяка и бора в слое Si^* практически не отличается от соответствующих распределений в Si . При окислении Si^* возникают те же эффекты, что и при окислении Si : константы роста окисла зависят от температуры, давления окислителей, концентрации примесей и их ориентации. Однако детально необходимо рассмотреть процесс роста зерна и процесс диффузии в системе $Si-Si^*$.

Структурная схема расчета Si^* -слоев (рис. 3.8). Рассмотрим структурную схему расчета электрофизических параметров поликремниевых слоев при моделировании технологии изготовления БИС. Технология первоначального осаждения, температуры, толщина и легирующая примесь определяют первоначальный размер зерна, распределение (сегрегацию) примеси и другие характеристики слоя. Затем на каждом временном шаге Δt процесса в модели роста зерна используют текущие параметры поликремния вместе с температурой, параметрами

окружающей среды, дополнительными легирующими веществами, чтобы определить новый размер зерен. Новый размер зерен применяют для коррекции распределения примеси в зернах и сегрегации на границах зерен, чтобы затем найти новую электрически активную концентрацию примеси внутри зерен. Эти скорректированные значения оказывают затем влияние на окисление поликремния, происходящее за тот же временной шаг ΔS . Рассчитывают также

диффузию примеси через границы раздела поликремния внутри многослойной структуры. За следующий временной шаг повторяют операцию, чтобы найти новые размеры зерна. После завершения всего цикла процесса определяют окончательные размеры зерна, распределение примеси и толщину слоев с захватом носителя заряда на границах зерен, затем по этим данным рассчитывают удельное сопротивление поликремния.

Модели роста зерна и сегрегации при термообработке. Поликремний состоит из множества кристаллитов или зерен монокристаллического кремния, причем размеры и электрофизические характеристики зерен изменяются в зависимости от условий технологического процесса. Согласно рассматриваемой модели, причиной роста зерна является «движущая сила» в процессе управляемой диффузии. Граница между кристаллитами может перемещаться в результате ряда диффузионных скачков атомов кремния через границу. В неравновесном состоянии «движущая сила», вызывающая суммарную диффузию, может существовать благодаря разности свободных энергий атомов, расположенных на двух сторонах границы (за счет ориентации зерен). Кроме того, в самой границе зерен содержится определенная энергия границы раздела, следовательно, «движущая сила» имеет отношение к кривизне границ и размерам зерен.

«Движущая сила» F нарушает равновесие при миграции атомов Si через границу зерна:

$$F = a\lambda b^2/L, \quad (3.18)$$

где a — геометрический фактор; λ — граничная энергия зерна; b — наименьшее расстояние между атомами Si; L — размер зерна.

Скорость движения границы зерна зависит также от подвижности атомов в границе зерна, определяемой выражением $\mu = D_G/\varphi_T$, где D_G — коэффициент самодиффузии кремния через границу зерна, значение которого рассчитывают по формуле

$$D_G = D^i \gamma \exp\left(\frac{\Delta G}{kT}\right) / \left(1 + A \frac{C_G}{C_{Si}} \exp\left(\frac{G_a}{kT}\right)\right). \quad (3.19)$$

Здесь $D^i = 1.2 \exp(-4.37/kT)$ мкм²/мин — коэффициент самодиффузии в зерне, зависящий от нормализованной концентрации заряженных вакансий; γ — коэффициент пропорциональности; $\Delta G = 0.45$ эВ — энергия активации в зерне; A — энтропийный коэффициент ($A = 3.02$ при концентрации $2 \cdot 10^{19}$ см⁻³ и $A = 2.17$ при концентрации $2 \cdot 10^{20}$ см⁻³); C_G — полная концентрация примеси внутри зерна; $G_a = 3.7$ эВ — разность энергий активации в границе и зерне.

Скорость распространения границы зерна можно выразить соотношением

$$\frac{dL}{dt} = \mu F = \frac{a\lambda b^2 D_G}{L_g kT}. \quad (3.20)$$

Рассматриваемая модель также учитывает изменение скорости роста зерна в зависимости от толщины нанесенного слоя Si из-за изменения величины

$$\lambda = \lambda_0 / (1 + L_g/d_s), \quad (3.21)$$

где λ_0 — поверхностная энергия зерна; d_s — толщина слоя Si*.

Формула (3.21) справедлива только в предположении, что все зерна имеют одинаковый размер и сферическую форму. В противном случае необходимо вводить поправочные коэффициенты.

Таким образом, результирующее дифференциальное изменение размера зерен, используемых в технологической программе, согласно выражениям (3.18) — (3.21), на i -м временном шаге записывается в виде

$$\Delta L'_g = \frac{ab^2}{kT} \left(\frac{D_G^i \lambda^i}{L_g^i} \right) \Delta t, \quad (3.22)$$

где D_G^i , λ^i , L_g^i — дифференциальные параметры D_G (3.19), λ (3.21) и L на i -м временном шаге.

Интегрирование (3.22) по времени t всего процесса дает уравнение

$$L_g = \left[L_{g0}^2 + \left(\frac{2ab^2}{kT} \right) D_G \lambda t \right]^{1/2}, \quad (3.23)$$

где L_{g0} — размер зерна перед рассматриваемым шагом процесса.

Если значение L_{g0} относительно мало, то рост зерна имеет параболическую зависимость от времени, что хорошо согласуется с большинством экспериментальных исследований. Поскольку коэффициент D_G (3.19) учитывает эффекты сильного легирования, соотношение (3.23) позволяет обосновать найденную экспериментальную зависимость усиления роста зерна Si* при высоких уровнях легирования полупроводника.

Необходимо еще раз подчеркнуть, что границы зерна являются фактором, влияющим на свойства поликремниевых слоев. Экспериментально обнаружено, что примеси типа фосфора и мышьяка (но не бора) перераспределяются на границах зерен, у которых они электрически неактивны. Рассматриваемый процесс сегрегации является функцией времени отжига, температуры, размеров зерен и концентрации примеси. Установившаяся сегрегация примеси достигается при достаточно долгом отжиге при низких температурах, когда большая часть примесных атомов покоится у границ зерен. Это установившееся состояние моделируется в предположении, что оставшиеся примесные атомы внутри зерна распределены равномерно:

$$\ln \left(\frac{C_{GB}}{C_G} \right) = \ln \left(\frac{AQ_s}{C_{Si}} \right) + \frac{G_a}{kT}, \quad (3.24)$$

где C_{GB} , C_G — концентрации примеси у границы и внутри зерна соответственно; $Q_s = 2.64 \cdot 10^{15} \text{ см}^{-2}$ — эффективная плотность центров сегрегации в 1 см^3 ; $C_{Si} = 5.5 \cdot 10^{22} \text{ см}^{-3}$ — концентрация атомов кремния. Коэффициенты A , G_a рассмотрены выше.

Эффект сегрегации примесей к границам зерна описывается коэффициентом сегрегации $m_s = C_{GB}/C_T$, где $C_T = C_{GB} + C_G$ — полная концентрация. Для бора $m_s = 0$, а для фосфора и мышьяка на основании соотношения (3.24)

$$m_s = 1 - \left[\frac{AQ_s}{L_g C_{S1}} \exp\left(\frac{G_a}{kT} + 1\right) \right]^{-1}. \quad (3.25)$$

Формула (3.25) позволяет рассчитать значения равновесного коэффициента сегрегации, который устанавливается с постоянной времени $\tau_{ms} \approx 0.027 \exp(0.86/kT)$ мин. На рис. 3.9, а представлена зависимость $m_s(L_g)$ для температуры отжига $T = 1000^\circ\text{C}$, а на рис. 3.9, б — расчетные зависимости $L_g(C_T)$ для мышьяка при $d_s = 0.25$ мкм и $t = 1$ ч. Наилучшее совпадение теоретических расчетов с экспериментальными данными, обозначенными квадратами, получено при соотношении $a\lambda_0 b^2 \gamma = 0.24$ эВ.

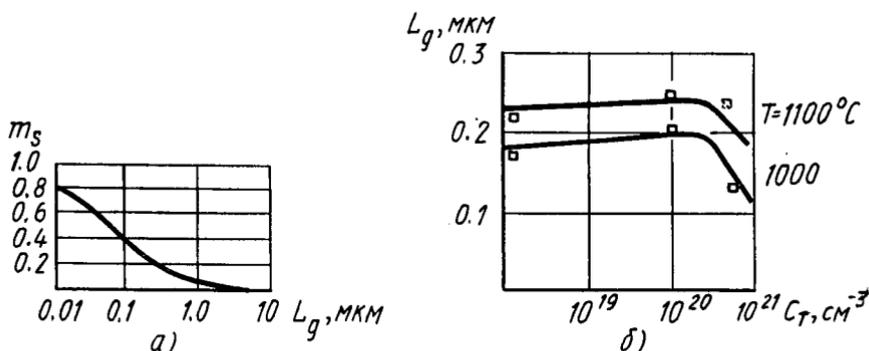


Рис. 3.9. Зависимости коэффициента сегрегации от размеров зерна поликремния (а) и размеров зерна поликремния от концентрации (б)

Термическое окисление поликремния. Окисление Si^* , являющееся важнейшим технологическим процессом, в общем аналогично рассмотренному окислению Si , но имеет следующие принципиальные в исследованиях особенности:

1) коэффициент B/A , который в Si зависит от ориентации, для Si^* имеет разброс примерно $\pm 30\%$, соответствующий хаотическому распределению ориентации зерен Si в пленке Si^* ;

2) возрастание коэффициентов B/A и B при окислении сильно легированного Si^* несколько больше, чем для Si с той же концентрацией примеси;

3) тонкие легированные пленки Si^* окисляются быстрее толстых.

Модель диффузии в системе $\text{Si}-\text{Si}^*$. При моделировании диффузии в системе $\text{Si}-\text{Si}^*$ необходим совместный расчет уравнений диффузии в Si , Si^* и уравнений модели, определяющей поведение $L_g(t)$. Даже в одномерном приближении эта задача очень сложна из-за сложности диффузионных процессов в Si . Многочисленные

экспериментальные исследования диффузии в Si показали, что скорость диффузии вдоль границ зерен в 30—100 раз больше скорости диффузии в зернах, что объясняется наличием большого числа дефектов в аморфном материале границы зерна. Это приводит к тому, что часть внедренной в Si примеси очень быстро достигает границы Si—Si* и начинает легировать кремниевую подложку. На этом факте основана простейшая модель, в которой предполагается, что концентрация примеси в слое Si* мгновенно устанавливается постоянной и затем изменяется со временем в соответствии с потоком примеси через границу Si—Si*. Это предположение справедливо, если время, за которое примесь вдоль границ зерен достигает границы Si—Si*, много меньше времени всего диффузионного процесса. Оценки показывают, что данное условие может не выполняться при формировании мелких эмиттерных профилей диффузией имплантированной примеси из слоя Si*. Например, вдоль

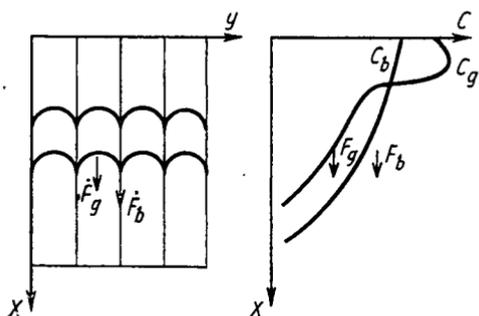


Рис. 3.10. Схематическая модель диффузии примеси в поликремнии F_g, F_b — потоки примеси, C_g, C_b — концентрации примеси в зерне и у границы зерна)

границы зерна Si* при $T=1000^\circ\text{C}$ значение коэффициента диффузии мышьяка $D_b=0,02$ мкм/мин. Для того чтобы в пленке Si толщиной $W=0,14$ мкм установилась равномерная концентрация, требуется время $t^*=(2\div 3)W^2/D_b=2\div 3$ мин. В действительности это значение t^* будет существенно больше, так как, во-первых, нужно некоторое время для протекания примеси из зерен в границу в сильнолегированной приповерхностной области пленки Si* (т. е. для установления равновесного коэффициента сегрегации), а во-вторых, вблизи границы Si—Si* примесь будет перетекать из границ зерен в слаболегированные зерна и концентрация примеси на границе Si—Si* будет возрастать не столь быстро. Время установления равновесного коэффициента сегрегации в Si* равно 60 мин при $T=1000^\circ\text{C}$, что является, по-видимому, грубой оценкой сверху.

В связи с этим большой интерес представляет более «физичная» модель диффузии в Si, учитывающая отдельно диффузию в зернах и вдоль границ зерен. В ней предполагается, что пленка Si состоит из одинаковых «колонн» (рис. 3.10), диффузия идет вдоль этих колонн, коэффициент диффузии в зерне D_g равен соответствующему коэффициенту в Si, а в границе зерна D_b не зависит от концентрации. Уравнения модели имеют вид

$$\frac{\partial C_g}{\partial t} = \frac{\partial}{\partial x} \left(D_g \frac{\partial C_g}{\partial x} \right) - h' (mC_{gs} - C_b), \quad (3.26)$$

$$\frac{\partial C_b}{\partial t} = \frac{\partial^2 C_b}{\partial x^2} D_b + h' (mC_{gs} - C_b),$$

где h' — коэффициент поверхностного массопереноса, равный нулю, если C_{gs} превышает предел растворимости; m — коэффициент сегрегации; C_{gs} — концентрация в зерне вблизи границы, для которой предполагается соотношение $C_{gs} = C_g L_g / \min\{L_g, \sqrt{4D_g t}\}$.

При диффузии под окисляющейся поверхностью Si^* наблюдается увеличение коэффициента диффузии аналогично рассмотренному случаю случайного окисления Si . Отличие состоит в том, что выражение для D^I умножается на $e^{-W/L}$, где W — толщина слоя Si^* , а характерная длина $L = 0.1$ мкм.

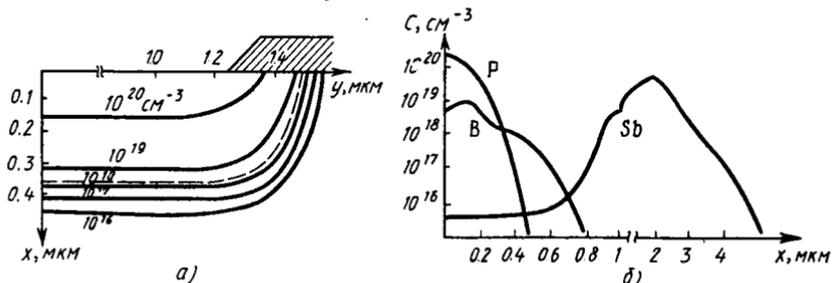


Рис. 3.11. Профили распределения примесей и эквипотенциальные линии эмиттерной примеси интегрального БТ (пунктиром отмечено залегание p - n -перехода)

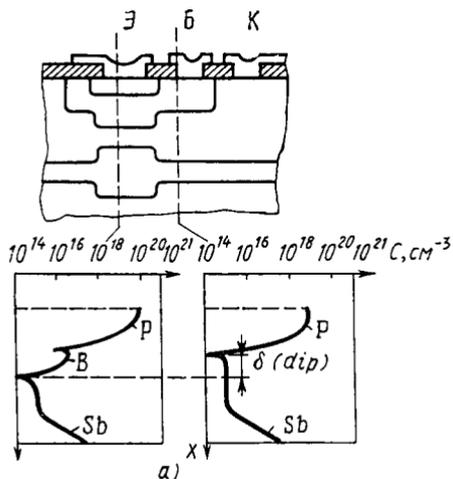
В заключение обсуждения диффузии в системе $Si-Si^*$ отметим, что очень тонкий промежуточный квазиокисный слой, который иногда специально выращивают перед осаждением Si^* на Si , не представляет собой барьера для диффузии. Это объясняется тем, что получаемый слой толщиной 10—20 Å имеет состав SiO_x , где $x = 1.7 \div 1.9$, т. е. сильнодефектную структуру. Более того, промежуточный слой препятствует упорядочению структуры Si^* в процессе отжига, что способствует более интенсивной диффузии в слое Si^* и, следовательно, большей степени легирования подложки.

3.5. ПРИМЕНЕНИЕ ПРОГРАММ ТЕХНОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ

По мере уменьшения горизонтальных и вертикальных размеров полупроводниковых приборов, повышения точности и эффективности их расчета с помощью программ технологического моделирования становится экономически оправданным предварительное моделирование отдельных этапов изготовления полупроводниковых приборов с целью расчета характеристик и параметров, экспериментальное определение которых сопряжено с большими затратами времени и средств. Примерами могут служить расчеты профилей распределения примесей в различных сечениях прибора, поверх-

ностных сопротивлений отдельных слоев, глубин залегания переходов, коэффициентов чувствительности параметров прибора к режимам отдельных операций, определение зависимости плотности дефектов от толщин SiO_2 , Si_3N_4 и др.

На рис. 3.11, а, б приведены рассчитанные концентрационные зависимости в одномерном и двумерном видах профилей БТ БИС,



типовой технологической маршрут изготовления которых включал следующие основные операции: 1) ионное легирование сурьмы в подложку; 2) разгонку сурьмы с окислением; 3) эпитаксию; 4) окисление; 5) ионное легирование бора; 6) разгонку бора; 7) ионное легирование фосфора; 8) разгонку фосфора. Отметим, что программа двумерного технологического моделирования позволяет просто осуществить двумерный расчет эмиттерно-базовой примеси, поскольку эмиттерная диффузия в современных процессах изготовления биполярных БИС проводится в инертной или слабоокислительной атмосфере (сухом кислороде). Сравнение эквипотенциальных линий эмиттерных примесей с учетом и без учета ($D_i = \text{const}$) рассмотренных концентрационных зависимостей коэффициентов диффузии дает относительную погрешность проникновения примесей вдоль координат x, y до 30—60%.

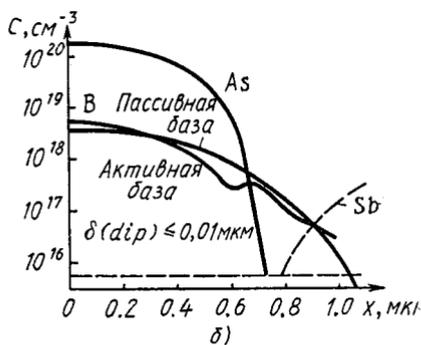


Рис. 3.12. Профили распределения примеси в активном и пассивном сечениях БТ: а — БТ с фосфорным эмиттером; б — БТ с мышьяковистым эмиттером

области боковой части перехода эмиттера концентрация боковой примеси значительно превосходит концентрацию донной его части, а следовательно, увеличивается удельная емкость C_j^b по сравнению с C_j^d , $C_j^b = N_j C_j^d$, $N_j > 2$. По этой же причине пробивное напряжение боковой части $p-n$ -перехода меньше пробивного напряжения донной части. Из рис. 3.11 видно, что глубина залегания эмиттерных $p-n$ -переходов по оси $x-x_j$ сравнима с глубиной залегания перехода по оси $y-y_j$, причем отношение $[y_j(t) - y_j(0)] / [x_j(t) -$

— $x_j(0)$] не постоянно и зависит от условий ионного легирования и диффузии. Это делает необходимым учет криволинейности перехода для прогнозирования БТ с микротопологией. Тенденция повышения степени интеграции БИС приводит к относительному росту отношения емкости боковой части к донной части эмиттерного перехода $\frac{C_3^6}{C_3^A} \approx \frac{h+l}{hl} H_j$ при уменьшении ширины h и длины

l эмиттера, а также увеличении фактора H_j . Из рис. 3.11 видно, что площадь боковой части p - n -перехода S_6 составляет значительную часть площади эмиттера S , даже для $x_{je}=0.34$ мкм; для контактного окна под эмиттер размером $2,5 \times 6$ мкм с учетом топологических особенностей мезаобласти, двумерности распределения примесей отношение $S_6/S_9 = 0.31$. Тенденция уменьшения горизонтальных размеров приводит к росту отношения периметра перехода к его площади и соответственно к относительному возрастанию составляющих поверхностной рекомбинации в общем токе базы, что вызывает некоторое уменьшение усиления и быстродействия БТ БИС.

На рис. 3.12, *а, б* изображены поперечное сечение типового БТ и одномерные профили примеси в областях эмиттера и базы при легировании эмиттера фосфором и мышьяком. При легировании эмиттера фосфором глубина залегания перехода коллектор — база различна в областях активного и пассивного БТ. Это указывает на то, что на диффузию бора существенно влияет сильное легирование эмиттера фосфором; в результате происходит подробно рассмотренное в гл. 2 эмиттерное выталкивание. Сравнение профилей примеси в активном и пассивном сечениях БТ с фосфорным и мышьяковистым эмиттером позволяет оценить «дип-эффект», вызванный ускоренным движением бора при его совместной диффузии с фосфором (мышьяком). Этот эффект значителен для БТ с фосфорным эмиттером (рис. 3.12, *а*) и пренебрежимо мал для БТ с мышьяковистым эмиттером (рис. 3.12, *б*). Отметим, что для последних расчеты показывают высокий градиент распределения примеси в эмиттерном переходе ($\sim 10^{24}$ см $^{-4}$), на порядок больший градиента аналогичных БТ с фосфорным эмиттером.

3.6. МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ОБРАЗОВАНИЯ ПОВЕРХНОСТНЫХ КОНФИГУРАЦИЙ

Как известно, фундаментальные ограничения на функциональные параметры и степень интеграции современных и перспективных БИС определяются на приборно-технологическом уровне. По мере того как прогресс микротехнологии БИС позволяет получать более тонкие линии с минимальными допусками, топографические особенности создаваемых полупроводниковых приборов приобретают все большее значение. Для высокой разрешающей способности не только нужен лучший контроль ширины линий, но и сам контроль

должен проводиться в условиях, когда размеры в третьем измерении соизмеримы с шириной линии. Топографические особенности, профили слоев полупроводниковых структур, получаемые с помощью микрофотографий поперечных шлифов кристаллов, дают возможность установить «обратную связь» с конкретным технологическим процессом. Однако получение таких микрофотографий с помощью растрового электронного микроскопа (РЭМ) сопряжено с большими затратами времени и средств. Кроме того, для систематического исследования экспериментальных топографических параметров требуется чрезвычайно точный контроль технологических процессов образования поверхностных конфигураций. Описание методов, позволяющих охарактеризовать и учесть при проектировании поверхностные непланарные особенности полупроводниковых приборов, является главной целью моделирования технологических процессов изготовления БИС, в первую очередь процессов образования поверхностных конфигураций подобных структур. Помимо таких достоинств, как управляемость и наблюдаемость, данное моделирование можно проводить быстрее и экономичнее, чем экспериментальное получение поперечных шлифов с соответствующими РЭМ — микрофотографиями структур.

Моделирование процессов образования поверхностных конфигураций — литографии, осаждения и травления пленок — распадается на две отдельные задачи. Во-первых, необходимо моделировать кинетику физических процессов, происходящих на границе обрабатываемого материала, с тем чтобы рассчитать зависимость скорости движения указанной границы от условий конкретного процесса. Например, при моделировании травления резиста сначала аналитически или методом Монте-Карло рассчитывают распределение энергии, поглощенной резистом. При этом учитывают рассеяние электронов не только молекулами резиста, но и границей резист — подложка, атомами подложки, уход электронов из резиста и т. д. Далее на больших пластинах экспериментально определяют зависимость скорости травления резиста от количества поглощенной энергии. Таким образом, становится известной скорость травления в каждой точке экспонированного резиста. При моделировании различных видов травления и осаждения эта скорость определяется, как правило, экспериментально. Она зависит от типа материала, энергии частиц, интенсивности и угла падения пучка. Большой интерес представляют работы, где эти данные получены расчетным путем, из «первых принципов».

Во-вторых, нужно решить геометрическую задачу о движении границы обрабатываемого материала с известной зависимостью скорости перемещения от условий процесса и положения рассматриваемого участка материала по отношению к соседним участкам и, возможно, к источнику воздействия (например, при реактивном травлении или напылении).

Моделирование процесса оптической литографии. Моделирование процесса оптической литографии стало возможным благода-

ря созданию модели экспонирования — проявления — травления для позитивного фоторезиста. Это позволяет моделировать процессы проекционной печати и полезно для оценки современных средств оптической литографии и резистивных материалов.

В оптической литографии используются фоторезисты, покрывающие тонкой пленкой поверхность кристалла, на которую должно быть нанесено изображение. Фоторезистивная пленка экспонируется в синем или ультрафиолетовом свете, после чего в соответствии с рисунком фотошаблона в ней образуются засвеченные и незасвеченные области. При проявлении происходит селективное удаление фоторезиста в соответствии с полученной экспозицией. При проявлении позитивных фоторезистов удаляются освещенные области, негативных — неосвещенные. Рассмотрим модель литографического процесса при использовании позитивного фоторезиста, состоящего из смолы, ингибитора (фотоактивного соединения) и легко испаряющегося растворителя, позволяющего наносить фоторезист на пластину в жидком виде.

Экспонирование светом вызывает химические видоизменения в фоторезисте, что приводит к разрушению ингибитора. Проявление заключается в процессе травления, при котором происходит селективное удаление резиста со скоростью, определяемой количеством разрушенного ингибитора.

Сильное поглощение, связанное при экспонировании с фоточувствительностью резиста, приводит к значительному оптическому поглощению на длине волны экспонирования. По мере разрушения ингибитора, концентрация которого составляет около 30% концентрации фоторезиста, поглощение уменьшается. Для описания процесса разрушения ингибитора введем понятие относительной стойкости $M(Z, t)$, характеризующей концентрацию ингибитора для произвольной координаты Z и времени экспонирования t по отношению к его концентрации перед экспонированием. Большинство фоторезистов характеризуется слабым рассеянием света, поэтому постоянная поглощения α определяется соотношением

$$\alpha = AM(Z, t) + B, \quad (3.27)$$

где A, B — экспериментально определяемые параметры материала, характеризующие соответственно зависящую и не зависящую от экспозиции части постоянной поглощения.

Скорость разрушения ингибитора связывается дифференциальным уравнением в частных производных с интенсивностью света в данной точке $I(Z, t)$, концентрацией ингибитора M и чувствительностью Φ :

$$\frac{\partial M}{\partial t} = -I(Z, t)M(Z, t)\Phi. \quad (3.28)$$

Изменение интенсивности света в бесконечно толстом слое фоторезиста определяется дифференциальным уравнением

$$\frac{\partial I(Z, t)}{\partial t} = -I(Z, t)[AM(Z, t) + B].$$

Не только параметры A , B , но и чувствительность Φ определяются экспериментально и зависят от типа фоторезиста и длины световой волны.

Уравнения (3.27), (3.28) решают методами численного интегрирования относительно $M(Z, t)$ и $I(Z, t)$ при заданных, определяемых экспериментально параметрах A, B, C, I_0 . Например, для типового фоторезиста $A=0.86$ мкм⁻², $B=0.07$ мкм⁻¹, $\Phi=0.018$ см²/мДж на длине волны 404.7 нм.

Поскольку оптические свойства фоторезиста в ходе экспонирования непрерывно изменяются и являются функцией глубины,

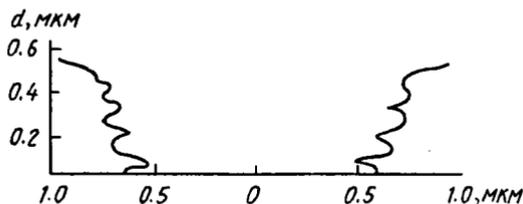


Рис. 3.13. Профиль линии позитивного фоторезиста толщиной 1 мкм после обработки

пленку фоторезиста при моделировании разбивают на тонкие слои, внутри которых свойства фоторезиста можно считать однородными и изотропными. Расчет проводят в достаточно малые дискретные интервалы времени (экспонирование). Обозначив I_j и M_j интенсивность светового потока и

концентрацию ингибитора в j -м слое, можно рассчитать I_j , поддерживая концентрацию ингибитора постоянной. Процедура расчета заключается в следующем: обозначив приращение времени экспозиции Δt_e и определив новые значения M_j , повторяют расчет I_j . При этом концентрация M_j изменяется в ходе экспонирования согласно выражению

$$M_j|_{t_e+\Delta t_e} = M_j|_{t_e} \exp(-I_j C \Delta t_e) \quad (3.29)$$

при начальном условии $M_j|_{t_e} = 1$. Удовлетворительная точность достигается в тех случаях, когда толщина слоев δZ_j , на которые разбивают пленку фоторезиста, составляет менее 0.03λ , а приращение времени экспонирования выбирают таким образом, чтобы максимальное изменение для любой концентрации составляло не более 0.2.

Описание экспонированной в фоторезисте картины определяется двумерной матрицей концентраций ингибитора $M(x, t)$. Процесс проявления моделируется как реакция травления с поверхностным контролем, интенсивность протекания которой определяется локальными значениями M , неизменными внутри элемента с размерами Δx и ΔZ . Скорость травления определяется выражением

$$R(M) = a \exp[E_1 + E_2 M + E_3 M^3], \quad (3.30)$$

где a — коэффициент, определяющий скорость травления, мкм/с; E_1, E_2, E_3 — экспериментальные константы фоторезиста, зависящие от типа проявителя, температуры и условий проведения процесса.

Удаление отдельных элементов с поверхности контакта прояви-

теля и фоторезиста происходит в соответствии с уравнением (3.30) и зависит от числа элементов, находящихся в контакте с проявителем. После удаления предыдущих элементов начинается травление последующих элементов. Время, необходимое для удаления элемента, у которого экспонирована только верхняя сторона, определяется выражением

$$t_r = \Delta Z / R_{ij},$$

где ΔZ — толщина слоя фоторезиста; R_{ij} — скорость травления отдельного элемента.

Если экспонированы верхняя и одна из боковых сторон, то время для удаления элемента

$$t_r = \frac{\Delta x \Delta Z}{R_{ij} (\Delta x^2 + \Delta y^2)^{1/2}}.$$

На рис. 3.13 приведен расчетный профиль линии фоторезиста с номинальной толщиной 1 мкм.

Моделирование процессов травления и осаждения. Моделирование процессов травления и осаждения по своей сути является геометрической задачей. Граница между обработанной и необработанной областями при моделировании аппроксимируется набором точек, соединенных прямыми отрезками. Результирующий профиль обработанной поверхности определяется положением первоначального профиля, который движется через сред с учетом того, что скорость распространения в каждой точке является функцией локальных переменных.

Все процессы травления могут быть классифицированы на изотропные и анизотропные. Рассмотрим пример, иллюстрирующий приложение модели струны к процессам изотропного и анизотропного травления (рис. 3.14). Моделирование изотропного процесса травления заключается в придании всем точкам струны постоянной скорости перемещения в направлении биссектрисы угла, образуемого между двумя перпендикулярами к соседним прямолинейным отрезкам (рис. 3.14, а). Скорость анизотропного травления пропорциональна косинусу угла между направлением потока ча-

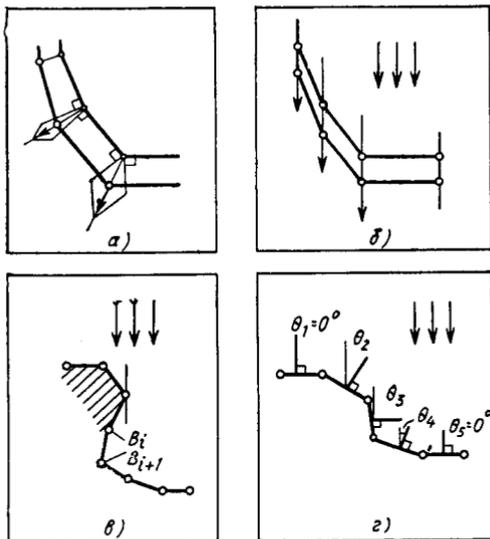


Рис. 3.14. Иллюстрация модели травления:

а — изотропное; б — анизотропное поступательное перемещение точек, принадлежащих профилю края обрабатываемой поверхности; в — затененные точки струны B_i, B_{i+1} , расположенные в теневых участках; г — определение локальных угловых ориентаций в модели ионного распыления

стиц и нормалью к поверхности (рис. 3.14, б). С другой стороны, в крайних ситуациях, возникающих при ионном распылении, может происходить образование теневых участков (рис. 3.14, в). Скорость травления точек, которые затенены другими отрезками, принимается равной скорости изотропного травления. При моделировании процессов ионного распыления для соседних отрезков берут среднее значение угла по отношению к направлению ионного пучка; процедура моделирования представлена на рис. 3.14, г. В программе моделирования процессов образования поверхностных конфигураций БИС скорость травления вдоль этих линий рассчитывают в соответствии с эффективностью распыления

$$S(\theta) = S_0 (\Pi/\rho) (A \cos \theta + B \cos^2 \theta + C \cos^4 \theta),$$

где параметры S_0 , A , B , C характеризуют эффективность распыления материала, подвергаемого ионному травлению; Π — плотность тока в ионном пучке; ρ — плотность атомных слоев.

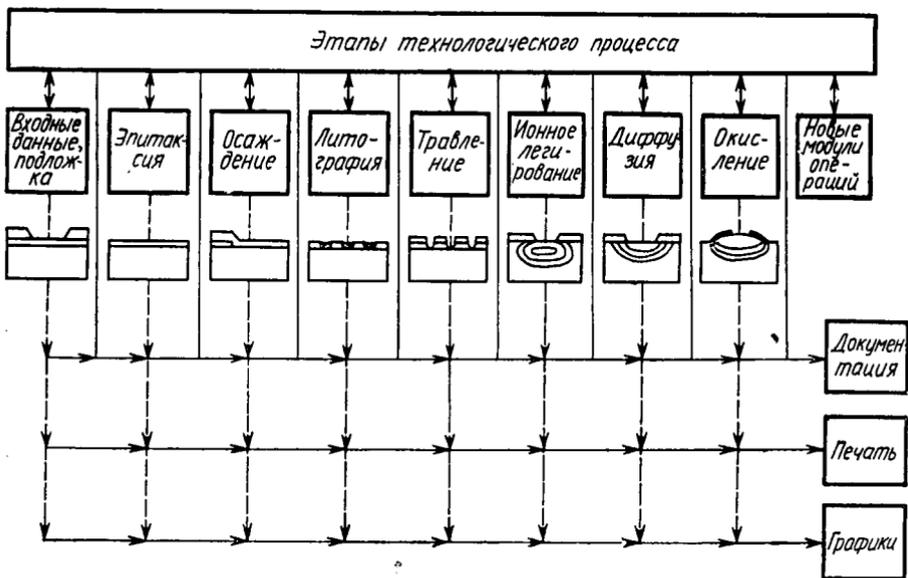


Рис. 3.15. Модульная обобщенная структура программы моделирования полного технологического цикла изготовления кремниевых БИС

Граница обрабатываемого материала представляется в виде ломаной линии, координаты узлов которой хранятся в памяти ЭВМ. В каждый момент времени вычисляются скорости движения границы и происходит сдвиг либо узлов в направлении биссектрис углов между сегментами, либо сегментов в перпендикулярном им направлении. В рамках рассмотренной модели струны рост пленок в ходе осаждения моделируется как травление с отрицательной скоростью. Продвижение каждой точки кусочно-линейного профи-

ля определяется условиями осаждения. В процессе такого продвижения за счет автоматического подбора происходит добавление или отбрасывание точек на отрезках струны, когда они становятся длинными или очень короткими.

В алгоритме ячеек, который используется, например, при моделировании процесса обработки резиста, вся область моделирования разбивается на массив прямоугольных или кубических (в трехмерном случае) ячеек. Каждой ячейке приписывается среднее количество поглощенной энергии и, следовательно, среднее время травления. Это время зависит также от количества граней ячеек, находящихся в контакте с травителем. Массив ячеек, хранящийся в памяти ЭВМ, просматривается в каждый дискретный момент времени моделирования. Различными признаками идентифицируются ячейки, в которых резист уже стравился, находится под обработкой или не обрабатывается. Таким образом, в любой момент времени ЭВМ может построить профиль обработанного резиста. Типичное время моделирования составляет десятки секунд на ЭВМ с производительностью 1—10 млн. опер/с.

Несмотря на то что рассмотренные модели, алгоритмы основаны на использовании большого количества параметров, характерных только для конкретных материалов и режимов изготовления, и не носят фундаментальный характер, они уже сегодня входят составной частью в ряд других универсальных технологических программ моделирования современных и перспективных БИС (рис. 3.15). Подчеркнем, что по мере совершенствования технологии БИС моделирование рассмотренных процессов создания поверхностных конфигураций становится все более важным, причем главное внимание будет уделяться разработке новых кинетических моделей поверхностных и объемных реакций с включением их в модули разрабатываемых универсальных программ технологического изготовления перспективных БИС общего и специализированного назначения.

**ФИЗИКО-ТОПОЛОГИЧЕСКИЕ МОДЕЛИ
ПОЛУПРОВОДНИКОВЫХ СТРУКТУР****4.1. НАЗНАЧЕНИЕ И ПРОБЛЕМЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ
ПОЛУПРОВОДНИКОВЫХ СТРУКТУР**

С помощью численных физико-топологических (ФТ) моделей по их входным параметрам (профиль легирующей примеси, расположение и размеры контактов, зависимости подвижности, времена жизни носителей, сужение запрещенной зоны от концентрации носителей, скорость поверхностной рекомбинации и др.) рассчитывают как «стационарные» электрические характеристики компонентов — ВАХ, емкости, времена задержки, так и «нестационарные» — переходные характеристики токов с соответствующими параметрами. ФТ-моделированию, основанному на численно-разностном решении фундаментальной системы уравнений (ФСУ) полупроводника с соответствующими граничными и начальными условиями в одно-, двух-, трехмерном и квазимономерном приближениях, посвящено начиная с середины 60-х годов большое количество работ. Причем в середине 80-х годов интерес к разработке и использованию ФТ-моделей, численным методам, алгоритмизации и организации процесса численного моделирования субмикронных компонентов БИС резко возрос. Отметим, что в понятие «фундаментальности» уравнений полупроводника большинством авторов не вкладывается их незыблемость — скорее это один из наборов основополагающих уравнений, описывающих с достаточной адекватностью электрофизические процессы в современных и перспективных полупроводниковых структурах. Кроме того, в зависимости от сделанных физических допущений и подходов к решению уравнений ФТ-модели согласно концепциям разработчиков ФТ-программ получается та или иная эффективность численного расчета при ориентации на тип и электрические режимы полупроводниковых структур. В наиболее распространенном до 70-х годов ФТ-методе разделения прибора на ряд областей квазинейтральности и объемного заряда с выделением в них доминирующего физического процесса достаточно просто расщепляется исходная ФСУ (сшиваемая на границах аналитическими приближениями) и далее осуществляется многомерное моделирование. Несмотря на достаточно широкую область применения ФТ-программ, основанных на таком подходе, для компонентов ИМС и БИС невысокого быстродействия, например БТ, работающих в режиме средних

уровней инжекции, они не могут обеспечить приемлемой точности моделирования тонкослойных БТ с микротопологией — элементов С³БИС, работающих с высокими плотностями тока.

Существенными недостатками подобных ФТ-моделей и программ, приводящих к неадекватности численного моделирования, являются: 1) идеализированное распределение примесей с ортогональными *p-n*-переходами; 2) задание средних значений электрофизических параметров в квазинейтральных областях, неучет или приблизительный учет изменения положения границ выделяемых областей при изменении уровня инжекции; 3) некоторые существенные ограничения на микротопологию и тип изоляции; 4) неучет большого количества физических эффектов (сильного легирования, кинетических и др.).

Большой универсальностью характеризуется подход и программное обеспечение, в которых ФСУ полупроводника решается методами конечных разностей или конечных элементов без выделения характерных областей, единообразно для всей полупроводниковой структуры. Наблюдающийся в последнее время значительный прогресс в развитии численных методов такого многомерного подхода, широкое внедрение программного обеспечения численного расчета элементов в промышленные разработки ИМС и БИС являются гарантией успешного развития актуального направления — сквозного автоматизированного проектирования перспективных БИС. Немаловажную роль играют и будут играть в дальнейшем факты установления адекватности ФТ-моделирования субмикронных элементов БИС, необходимости включения в ФТ-модели новых физических эффектов в тонкослойных сильнолегированных компонентах, идентификации основных статических и динамических параметров приборных структур. В то же время можно констатировать, что большинство разработанных ФТ-моделей универсального подхода применимо к задачам расчета, например БТ крайне ограниченного назначения, без учета специфики распределения примесей в мелких субмикронных БТ, в режимах низкого и среднего уровней инжекции и т. д.

Подчеркнем, что в вопросах физического и программного обеспечения физико-топологического моделирования ориентируются прежде всего на разработчиков элементной базы — полупроводниковых приборов, БИС, которым в своей работе следует постоянно учитывать физику явлений, конструктивные особенности, лежащие в основе проектируемых и разрабатываемых полупроводниковых приборов. Современный уровень численного моделирования многомерных полупроводниковых структур при использовании даже самого мощного арсенала математических и алгоритмических методов не позволяет, к сожалению, избежать огромных затрат машинного времени и оперативной памяти. По оценкам специалистов, задачи трехмерного численного моделирования субмикронных тонкослойных сильнолегированных полупроводниковых структур будут сопряжены с вычислительными ресурсами, сравнимыми или

превосходящими ресурсы при решении задач точного прогнозирования погоды или проектирования авиалайнеров. Математическая и физическая сложность задач двух- и трехмерного моделирования такова, что, по крайней мере на сегодня, несмотря на различные ухищрения численных методов, надежное решение данного класса задач возможно лишь достаточно трудоемкими численными методами, требующими быстродействия $\geq 10\text{--}20$ млн. опер/с и объема оперативной памяти $\geq 1\text{--}4$ Мбайт высокопроизводительных ЭВМ. С этой точки зрения даже разработчикам алгоритмического, программного обеспечения для физико-технологического моделирования полупроводниковых структур необходимо соотносить требования точного расчета и выбора альтернативы упрощенного рассмотрения сложных многомерных структур, кинетических уравнений, описывающих физические процессы, граничных условий и топологических упрощений.

Проблема уточнения геометрических и электрофизических параметров моделей, модификации граничных и других условий, физических механизмов, составляющих суть моделей тех или иных процессов в полупроводниковых структурах, не теряет своей огромной актуальности при разработке приборов современных и перспективных БИС, поскольку с уменьшением горизонтальных и вертикальных размеров приборов заметное влияние на их характеристики и параметры начинают оказывать все новые и новые физические эффекты.

Поскольку наибольший теоретический и практический интерес представляет численный расчет элементов БИС в рамках совместного физико-технологического моделирования элементов БИС с выходом на расчет схем, основное внимание будет уделено именно проблемам численного ФТ-моделирования полупроводниковых приборов, в первую очередь кремниевых БТ, исследованию их характеристик, взаимодействию технологических и ФТ-программ с целью идентификации и оптимизации электрических параметров приборов.

4.2. ОСНОВНЫЕ УРАВНЕНИЯ ПОЛУПРОВОДНИКА

ФСУ — Пуассона, непрерывности и плотностей тока для дырок и электронов — записывается следующим образом:

$$\varepsilon\varepsilon_0\Delta V = -q(p - n - C_A + C_D), \quad (4.1)$$

$$\frac{\partial p}{\partial t} + \frac{1}{q} \operatorname{div} \mathbf{J}_p = -R_p, \quad (4.2)$$

$$\frac{\partial n}{\partial t} - \frac{1}{q} \operatorname{div} \mathbf{J}_n = -R_n, \quad (4.3)$$

$$\mathbf{J}_n = qD_n \nabla n - qn\mu_n \mathbf{E}_n = \mu_n n \nabla E_{F_n} = -q\mu_n n \nabla \varphi_n, \quad (4.4)$$

$$\mathbf{J}_p = -qD_p \nabla p + qp\mu_p \mathbf{E}_p = \mu_p p \nabla E_{F_p} = -q\mu_p p \nabla \varphi_p. \quad (4.5)$$

Здесь V , p , n — потенциал, концентрации дырок и электронов; C_A , C_D — концентрации ионизированных акцепторной и донорной примесей; ϵ — относительная диэлектрическая проницаемость кремния; ϵ_0 — абсолютная диэлектрическая проницаемость вакуума, $\epsilon_0 = 8.85 \cdot 10^{-12}$ Ф/М; q — заряд электрона, $q = 1.6 \cdot 10^{-14}$ Кл; J_p , J_n — плотности дырочного и электронного токов; R_p , R_n — суммарные скорости рекомбинации для дырок и электронов; E_p , E_n — напряженности квазиэлектрического поля для дырок и электронов; μ_p , μ_n — коэффициенты подвижности для дырок и электронов; E_{Fp} , E_{Fn} — уровни энергии Ферми относительно уровня вакуума; φ_n , φ_p — квазипотенциалы Ферми для электронов и дырок.

Отличительная особенность записи уравнений для плотности дырочного и электронного токов вытекает из следствия термодинамики необратимых процессов в изотермических условиях, что позволяет связать плотности токов и градиенты уровней энергии Ферми. Для полупроводниковых структур, учитывающих влияние эффектов сильного легирования (ЭСЛ) — уменьшение ширины запрещенной зоны и вырождение полупроводников, — использование обобщенных соотношений $J_n(E_{Fn})$, $J_p(E_{Fp})$ позволяет, во-первых, исключить дополнительные члены, необходимые в ряде случаев для разделения дрейфовой и диффузионной составляющих токов, и, во-вторых, достаточно просто учесть квазиэлектрическое поле: асимметрию сужения запрещенной зоны в области сильного легирования и следствие статистики Ферми—Дирака при проявлении ЭСЛ.

Уравнение Пуассона. Уравнение (4.1), используемое для численного моделирования структур, является удобным для расчетов E , V следствием одного из четырех обобщенных уравнений Максвелла:

$$\operatorname{div} \mathbf{D} = \rho, \quad \operatorname{div}(\epsilon \epsilon_0 \mathbf{E}) = \rho, \quad (4.6)$$

где \mathbf{D} — вектор электрического смещения; ρ — плотность электрического заряда.

Пренебрегая магнитным полем и связывая потенциал $V(x, y, z, t)$ с вектором напряженности электрического поля \mathbf{E} соотношением

$$\mathbf{E} = -\operatorname{grad} V = -\nabla V, \quad (4.7)$$

из (4.6), (4.7) получим

$$\operatorname{div}(\epsilon \epsilon_0 \operatorname{grad} V) = -\rho. \quad (4.8)$$

Обозначая объемную плотность заряда $\rho = q(p - n + C_D - C_A)$, найдем искомое уравнение Пуассона

$$\operatorname{div} \operatorname{grad} V = -\frac{q}{\epsilon \epsilon_0} (p - n + C_D - C_A), \quad (4.9)$$

которое можно записать также в виде (4.1), где значения параметров ϵ для различных материалов: Si, SiO₂, Si₃N₄, GaAs, Ge — равны соответственно 11.7, 3.9, 7.2, 12.5, 16.1.

Уравнения непрерывности. Рассмотрим полупроводниковую структуру с концентрацией C_A, C_D, n, p , заполняющую некоторый объем V и ограниченную замкнутой поверхностью S . Предположим, что в объеме V , где протекают потоки электронов и дырок с плотностями J_n, J_p , происходит рекомбинация частиц со скоростью $R(n, p)$. Число электронов, покидающих произвольный объем V_i , ограниченный поверхностью S_i (в общем объеме V) за единицу времени, равно $\int_{S_i} J_n n d\sigma/q$, где n — единичный вектор внешней нормали к поверхности. Число электронов, исчезающих из объема V_i вследствие рекомбинации за единицу времени, равно $-\int_{V_i} R dv$.

В то же время изменение числа электронов в объеме V_i за единицу времени определяется величиной $\frac{\partial}{\partial t} \int_{V_i} n dV$. Составим уравнение

баланса общего числа электронов за единицу времени в объеме V_i :

$$-\int_{S_i} (J_n n) \frac{d\sigma}{q} - \int_{V_i} R dV = \int_{V_i} \frac{\partial n}{\partial t} dV. \quad (4.10)$$

Аналогичное уравнение может быть выведено и для дырок:

$$-\int_{S_i} (J_p n) \frac{d\sigma}{q} - \int_{V_i} R dV = \int_{V_i} \frac{\partial p}{\partial t} dV. \quad (4.11)$$

Отметим, что уравнения (4.10), (4.11) выражают законы сохранения числа электронов и дырок в объеме V_i , отрицательного и положительного зарядов, а также плотностей электронного и дырочного токов. Используя формулу Остроградского — Гаусса, преобразуем поверхностные интегралы в объемные:

$$\int_{V_i} \left(\operatorname{div} J_n - qR - q \frac{\partial n}{\partial t} \right) dV = 0, \quad (4.12)$$

$$\int_{V_i} \left(\operatorname{div} J_p + qR + q \frac{\partial p}{\partial t} \right) dV = 0. \quad (4.13)$$

Ввиду произвольности выбранного объема V_i из (4.12), (4.13) следуют уравнения непрерывности (4.2), (4.3), выражающие те же законы сохранения, что и уравнения (4.10), (4.11).

Кинетические уравнения переноса носителей заряда. В общем виде векторы плотностей электронного и дырочного токов определяются концентрацией и средней дрейфовой скоростью частиц:

$$J_n = -qnv_n, \quad J_p = qpvn. \quad (4.14)$$

Главной проблемой описания кинетических явлений переноса носителей заряда в полупроводнике является выявление связи

средних скоростей носителей с концентрацией и напряженностью электрического поля.

В качестве базовой «квазиклассической» модели переноса носителей заряда примем модель, основанную на следующих допущениях: 1) свободные носители заряда в полупроводниковой структуре можно рассматривать как точечные частицы в фазовом пространстве координат и моментов. Квантовые эффекты учитываются косвенно в эффективной массе; 2) количество носителей заряда в структуре достаточно велико, поэтому правомочно использование аппарата статистического анализа; 3) носители заряда в структуре можно считать практически не взаимодействующими, т. е. функцию распределения нескольких частиц можно записать как произведение отдельных функций распределения.

Для описания кинетических явлений в полупроводнике, обусловленных движением носителей заряда при наличии внешних и внутренних полей, градиента температур, используют кинетическое уравнение Больцмана. Поскольку полное число состояний в полупроводнике — величина постоянная, полная производная по времени от функций распределения частиц по состояниям $f(x, k, t)$ [в пространстве семи измерений: координат $x(x, y, z)$, моментов $k(k_x, k_y, k_z)$ и времени t] равна нулю: $df/dt=0$. Дифференцируя $f(x, k, t)$ по времени с учетом того, что $df/dt=0$ и x, k являются функциями времени, получим

$$\frac{\partial f_{n,p}}{\partial t} + \text{grad}_k f_{n,p} \frac{dk_{n,p}}{dt} + \text{grad}_x \frac{dx_{n,p}}{dt} = 0. \quad (4.15)$$

Уравнение (4.15) показывает, что изменение во времени функций распределения для электронов и дырок $f_{n,p}$ (в дальнейшем будем рассматривать f_n аналогично f_p) в каждой точке фазового пространства (x, k) вызвано движением частиц в пространстве координат x и моментов k в результате действия внешних F_e и внутренних F_i сил. Причем функция распределения f_n определяется как вероятность согласно формуле расчета концентраций n в полном объеме моментов V_k :

$$\frac{1}{4\pi^3} \int_{V_k} f_n(x, k, t) dk = V(x, t). \quad (4.16)$$

Изменение во времени функции распределения (4.15) представляется в виде суммы двух членов — полевого и столкновений:

$$\frac{\partial f_n}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{пол}} + \left(\frac{\partial f}{\partial t} \right)_{\text{ст}}.$$

Производная по времени вектора k_n связана с суммой внешних и внутренних сил в полупроводнике $F_n = F_e + F_i$ соотношением

$$\frac{dk_n}{dt} = \frac{F_n}{\hbar}, \quad (4.17)$$

где $\hbar = \hbar 2\pi$ — постоянная Планка.

Для нахождения $(\partial f/\partial t)_{ст}$ используют статистические методы описания физических явлений.

Столкновения приводят к переходу частиц из одного состояния в другие с вероятностью $S_n(\mathbf{k}, \mathbf{k}')$. Тогда с помощью члена $S_n(\mathbf{k}, \mathbf{k}')dk'$, означающего вероятность столкновений в объеме моментов dk' , можно записать интеграл члена столкновений:

$$\text{grad}_{\mathbf{k}} f_n \frac{F_i}{\hbar} = \int_{V_{\mathbf{k}}} [(f_n(\mathbf{x}, \mathbf{k}, t)(1 - f_n(\mathbf{x}, \mathbf{k}', t))S_n(\mathbf{k}, \mathbf{k}') - f_n(\mathbf{x}, \mathbf{k}', t)(1 - f_n(\mathbf{x}, \mathbf{k}, t))S_n(\mathbf{k}', \mathbf{k})] dk'. \quad (4.18)$$

Первый член интеграла описывает уменьшение количества частиц в элементе объема dk' в результате прямых переходов из состояний \mathbf{k} в состояние \mathbf{k}' . Второе слагаемое определяет увеличение количества частиц в dk' в результате обратных переходов из состояния \mathbf{k}' в \mathbf{k} с вероятностью $S_n(\mathbf{k}', \mathbf{k})$.

Производная по времени вектора \mathbf{x}_n представляет групповую скорость носителей заряда:

$$\frac{d\mathbf{x}_n}{dt} = \mathbf{v}_n. \quad (4.19)$$

Подставляя соотношения (4.17)–(4.19) в (4.15), получаем так называемое обобщенное кинетическое уравнение Больцмана (КУБ):

$$\begin{aligned} \frac{\partial f_n}{\partial t} &= \frac{F_e}{\hbar} \text{grad}_{\mathbf{k}} f_n + \mathbf{v}_n \text{grad}_{\mathbf{x}} f_n = \\ &= - \int_{V_{\mathbf{k}}} [f_n(\mathbf{x}, \mathbf{k}, t)(1 - f_n(\mathbf{x}, \mathbf{k}', t))S_n(\mathbf{k}, \mathbf{k}') - \\ &\quad - f_n(\mathbf{x}, \mathbf{k}', t)(1 - f_n(\mathbf{x}, \mathbf{k}, t))S_n(\mathbf{k}', \mathbf{k})] dk'. \end{aligned} \quad (4.20)$$

В стационарном состоянии, когда $\frac{\partial f}{\partial t} = 0$ и $\left(\frac{\partial f}{\partial t}\right)_{ст} = \left(\frac{\partial f}{\partial t}\right)_{пол}$, изменения функции распределения, создаваемые внешними полями и движением частиц, компенсируются столкновениями частиц. Система кинетических уравнений (4.20) (для электронов и дырок) при имитационном розыгрыше вероятных (по Монте-Карло) сценариев столкновений является чрезвычайно сложной, ее эффективное использование невозможно без определенных упрощений, учитывающих явление релаксации. Процессы столкновений приводят к восстановлению нарушаемого полями равновесного распределения электронов и дырок. Их действие можно описать временем релаксации импульса (инерции) $\tau_p(\mathbf{k})$, равным среднему времени существования неравновесного состояния после выключения полей, вызвавших это отклонение.

В предположении, что время релаксации не зависит от внешних полей и нет вырождения полупроводника, КУБ имеет вид

$$\frac{\partial f_n}{\partial t} + \frac{F_e}{\hbar} \text{grad}_k f_n + v_n \text{grad}_x f_n = -\frac{f_n - f_{n0}}{\tau_p}, \quad (4.21)$$

где f_{n0} — функция распределения в равновесном состоянии.

Умножим члены уравнения (4.21) на групповую скорость v и проинтегрируем их по объему моментов:

$$\int_{V_k} v_n \frac{\partial f_n}{\partial t} dk + \int_{V_k} v_n \left(\frac{F_e}{\hbar} \text{grad}_k f_n \right) dk + \\ + \int_{V_k} v_n (v_n \text{grad}_x f_n) dk = - \int_{V_k} v_n \frac{f_n - f_{n0}}{\tau_p} dk. \quad (4.22)$$

Интегралы в (4.22) записываются следующим образом:

$$\int_{V_k} v_n \frac{\partial f_n}{\partial t} dk = 4\pi^3 \frac{\partial}{\partial t} (n v_n), \quad (4.23)$$

$$\int_{V_k} v_n \left(\frac{F_e}{\hbar} \text{grad}_k f_n \right) dk = \frac{F_e}{\hbar} \int_{V_k} v_n \text{grad}_k f_n dk = -4\pi^3 F_e \frac{n}{m_n^*}, \quad (4.24)$$

$$\int_{V_k} v_n (v_n \text{grad}_x f_n) dk = \frac{4\pi^3}{m_n^*} \text{grad}_x (n k T), \quad (4.25)$$

$$\int_{V_k} v_n \frac{f_n - f_{n0}}{\tau_n} dk = \frac{1}{\tau_n} \int_{V_k} v_n (f_n - f_{n0}) dk = 4\pi^3 \frac{n v_n}{\tau_p}, \quad (4.26)$$

где m_n^* — эффективная масса; T — температура решетки; v_n — дрейфовая скорость.

Пренебрегая магнитными полями и используя соотношения $F_{ne} = -qE$, $F_{pe} = qE$, с помощью выражений (4.22), (4.23)–(4.26) можно получить дифференциальное уравнение для дрейфовой скорости и напряженности электрического поля:

$$\frac{\partial}{\partial t} (n v_n) + \frac{q}{m_n^*} n E + \frac{1}{m_n^*} \text{grad} (n k T) = -\frac{n v_n}{\tau_n}, \quad (4.27)$$

где $n = \int f_n d^3k$, $n v = \int v(k) f_n d^3k$.

Уравнение (4.27) для дырок записывается в виде

$$\frac{\partial}{\partial t} (p v_p) - \frac{q}{m_p^*} p E + \frac{1}{m_p^*} \text{grad} (p k T) = -\frac{p v_p}{\tau_p}. \quad (4.27a)$$

Аналогично, интегрируя уравнения Больцмана с учетом дополнительного соотношения для энергии, справедливого для полупро-

водника с параболической зонной структурой: $E = -\frac{1}{2} m^* v^2 + \frac{3}{2} kT$, можно получить уравнение

$$\frac{\partial E}{\partial t} + v \nabla E = qvE - \frac{1}{n} \nabla (nvkT) + \frac{E - E_0}{\tau_E}, \quad (4.28)$$

где использовано время релаксации энергии τ_E , являющееся функцией локальной средней энергии носителей заряда согласно определению: $\frac{E - E_0}{\tau_E(E)} = \left(\frac{dE}{dt} \right)_{\text{ст}}$.

Вместе с уравнением непрерывности вида

$$\frac{\partial n}{\partial t} + \nabla (nv) = \left(\frac{\partial n}{\partial t} \right)_{\text{ст}} \quad (4.29)$$

или упрощенной записью формул (4.2), (4.3) уравнения (4.27), (4.28) составляют ядро квазигидродинамической или электронно-температурной математической модели полупроводниковой структуры.

Поскольку точное решение этих уравнений невозможно, для получения приближенного решения введем понятие эффективных подвижностей μ_n и μ_p : $\mu_n = q\tau_n/m_n^*$, $\mu_p = q\tau_p/m_p^*$.

Уравнения (4.27), (4.27a) с учетом величин μ_n , μ_p можно переписать таким образом:

$$\tau_n \frac{\partial J_n}{\partial t} + J_n = q\mu_n n \left[E + \frac{1}{n} \text{grad} \left(n \frac{kT}{q} \right) \right], \quad (4.30)$$

$$\tau_p \frac{\partial J_p}{\partial t} + J_p = q\mu_p p \left[E - \frac{1}{p} \text{grad} \left(p \frac{kT}{q} \right) \right]. \quad (4.31)$$

Для малых значений τ_p можно получить приближенные выражения векторов плотностей тока первого порядка:

$$J_{n0} = q\mu_n n \left[E + \frac{1}{n} \text{grad} \left(n \frac{kT}{q} \right) \right], \quad (4.32)$$

$$J_{p0} = q\mu_p p \left[E - \frac{1}{p} \text{grad} \left(p \frac{kT}{q} \right) \right], \quad (4.33)$$

$$J_n = J_{n0} + 0(\tau_n), \quad J_p = J_{p0} + 0(\tau_p). \quad (4.34)$$

В предположении постоянства температуры решетки и выполнения соотношений Эйнштейна

$$D_n = \mu_n \frac{kT}{q}, \quad D_p = \mu_p \frac{kT}{q} \quad (4.35)$$

выражения для векторов плотностей тока с учетом (4.33), (4.34) записываются в виде суммы диффузионного и дрейфового членов, т. е. сводятся к каноническим выражениям (4.4), (4.5).

Таким образом, в рассматриваемом случае электронные и дырочные потоки оказываются функциями концентраций, температур, напряженностей электрического поля и градиентов концентраций и температур, при этом эффективные температуры полупроводника можно считать локальными функциями электрического поля. Система уравнений квазигидродинамической модели дополнительно упрощается и соответствует дрейфово-диффузионному приближению, наиболее распространенному в моделировании полупроводниковых приборов.

4.3. ОГРАНИЧЕНИЯ ФИЗИКО-ТОПОЛОГИЧЕСКИХ МОДЕЛЕЙ. ВОЗМОЖНЫЕ ПОДХОДЫ К МОДЕЛИРОВАНИЮ

Особенности физических явлений в субмикронных полупроводниковых структурах. При уменьшении линейных размеров полупроводниковых структур, а также снижении рабочих температур размеры неоднородностей электронно-дырочной плазмы в структурах становятся соизмеримыми с фундаментальными длинами, характеризующими физические свойства плазмы. К таким фундаментальным длинам относятся:

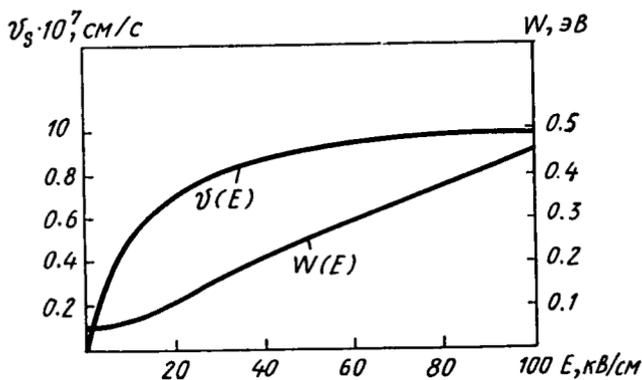


Рис. 4.1. Зависимости скорости дрейфа и средней энергии от напряженности электрического поля ($T=300$ К, $C_D=10^{17}$ см $^{-3}$)

дебройлевская длина волны электронов (дырок)

$$\lambda = \frac{2\pi\hbar}{mv_\tau}, \quad (4.36)$$

длина свободного пробега, или длина релаксации импульса,

$$\lambda_p = v_\tau \tau_p, \quad (4.37)$$

длина «остывания», или длина релаксации энергии,

$$\lambda_E = v_\tau \sqrt{\tau_p \tau_E}, \quad (4.38)$$

где m , v_t , τ_p , τ_E — соответственно характерная эффективная масса, тепловая скорость, времена релаксации импульса и энергии электронов.

Из экспериментальных зависимостей скорости и энергии от напряженности электрического поля (для кремния) (рис. 4.1) определяются соответствующие зависимости времен релаксации τ_p и τ_E от энергии (рис. 4.2), для которых в целом выполняются соотношения (4.37), (4.38).

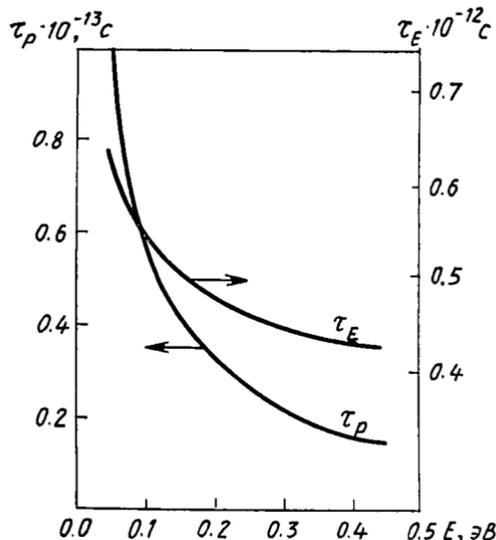


Рис. 4.2. Зависимости времени релаксации инерции и энергии от средней энергии носителей заряда в кремнии ($T = 300$ К, $C_D = 10^{17}$ см $^{-3}$)

Если рассеяние носителей заряда в полупроводнике имеет квазиупругий характер, то $\tau_E \gg \tau_p$. Так, при рассеянии на акустических фононах $\lambda_E/\lambda_p \approx (T/T_s)^{1/2}$, где $T_s \approx ms^2$; s — скорость звука. При эффективной массе носителей заряда $m \approx 10^{-28}$ г $s \approx 5 \cdot 10^5$ см/с, $T_s \approx 0,1$ К.

Таким образом, значения величины λ_E в широком диапазоне температур могут на порядок или более превышать λ_p . Поэтому можно считать справедливым неравенство

$$\lambda \ll \lambda_p \ll \lambda_E. \quad (4.39)$$

Например, для азотных температур ($T \approx 77$ К), $m \approx 10^{-28}$ г, дебройлевская длина волны $\lambda \sim 0,1$ мкм. При подвижных носителях, например в достаточно чистом GaAs, $\mu = 2 \cdot 10^5$ см/(В·с); для тех же азотных температур и эффективных масс $\lambda_p = 0,5 \div 1$ мкм. В то же время при относительно высоких (например, комнатных) температурах в кремнии длины релаксации импульса и «остывания» меньше 0,1 мкм, однако неравенство (4.39) будет справедливо и в этом случае.

Если один из характерных размеров полупроводниковой структуры $l \sim \lambda$, то оказываются существенными квантовые эффекты, которые могут сильно влиять на электрические характеристики и параметры тонкослойных МОП-структур, а также полевых и биполярных гетероструктур с селективным легированием.

Сделаем допущения о несущественном влиянии квантовых эффектов, связанных с соизмеримостью характерных размеров $\lambda \approx l$, и в дальнейшем ограничимся рассмотрением физических явлений при условии $l \gg \lambda$.

Квазигидродинамическое приближение (электронно-температурная модель полупроводниковой структуры). Интегрированием КУБ с учетом дополнительных соотношений для полупроводника с параболической зонной структурой получены три основных уравнения, отражающие законы сохранения количества электронов (4.29), импульса (4.27) и энергии (4.28). Данная система уравнений получена из КУБ достаточно строго, поэтому их области применения для большинства полупроводниковых структур (особенно кремниевых биполярных структур) практически совпадают.

При $\lambda_p \ll \lambda_E$ и l (это означает, что электрон или дырка, проходя расстояние порядка l , испытывает значительное число актов рассеяния) парные столкновения достаточно часты, так что соответствующая длина $\lambda_e \ll \lambda_E$ (однако $\lambda_E \gg \lambda_p$), т. е. возможно описание электронно-дырочной плазмы смещенной функцией распределения Максвелла. Выполнение условия квазигидродинамического приближения задается при этом соотношением для времен релаксации импульса, энергии и частоты парных столкновений электронов:

$$\tau_p \approx \tau_e \ll \tau_E. \quad (4.40)$$

По оценкам, $\tau_e = 10^{13}$ нс/ n , где n — концентрация носителей заряда, см⁻³. Согласно зависимостям на рис. 4.1, 4.2 τ_e ($E = 10^4$ В/см) = 0,6 пс, τ_E ($E = 10^5$ В/см) = 0,43 пс. При $\tau_p \sim 0,05 \div 0,1$ пс условие (4.40) выполняется, строго говоря, для режима в полупроводниковых структурах, когда $n \geq 10^{14}$ см⁻³, $\tau_e \leq 0,1$ пс.

Суть данной модели — учет конечного времени релаксации импульса и энергии, дрейфовой компоненты полной энергии подвижных носителей заряда и температуры электронно-дырочной плазмы, не равной температуре кристаллической решетки полупроводника. Использование системы квазигидродинамических уравнений (4.27) — (4.29) вместо уравнений диффузионно-дрейфовой модели (4.1) — (4.5) позволяет эффективно моделировать полупроводниковые структуры, в которых существенны такие физические эффекты, как разогрев электронно-дырочной плазмы, превышение скорости носителей заряда над скоростью насыщения в областях сильного изменения электрического поля и др. Численные методы решений уравнений квазигидродинамической модели практически не отличаются от методов решений диффузионно-дрейфовой модели и ресурсы ЭВМ в них отличаются незначительно. Однако включение входной информации об электрофизических параметрах квазигидродинамической модели для субмикронных полупроводниковых структур (см., например, рис. 4.2) сопряжено с некоторыми сложностями в определении априорной информации и подтверждении необходимой адекватности для реальных приборных структур, выполненных по конкретной технологии изготовления БИС.

Диффузионно-дрейфовое приближение. Ограничения ФСУ полупроводника. Эффективные температуры полупроводниковой структуры считаются локальными функциями электрического поля для характерных размеров структуры $l \gg \lambda$, при этом система квазигид-

родинамических уравнений переходит в уравнение диффузионно-дрейфового приближения — ФСУ полупроводника (4.1) — (4.5).

Каждое из соотношений ФСУ (4.1) — (4.5), несмотря на достаточно большую общность, универсальность и правомочность, показанную на огромном теоретическом и экспериментальном материале с реальными полупроводниковыми структурами, вплоть до последнего времени имеет ограниченное применение, особенно в свете неуклонной тенденции разработки компонентов БИС с чрезвычайно мелкими геометрическими размерами (вертикальными и горизонтальными), высокими уровнями легирования соответствующих областей, работой при высоких и сверхвысоких плотностях тока в БИС.

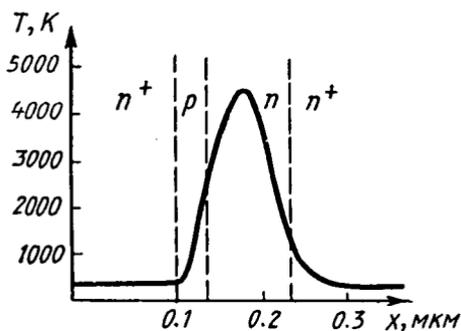


Рис. 4.3. Зависимость температуры электронов БТ от координаты

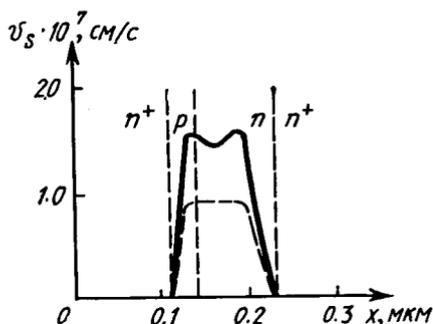


Рис. 4.4. Зависимость дрейфовой скорости электронов БТ от координаты с учетом и без учета (пунктир) разогрева носителей заряда

Вопросам ограничений, присущим «классической» ФСУ при моделировании субмикронных, тонкослойных сильнолегированных компонентов (биполярных, униполярных и др.), посвящено много теоретических и экспериментальных исследований.

На основе анализа выводов этих работ можно выделить ряд основных ограничений: 1) при характерных временах изменения концентраций электронов и дырок, близких к временам максвелловской релаксации 10^{-12} — 10^{-13} с, необходимо учитывать электромагнитный характер потенциала, что приводит к появлению дополнительных членов в уравнении Пуассона; 2) величину E можно считать практически не зависящей от концентрации примесей при $\text{max}(C_A C_D) \leq 10^{21} \text{ cm}^{-3}$; 3) феноменологические электрофизические параметры полупроводника вводят теоретически и измеряют экспериментально при постоянных концентрациях легирующих примесей и носителей заряда. Поэтому на максимальные градиенты концентраций накладываются ограничения $\text{grad } C < C/l_{\text{пр}}$, где $l_{\text{пр}}$ — длина свободного пробега носителей заряда; практически можно полагать

град $C < 10^{24} \div 10^{25} \text{ см}^{-4}$; 4) при уменьшении характерных областей приборных структур до $L \approx 50 \div 150 \text{ нм}$ ставится вопрос о правомочности ФСУ, корректности использования диффузионно-дрейфовых моделей переноса носителей заряда, электрофизических параметров подвижности и др.

Однако для кремниевых БТ даже с характерными размерами областей $\sim 1000 \text{ \AA}$ большинство кинетических, баллистических, квазибаллистических эффектов являются эффектами второго порядка в отличие, например, от аналогичных компонентов, изготовленных по GaAs-технологии ССБИС. Но этот факт позволяет модифицировать классические диффузионно-дрейфовые ФТ-модели в упрощенные электронно-температурные модели, где существует отличие температуры электронного газа от температуры кристаллической решетки. На рис. 4.3—4.5 представлены численные эксперименты моделирования кремниевой тонкослойной БТ с толщиной базы около 300 \AA при учете совокупности физических эффектов, возникающих из-за нагрева электронного газа. С помощью электронно-температурной ФТ-модели были рассчитаны зависи-

мости электронной температуры, дрейфовой скорости электронов от координаты для различных токов (рис. 4.3, 4.4). На основании расчетов был сделан вывод о том, что электронный газ сильно нагревается в области коллекторного перехода и остается практически неизменным в области эмиттерного перехода. Проявлением температурных эффектов является «всплеск» скорости носителей заряда в области коллекторного перехода, что приводит к соответствующему спаду концентрации электронов при неизменной плотности тока ($J_n = \text{const}$). Несмотря на то что «внутренние» физические процессы, предсказываемые «классической» и «электронно-температурной» моделями, принципиально различны, на внешних выходных характеристиках это различие практически (в пределах 10%) не проявляется (рис. 4.5) из-за уменьшения как времени переноса коллекторной области вследствие увеличения дрейфовой скорости, так и времени переноса базовой области вследствие увеличения градиента концентраций неосновных носителей заряда в квазинейтральной базе.

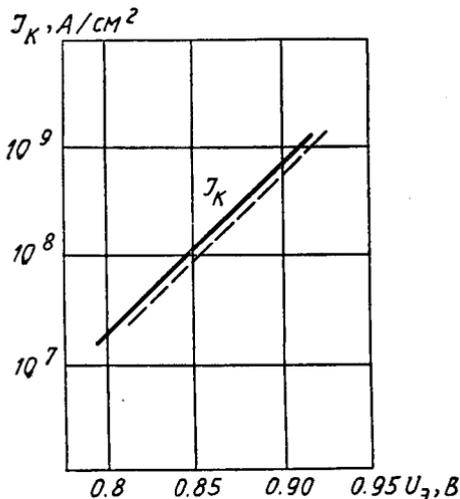


Рис. 4.5. Зависимость плотности коллекторного тока от напряжения на эмиттерном переходе для электронно-температурной и обычной диффузионно-дрейфовой (пунктир) ФТ-моделей

Модифицированный метод Монте-Карло, используемый для расчета кинетических характеристик полевых приборов для сильных электрических полей, был использован для расчета статических и динамических характеристик тонкослойных БТ с заданием реальных профилей распределения примесей, получаемых при изготовлении БИС. Несмотря на значительную разницу скорости и энергии электронов, а также концентрации электронов в базе и коллекторе, рассчитанных с помощью классической модели и метода Монте-Карло (рис. 4.6, а, б), различие выходных ВАХ и времени задержки эмиттер — коллектор БТ из-за «всплеска» скорости не превышало

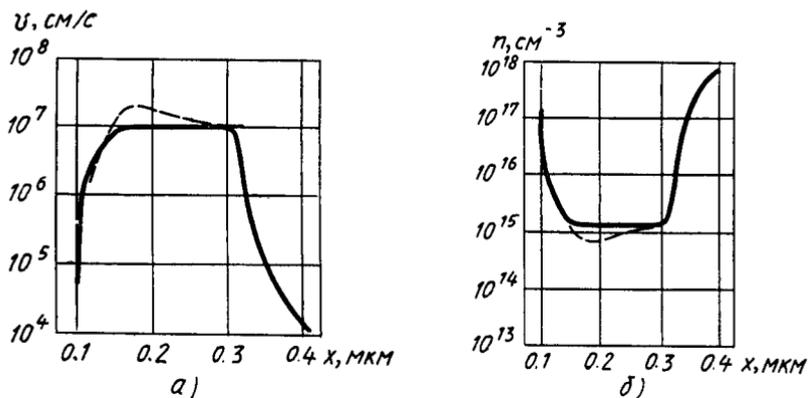


Рис. 4.6. Зависимости скорости (а) и концентрации (б) электронов от координаты тонкослойного БТ, рассчитанные с помощью диффузионно-дрейфовой модели и метода Монте-Карло (пунктир)

также 5—7%. Таким образом, даже для достаточно тонкослойных сильнолегированных биполярных БТ с помощью детальных численных экспериментов была показана правомочность использования модифицированной диффузионно-дрейфовой модели с принципиально достижимой возможностью получения приемлемой адекватности моделирования в широком режимном диапазоне.

4.4. ГРАНИЧНЫЕ УСЛОВИЯ В ЧИСЛЕННОМ МОДЕЛИРОВАНИИ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

Для полной математической формулировки краевой задачи, которую требуется решить в рамках ФТ-моделирования, ФСУ дополняют граничными условиями с учетом специфики топологии полупроводниковых структур. Для рассматриваемых в дальнейшем полупроводниковых структур (в первую очередь БТ) характерны следующие граничные условия.

Для концентраций дырок и электронов граничные условия в общем виде задаются с помощью уравнений для нормальной компоненты соответствующих плотностей токов

$$J_n = qR_s, \quad (4.41)$$

$$J_p = -qR_s, \quad (4.42)$$

где R_s — поверхностный рекомбинационный член.

Для границы раздела Si—SiO₂

$$R_s = \frac{pn - n_{ie}^2}{(p + n_{ie})/s_p + (n + n_{ie})/s_n}. \quad (4.43)$$

Здесь s_p, s_n — скорости поверхностной рекомбинации, сильно зависящие от условий технологического процесса. Обычно $10^3 \text{ см/с} < s_p, s_n < 10^5 \text{ см/с}$.

Для границы раздела Si — металл выражение (4.43) упрощается, поскольку там всегда $n \gg p$ или $p \gg n$. Поэтому условия (4.11), (4.42) можно записать в виде

$$J_n = qs_n(n - n_0), \quad (4.44)$$

$$J_p = -qs_p(p - p_0), \quad (4.45)$$

где n_0, p_0 — равновесные концентрации, которые определяются из уравнений

$$n_0 - p_0 = C_D - C_A,$$

$$n_0 p_0 = n_{ie}^2.$$

Если контакт омический, то скорости s_n и s_p очень велики (10^6 см/с и более). Тогда из выражений (4.44) и (4.45) следует, что

$$n = n_0, \quad p = p_0. \quad (4.46)$$

Особый интерес представляет контакт системы кремний — поликремний (Si—Si*). Весьма эффективным является описание граничных условий на границе раздела Si—Si* с помощью конечной скорости рекомбинации неосновных носителей s_p . В теоретических и экспериментальных исследованиях показано, что параметр s_p существенно зависит от технологии изготовления прибора, в первую очередь от качества и типа поверхностной обработки кремния до осаждения на него поликремниевого слоя. Для различных технологий БИС значение s_p имеет значительный разброс: $s_p \approx (0.3 \div 9) \times 10^4 \text{ см/с}$.

В данном случае необходимо пользоваться условиями (4.44), (4.45) (для контакта Si с кремнием n -типа). Для контакта Шотки также используются выражения (4.44), (4.45), где $s_n = A_n T^2 / N_c$, $s_p = A_p T^2 / N_v$; A_n, A_p — эффективные константы Ричардсона для электронов и дырок; N_c, N_v — эффективные плотности состояний в зоне проводимости и валентной зоне.

Для потенциала V граничное условие на омическом контакте задается из условия равновесия основных носителей заряда:

$$V = U_a + \varphi_T \ln \frac{n}{n_{ie}}, \quad (4.47)$$

где U_a — приложенное внешнее напряжение.

Для контакта Шотки поверхностный электростатический потенциал задается выражением

$$V = \Phi_m - |\Phi_b| + \frac{1}{2} |E|_g, \quad (4.48)$$

где Φ_m — энергия Ферми металла; Φ_b — высота барьера Шотки; E_g — ширина запрещенной зоны.

На границе Si—SiO₂ нормальная составляющая электростатического поля задается плотностью эффективного поверхностного заряда ρ_s на границе Si—SiO₂:

$$-\varepsilon \varepsilon_0 \nabla V n = q \rho_s. \quad (4.49)$$

Величина ρ_s также сильно зависит от технологического процесса.

На остальных топологических «отражающих» границах нормальные составляющие плотности токов дырок и электронов, а также напряженность электрического поля принимаются равными нулю:

$$n \mathbf{J}_n = n \mathbf{J}_p = n \nabla V = 0. \quad (4.50)$$

В рамках двумерного ФТ-моделирования общие граничные условия Неймана (4.50) трансформируются в более простые для y -координаты:

$$\frac{\partial V}{\partial y} = \frac{\partial n}{\partial y} = \frac{\partial p}{\partial y} = 0. \quad (4.51)$$

Кратко остановимся на некоторых специфических особенностях, возникавших при моделировании МОП-структур. В этом случае для определения потенциала электрического поля помимо области кремния необходимо рассматривать область окисла (под затвором, между истоком и стоком). Полагается, что в этой области объемная плотность заряда $\rho = 0$ и соответственно потенциал электрического поля удовлетворяет условию

$$\Delta V = 0. \quad (4.52)$$

На границе раздела диэлектрик—полупроводник выполняются граничные условия

$$V^+ = V^-, \quad \left(\frac{\partial V}{\partial y} \right)^+ = \left(\frac{\partial V}{\partial y} \right)^-, \quad (4.53)$$

$$\varepsilon_{Si} \varepsilon_{SiO_2} \left(\frac{\partial V}{\partial x} \right)^+ - \varepsilon_{SiO_2} \varepsilon_{Si} \left(\frac{\partial V}{\partial x} \right)^- = Q_s, \quad (4.54)$$

где V^+ , $\left(\frac{\partial V}{\partial x} \right)^+$, $\left(\frac{\partial V}{\partial y} \right)^+$, V^- , $\left(\frac{\partial V}{\partial y} \right)^-$, $\left(\frac{\partial V}{\partial x} \right)^-$ — предельные значения потенциала и его производных со стороны кремния и диэлектрика соответственно; ε_{Si} , ε_{SiO_2} — диэлектрическая проницаемость кремния и окисла; Q_s — поверхностная плотность за-

ряда, которая в общем случае считается известной (обычно $Q_s=0$).

На границе раздела кремний — окисел ставятся условия (4.41), (4.42), при этом уравнения непрерывности рассматриваются только в области кремния. На контакте затвора задается значение потенциала электрического поля. Поскольку толщина окисла обычно много меньше его ширины, для упрощения на краях окисла задают значение потенциала с помощью линейной интерполяции.

Некоторая искусственность граничных условий (4.50), (4.51) для реальных топологий полупроводниковых структур связана с одной из узловых проблем адекватности численного ФТ-расчета: насколько велика погрешность идеализации при моделировании реальных приборов с произвольной топологией БИС? В качестве одного из альтернативных подходов при ФТ-моделировании полупроводниковых структур со сложной топологией используют размещение соответствующих границ достаточно далеко от области структуры, определяющей основные электрические характеристики исследуемого объекта. Применение систематизированного набора граничных условий (4.41) — (4.51) в той или иной программе ФТ-моделирования, естественно, не исчерпывает всех возможностей и подходов адекватного описания сложных полупроводниковых структур.

4.5. МОДЕЛЬ ИСКАЖЕНИЙ ЭНЕРГЕТИЧЕСКИХ ЗОН В СИЛЬНОЛЕГИРОВАННОМ КРЕМНИИ

Несмотря на то что в производстве дискретных полупроводниковых приборов, ИМС и БИС кремний давно является самым распространенным материалом, его фундаментальные электрические характеристики, связанные с зонной структурой сильнолегированных областей, изучены недостаточно.

В сильнолегированном кремнии зонную структуру уже нельзя считать не зависящей от концентрации примесей и подвижных носителей заряда вследствие:

- 1) взаимодействия между подвижными носителями, а также носителями и ионизированными примесями;
- 2) флуктуации локального электростатического потенциала из-за случайного распределения примесей высоких концентраций;
- 3) образования из дискретных энергетических уровней примесных атомов отдельной примесной энергетической зоны;
- 4) вырождения основных носителей заряда при высоких концентрациях примесей.

Данные физические эффекты приводят к соответствующим смещениям (с произвольным или достаточно «жестким» сдвигом) краев валентной зоны и зоны проводимости с уменьшением ширины запрещенной зоны.

Точный теоретический расчет искажений энергетических зон, сужения запрещенной зоны в сильнолегированном кремнии весьма затруднителен, сложен и в целом не может удовлетворить требова-

ниям численного ФТ-моделирования сильнолегированных полупроводниковых структур. Более надежным и практически необходимым является развитие феноменологической модели искажения энергетических зон в сильнолегированном кремнии на основе систематизированных экспериментальных данных об «оптической» $E_g^{оп}$ и «электрической» $E_g^{эл}$ ширине запрещенной зоны с описанием физических механизмов ее уменьшения при высоких концентрациях примеси и подвижных носителей заряда. Величины зоны $E_g^{оп}$ и ее сужения $\Delta E_g^{оп}$ оценивались из оптического эксперимента определения края зоны поглощения инфракрасного излучения. Оценки величин зоны $E_g^{эл}$ и ее сужения $\Delta E_g^{эл}$ основаны в общем случае на измерениях электрических параметров, связанных с рекомбинацией неосновных носителей заряда в сильнолегированных областях полупроводника. В рамках модели искажения зон не только отражены смещения краев зоны проводимости ΔE_c и валентной зоны ΔE_v и соответствующие сужения зон $\Delta E_g^{оп}$, $\Delta E_g^{эл}$, но и объяснено принципиальное соотношение $\Delta E_g^{эл} - \Delta E_g^{оп} > 0$ в средне- и сильнолегированных полупроводниках, наблюдаемая экспериментально.

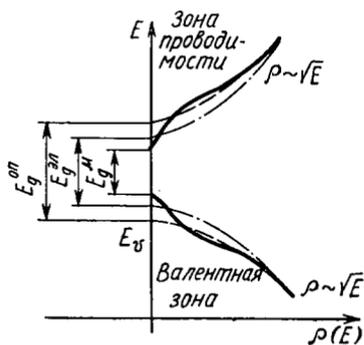


Рис. 4.7. Зависимость плотности состояний от энергии в сильнолегированном кремнии. Модельные представления о трех типах «запрещенных» зон в полупроводнике

Суть модели искажения зон с выделением трех типов запрещенных зон проиллюстрирована на рис. 4.7 для обобщенной функции квантовой плотности состояний в n^+ -кремнии, которая включает в себя эффекты «многих тел», связанных как с многодолинными взаимодействиями примесей и носителей заряда, так и со случайными распределениями примесей. Оптическая запрещенная зона $E_g^{оп}$ является истинной, разделяющей зону проводимости и валентную зону и в конечном счете управляющей процессом поглощения фотонов. Эффекты «многих тел» приводят к некоторым изгибам у краев зон и обуславливают действительное сужение $\Delta E_g^{оп}$:

Суть модели искажения зон с выделением трех типов запрещенных зон

проиллюстрирована на рис. 4.7 для обобщенной функции квантовой плотности состояний в n^+ -кремнии, которая включает в себя эффекты «многих тел», связанных как с многодолинными взаимодействиями примесей и носителей заряда, так и со случайными распределениями примесей. Оптическая запрещенная зона $E_g^{оп}$ является истинной, разделяющей зону проводимости и валентную зону и в конечном счете управляющей процессом поглощения фотонов. Эффекты «многих тел» приводят к некоторым изгибам у краев зон и обуславливают действительное сужение $\Delta E_g^{оп}$:

$$\Delta E_g^{оп} = E_{g0} - E_{g0}^{оп} = \Delta E_c^{mb} + \Delta E_v^{mb} \quad (4.55)$$

относительно фундаментальной величины запрещенной зоны собственного кремния (при низких концентрациях легирующих примесей).

Эффекты образования хвостов валентной зоны и зоны проводимости не влияют на изменения $\Delta E_g^{оп}$ (4.55), поскольку в хвостовой части функция плотности состояний (рис. 4.7) представляет собой статистические усреднения локальных плотностей состояний по всей кристаллической решетке. Эти эффекты приводят к «жесткому» сдвигу зон при сохранении их параболичности, т. е. сохранении за-

висимости плотности состояний пропорционально $E^{1/2}$. Причем «жесткие» сдвиги краев зоны проводимости ΔE_c^{bt} и валентной зоны ΔE_v^{bt} обуславливают подвижную запрещенную зону E_g^μ , которая в конечном итоге определяет макроскопические плотности подвижных носителей заряда, используемые в расчете электрических характеристик полупроводниковых структур. В модели искажений зон электрическая запрещенная зона $E_g^{\text{эл}}$ является эффективной величиной, отражающей суть «жестких» сдвигов подвижной зоны E_g^μ . Таким образом, сужение электрической запрещенной зоны

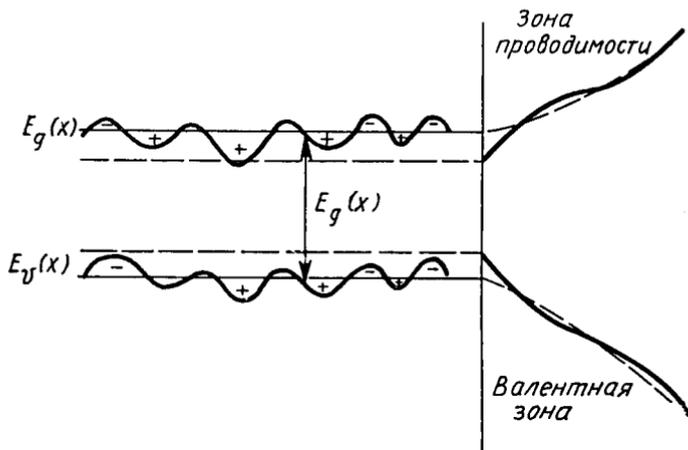


Рис. 4.8. Образование хвостов в зоне проводимости и валентной зоне из-за флуктуации электростатического потенциала

$\Delta E_g^{\text{эл}}$, обуславливающей изменение концентраций подвижных носителей заряда в полупроводниковых структурах, определяется как эффектами «многих тел», так и «хвостовыми» эффектами:

$$\Delta E_g^{\text{эл}} = E_{g0} - E_g^{\text{эл}} = \Delta E_c^{mb} + \Delta E_c^{bt} + \Delta E_v^{mb} + \Delta E_v^{bt}. \quad (4.56)$$

Величина смещения $\Delta E_c^{mb} \sim (E_F - E_c)^{1/2}$ учитывает энергию квантово-механического взаимодействия электрон — электрон в системе «многих тел» для высоких концентраций электронов в полупроводнике. Величина смещения ΔE_v^{mb} зависит от концентрации электронов $\Delta E_v^{mb} \sim n^{1/4}$ и учитывает кулоновское взаимодействие электрон — дырка, приводящее к уменьшению потенциальной энергии дырок за счет эффекта экранирования их основными носителями заряда — электронами. Теоретические исследования показали, что взаимодействие подвижных носителей заряда с примесями, а также взаимодействие дырка — дырка в системе «многих тел» оказывают пренебрежимо малое влияние на смещения E_c , E_v и, следовательно, на ΔE_g .

Так как действительное распределение ионизированных примесей в легированном кремнии не однородное, а случайное, возника-

ют флуктуации локального потенциала, приводящие к искажению квантовой плотности состояний, что вызывает нарушение параболичности энергетических зон около экстремума (рис. 4.8). Функции макроскопической плотности состояний в общем случае позволяют произвести расчет эффективного «жесткого» сдвига зоны проводимости $\Delta E_c^{bt} = E_c - E_{c\text{эф}}$ и валентной зоны $\Delta E_v^{bt} = E_{v\text{эф}} - E_v$, где энергии эффективных краев параболических зон (проводимости и валентной) определяются интегралами плотностей занятых состояний в соответствующих зонах.

Отметим пренебрежимо малое влияние примесной зоны на смещения E_c , E_v и ΔE_g .

Основываясь на рассмотренных представлениях искажений энергетических зон с учетом их «жесткого» сдвига и сохранением параболичности, для сильнолегированного кремния концентрации дырок и электронов можно выразить следующим образом:

$$n_0 = N_c F_{1/2} \left(\frac{E_E - E_{c\text{эф}}}{kT} \right) = N_c F_{1/2}(\eta_c), \quad (4.57)$$

$$p_0 = N_v \exp \left(\frac{E_{v\text{эф}} - E_F}{kT} \right) = N_v \exp(-\eta_v). \quad (4.58)$$

Здесь N_c , N_v — эффективные плотности состояний в зоне проводимости и валентной зоне; $N_c = 3 \cdot 1 \cdot 10^{19} \text{ см}^{-3}$, $N_v = 1.8 \cdot 10^{19} \text{ см}^{-3}$ при $T = 300 \text{ К}$; $F_{1/2}$ — интеграл Ферми—Дирака половинного индекса.

Важнейший параметр полупроводника — произведение

$$p_0 n_0 = n_{ie}^2, \quad (4.59)$$

которое, согласно выражениям (4.57), (4.58), записывается в виде

$$n_{ie}^2 = n_{i0}^2 \exp \frac{\Delta E_g^{\text{эл}}}{kT} \frac{F_{1/2} \left(\frac{E_F - E_{c\text{эф}}}{kT} \right)}{\exp \left(\frac{E_F - E_{v\text{эф}}}{kT} \right)}, \quad (4.60)$$

где

$$n_{i0}^2 = 9.6 \cdot 10^{32} T^3 \exp(1.206 \text{ эВ}/kT) \quad (4.61)$$

для нелегированного полупроводника.

Для произвольных уровней легирования $n_{ie} \geq n_{i0}$ в (4.60) отношение $F_{1/2}(\eta_c)/\exp \eta_v \leq 1$ для статистики Ферми—Дирака и равно единице для статистики Больцмана. В связи с этим в работах по экспериментальному определению $\Delta E_g^{\text{эл}}$, выполненных до начала 80-х годов, существенно занижена оценка этой величины, так как в основу расчета закладывалась статистика Больцмана и эффектами вырождения в полупроводниках пренебрегали.

Учет влияния вырождения в рамках статистики Ферми—Дирака позволяет скорректировать величину сужения электрической запрещенной зоны дополнительным членом ΔE_g^{FD} :

$$\Delta E_g^{эл,FD} = \Delta E_g^{эл} + \Delta E_g^{FD}, \quad \Delta E_g^{FD} = kT \ln \left[\frac{\exp(\eta_D)}{F_{1/2}(\eta_c)} \right]. \quad (4.62)$$

Зависимости ΔE_g^{FD} приведены на рис. 4.9 и учтены при пересчете старых экспериментальных результатов определения $\Delta E_g^{эл}$. На рис. 4.10 представлены для сравнительного анализа результаты идентификации $\Delta E_g^{эл,FD}$, $\Delta E_g^{оп}$ на основе электрических и оптических измерений.

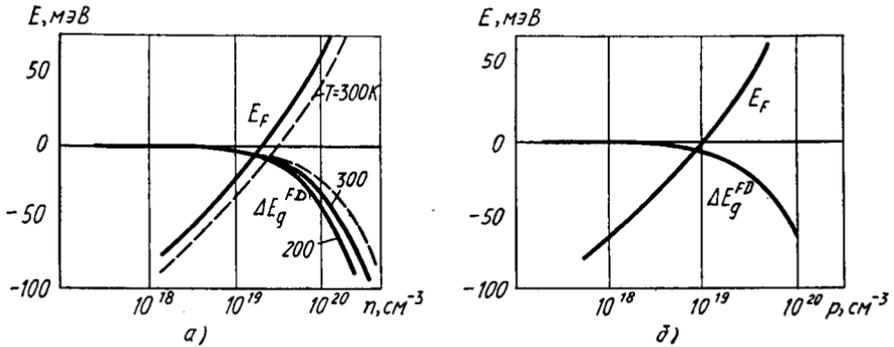


Рис. 4.9. Зависимость корректирующего фактора статистики Ферми — Дирака и уровня Ферми для полупроводника n -типа от концентрации основных носителей заряда и температуры (а): $C = 2.85 \cdot 10^{19} \text{ см}^{-3}$ — сплошные кривые, $C = 4.38 \cdot 10^{19} \text{ см}^{-3}$ — пунктирные кривые; зависимость корректирующего фактора статистики Ферми — Дирака и уровня Ферми для полупроводника p -типа (б): $C = 1.2 \cdot 10^{19} \text{ см}^{-3}$, 300 К

Несмотря на некоторые расхождения (и даже значительный разброс значений $\Delta E_g^{эл}$, $\Delta E_g^{оп}$), можно выделить два принципиально различных характера изменения ΔE_g в электрических и оптических измерениях. Возможно сгруппировать все имеющиеся на сегодня измерения ΔE_g вокруг

двух ветвей зависимости $\Delta E_g^{эл}$, $\Delta E_g^{оп}$ от концентрации, причем для всего диапазона концентраций $\Delta E_g^{эл} > \Delta E_g^{оп}$ (рис. 4.10).

Детальные многочисленные эксперименты показали, что влияние компенсирующей примеси на величину $\Delta E_g^{эл}$ незначительно вплоть до концентраций 10^{19} см^{-3} . Данные результаты являются достаточно сильным аргументом в пользу рассмот-

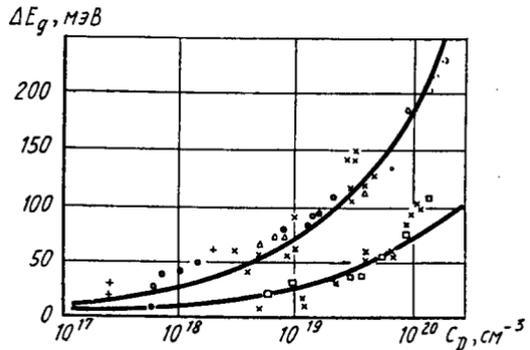


Рис. 4.10. Модельные и экспериментальные зависимости сужения запрещенной зоны оптической и электрической от концентрации донорной примеси

ренной модели искажений зон, в которой основное (оптическое) сужение запрещенной зоны обуславливается многодолинным взаимодействием подвижных носителей заряда, а компенсирующие примесные атомы создают дополнительный (электрический) эффект сужения зоны только при очень сильном легировании. Поэтому правомочно предполагать, что практически все электрофизические параметры легированного кремния — квазиуровни Ферми, сужение запрещенной зоны со стороны зон проводимости и валентной зоны, асимметрия сужения и др. — зависят только от разности концентраций n - p или p - n .

4.6. СИНТЕЗ МОДЕЛЕЙ ПЕРЕНОСА НОСИТЕЛЕЙ ЗАРЯДА ПОЛУПРОВОДНИКОВЫХ СТРУКТУР С УЧЕТОМ ЭФФЕКТОВ СИЛЬНОГО ЛЕГИРОВАНИЯ

Обобщенные соотношения для плотностей тока. При выводе обобщенных соотношений для плотностей тока J (4.4), (4.5) полупроводника с несильным легированием, параболической зонной структурой рассмотрим соотношения Больцмана в состоянии термодинамического равновесия:

$$n = n_{i0} \exp \frac{E_F - E_i}{kT} = n_{i0} \exp \left(\frac{\psi - \varphi}{\varphi_T} \right), \quad (4.63)$$

$$p = n_{i0} \exp \frac{E_i - E_F}{kT} = n_{i0} \exp \left(\frac{\varphi - \psi}{\varphi_T} \right), \quad (4.64)$$

где ψ , φ — потенциалы, соответствующие энергии середины запрещенной зоны E_i и уровня Ферми E_F ($\psi = E_i/q$, $\varphi = E_F/q$).

В состоянии термодинамического равновесия $pn = n^2_{i0}$, однако при подаче на полупроводниковую структуру рабочих напряжений происходит изменение концентрации — неосновных носителей заряда, при этом $pn \neq n^2_{i0}$. Определим энергию уровня Ферми и квазиуровня Ферми с помощью выражений

$$n = n_{i0} \exp \left(\frac{E_{Fn} - E_i}{kT} \right), \quad p = n_{i0} \exp \left(\frac{E_i - E_{Fp}}{kT} \right), \quad (4.65)$$

$$n = n_{i0} \exp \left(\frac{V - \varphi_n}{\varphi_T} \right), \quad p = n_{i0} \exp \left(\frac{\varphi_p - V}{\varphi_T} \right). \quad (4.66)$$

Из выражений (4.4), (4.64) с учетом того, что $E = -\nabla V$, получим

$$\begin{aligned} J_n &= q\mu_n \left(nE + \frac{kT}{q} \nabla n \right) = q\mu_n n (-\nabla V) - \\ &- q\mu_n \frac{kT}{q} \left[\frac{qn}{kT} (\nabla V - \nabla \varphi_n) \right] = -q\mu_n \nabla \varphi_n = \mu_n n \nabla E_{Fn}. \end{aligned} \quad (4.67)$$

Аналогично, плотность дырочного тока

$$J_p = -q\mu_p p \nabla \varphi_p = \mu_p p \nabla E_{Fp}. \quad (4.68)$$

Покажем, что выражения для плотностей тока (4.67), (4.68) будут выполняться и для сильнолегированного полупроводника, где состояние термодинамического равновесия описывается функцией распределения Ферми:

$$f_{n0}(\mathbf{x}, \mathbf{k}) = \frac{1}{1 + \exp \frac{E_c(\mathbf{x}, \mathbf{k}) - E_{Fn}(\mathbf{x})}{kT(\mathbf{x})}}, \quad (4.69)$$

$$f_{p0}(\mathbf{x}, \mathbf{k}) = \frac{1}{1 + \exp \frac{E_{Fn}(\mathbf{x}) - E_v(\mathbf{x}, \mathbf{k})}{kT(\mathbf{x})}}.$$

Воспользуемся КУБ (4.21). Предполагая

$$\left| \frac{\partial f_n}{\partial t} \right| \ll \left| \frac{F_e}{\hbar} \cdot \text{grad} f_n + \mathbf{v}_n \text{grad} f_n \right|, \quad (4.70)$$

получим упрощенную запись КУБ

$$\frac{F_e}{\hbar} \text{grad}_k f_n + \mathbf{v}_n \text{grad}_x f_n = -\frac{f_n - f_{n0}}{\tau_p}, \quad (4.71)$$

достаточно строгую только для небольших отклонений от состояния термодинамического равновесия, описываемого коррекцией функции распределения $f_{n1}(\mathbf{x}, \mathbf{k}, t)$:

$$f_n(\mathbf{x}, \mathbf{k}, t) = f_{n0}(\mathbf{x}, \mathbf{k}) + f_{n1}(\mathbf{x}, \mathbf{k}, t). \quad (4.72)$$

Из выражений (4.71), (4.72) легко оценить коррекцию функции распределения f_{n1} к функции распределения в состоянии термодинамического равновесия (4.69):

$$f_{n1} \approx -\tau_p \left(\frac{F_e}{\hbar} \text{grad}_k f_{n0} + \mathbf{v}_n \text{grad}_x f_{n0} \right). \quad (4.73)$$

Градиенты функции распределения f_{n0} и f_{p0} для электронов и дырок (4.69) записываются следующим образом:

$$\text{grad}_x f_{n0} = f_{n0} (1 - f_{n0}) \text{grad}_x \left(\frac{qV + E_{Fn}}{kT} \right), \quad (4.74)$$

$$\text{grad}_x f_{p0} = -f_{p0} (1 - f_{p0}) \text{grad}_x \left(\frac{qV + E_{Fp}}{kT} \right), \quad (4.75)$$

$$\text{grad}_k f_{n0} = -f_{n0} (1 - f_{n0}) \frac{\hbar^2 \mathbf{k}}{m_n^* kT}, \quad (4.76)$$

$$\text{grad}_k f_{p0} = -f_{p0} (1 - f_{p0}) \frac{\hbar^2 \mathbf{k}}{m_p^* kT}, \quad (4.77)$$

а групповые скорости электронов и дырок (4.19)

$$\mathbf{v}_n = \hbar \mathbf{k} / m_n^*, \quad \mathbf{v}_p = \hbar \mathbf{k} / m_p^*. \quad (4.78)$$

В предположении пренебрежимо малого изменения температуры ($\text{grad } T = 0$) при подстановке соотношений (4.74) — (4.78) в (4.73) получим

$$f_n \approx f_{n0} - \tau_p f_{n0} (1 - f_{n0}) \frac{v_n}{kT} \text{grad}_x E_{Fn}, \quad (4.79)$$

$$f_p \approx f_{p0} + \tau_p f_{p0} (1 - f_{p0}) \frac{v_p}{kT} \text{grad}_x E_{Fp} \quad (4.80)$$

Выражения для плотностей тока носителей заряда

$$J_n = -\frac{q}{4\pi^3} \int v_n f_n dk, \quad J_p = \frac{q}{4\pi^3} \int v_p f_p dk \quad (4.81)$$

при использовании соотношений (4.79), (4.80) для обычно выполняемых условий $f_{n0} \ll 1$, $f_{p0} \ll 1$ записываются в виде

$$J_n = -\frac{q}{4\pi^3} \int v_n f_n dk = -q \mu_n n \text{grad } \varphi_n = \mu_n n \nabla E_{Fn}, \quad (4.82)$$

$$J_p = \frac{q}{4\pi^3} \int v_p f_p dk = -q \mu_p p \text{grad } \varphi_p = \mu_p p \nabla E_{Fp}, \quad (4.83)$$

что аналогично выражениям (4.67), (4.68) в случае слабого легирования. Отсюда можно сделать вывод, что обобщенная запись уравнений переноса инвариантна произвольным уравнениям легирования.

Иерархический набор уравнений плотности тока при сильном легировании полупроводника. В модели искажения энергетических зон с переменным сильным легированием концентрация дырок зависит от двух независимых величин: координаты x и параметра $\eta_v = (E_v - E_{Fp})/kT$. Аналогично, концентрация электронов зависит от координаты x и параметра $\eta_c = (E_{Fn} - E_c)/kT$. «Жесткие» сдвиги краев энергетических

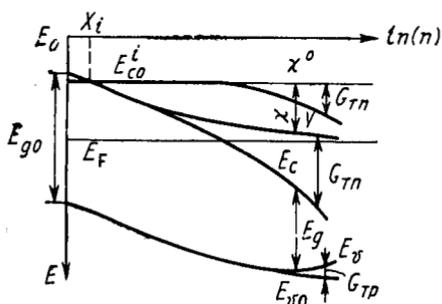


Рис. 4.11. Представление зонной структуры полупроводника с переменным сильным легированием при нулевых внешних потенциалах (E_0^i — энергетический уровень дна зоны проводимости в отсутствие сужения запрещенной зоны)

зон могут быть выражены через абсолютный энергетический уровень вакуума, свободного от электрических полей, а также электростатический потенциал V , ширину запрещенной зоны ΔE_g (электрической) и параметр «электронного сродства» полупроводника χ (рис. 4.11):

$$E_v = E_0 - \chi - E_g - qV, \quad (4.84)$$

$$E_c = E_0 - \chi - qV. \quad (4.85)$$

Соотношения для градиентов сдвига краев энергетических зон и электростатического потенциала записываются в виде

$$\nabla E_v = q \nabla V - \nabla (\chi + E_g), \quad (4.86)$$

$$\nabla E_c = -q \nabla V - \nabla \chi. \quad (4.87)$$

Так как координата квазиуровня Ферми для электронов относительно подвижного края энергетической зоны определяется параметром η_c , можно оценить изменение концентрации n :

$$\nabla n = \nabla_{\eta_c} n + \frac{\partial n}{\partial \eta_c} \nabla \eta_c, \quad (4.88)$$

где $\nabla_{\eta_c} n$ — градиент n при $\eta_c = \text{const}$.

Используя определение параметра $\eta_c = (E_{Fn} - E_c)/kT$ и соотношение (4.88), можно выразить градиент энергии уровня Ферми как

$$\nabla E_{Fn} = \nabla E_c + kT \left[\frac{\nabla n - \nabla_{\eta_c} n}{\frac{\partial n}{\partial \eta_c}} \right]. \quad (4.89)$$

Подставив это выражение в обобщенное уравнение плотности тока (4.82), получим

$$J_n = n \mu_n \nabla E_c - n \mu_n kT \frac{\nabla_{\eta_c} n}{\frac{\partial n}{\partial \eta_c}} + \frac{kT \mu_n n}{\frac{\partial n}{\partial \eta_c}} \nabla n. \quad (4.90)$$

Запишем обобщенное соотношение Эйнштейна для электронов:

$$\lambda_n = \frac{D_n}{\mu_n} = \frac{kT}{q} \frac{n}{\frac{\partial n}{\partial \eta_c}} = \frac{1}{q} \left(\frac{\partial E_{Fn}}{\partial \ln(n)} \right)_{E_c}, \quad (4.91)$$

а с помощью параметра Γ выразим эффект изменения плотности состояний для зоны проводимости в модели искажений энергетических зон:

$$\nabla \Gamma_n = \frac{kT}{\frac{\partial n}{\partial \eta_c}} \nabla_{\eta_c} n. \quad (4.92)$$

Тогда уравнение переноса плотности тока электронов в изотермических условиях с использованием соотношений (4.87) и (4.90) — (4.92) запишется как

$$J_n = -qn \mu_n \left[\nabla \left(V + \frac{\chi}{q} + \frac{\Gamma_n}{q} \right) \right] + q D_n \nabla n. \quad (4.93)$$

Вводя аналогично обобщенное соотношение Эйнштейна для дырок

$$\lambda_p = \frac{D_p}{\mu_p} = \frac{kT}{\frac{\partial p}{\partial \eta_v}} = \frac{1}{q} \left(\frac{\partial E_{Fp}}{\partial \ln(p)} \right)_{E_v} \quad (4.94)$$

и выражая с помощью параметра Γ_p эффект изменения плотности состояний для валентной зоны

$$\nabla \Gamma_p = \frac{kT}{\frac{\partial p}{\partial \eta_v}} \nabla_{\eta_v} p, \quad (4.95)$$

получим с учетом (4.86) уравнение переноса плотности тока дырок в изотермических условиях

$$\mathbf{J}_p = -q p \mu_p \left[\nabla \left(V + \frac{\chi + E_g}{q} - \frac{\Gamma_p}{q} \right) \right] - q D_p \nabla p. \quad (4.96)$$

Как было выяснено, экспериментально весьма затруднительно разделить эффекты, вызванные искажениями энергетических зон полупроводника и влиянием статистики Ферми — Дирака в условиях вырождения полупроводника при его сильном легировании. С этой точки зрения для удобства использования уравнений (4.93), (4.96) корректно сместить эффекты вырождения в диффузионных членах данных уравнений путем соответствующей записи коэффициентов D_n , D_p (4.91), (4.94) в дрейфовые члены (4.93), (4.96). Тогда уравнения (4.93), (4.96) будут приведены к виду

$$\mathbf{J}_p = -q \mu_p p \nabla \Phi_p - kT \mu_p \nabla p, \quad (4.97)$$

$$\mathbf{J}_n = -q \mu_n n \nabla \Phi_n + kT \mu_n \nabla n, \quad (4.98)$$

где Φ_p , Φ_n — эффективные квазипотенциалы, действующие на дырки в валентной зоне и на электроны в зоне проводимости:

$$\Phi_n = \frac{\Gamma_n + \chi}{q} + V, \quad \Phi_p = \frac{\chi + E_g - \Gamma_p}{q} + V. \quad (4.99)$$

С учетом соотношений для ширины электрической запрещенной зоны $E_g^{\text{эл}} = E_{g0} - G_{\text{тп}} - G_{\text{тн}}$ (рис. 4.11) и Φ_n (4.99) градиенты эффективных квазипотенциалов могут быть представлены в виде

$$\nabla \Phi_n = \frac{1}{q} \nabla [(G_{\text{тн}} + \Gamma_n) + V], \quad \nabla \Phi_p = -\frac{1}{q} \nabla [(G_{\text{тп}} + \Gamma_p) + V]. \quad (4.100)$$

Обобщенная запись уравнений переноса (4.97), (4.98) с учетом выражения (4.100) включает четыре параметра, обуславливающие сдвиги энергетических зон и сужение ширины запрещенной зоны полупроводника. Однако нет необходимости вычислять или экспериментально идентифицировать их по отдельности, так как в уравне-

ния для плотностей тока входят лишь суммы $\Gamma_n + G_{\tau n} = \Delta E_{gn}$ и $\Gamma_p + G_{\tau p} = \Delta E_{gp}$. Поэтому для численного моделирования полупроводниковых структур практическую ценность имеют только два модельных параметра:

сужение электрической запрещенной зоны

$$\Delta E_g = \Gamma_n + G_{\tau n} + \Gamma_p + G_{\tau p} \quad (4.101)$$

и фактор асимметрии сужения

$$A_{BN} = (\Gamma_n + G_{\tau n}) / \Delta E_g. \quad (4.102)$$

Фактор асимметрии сужения электрической запрещенной зоны (4.102) в общем случае показывает, какую часть глобального сужения зоны составляет сужение у зоны основных носителей заряда.

Тогда обобщенные уравнения для плотностей тока носителей можно записать в форме, удобной для численного моделирования:

$$\mathbf{J}_p = -q p \mu_p \left[\nabla \left(V - (1 - A_{BN}) \frac{\Delta E_g}{q} \right) \right] - k T \mu_p \nabla p, \quad (4.103)$$

$$\mathbf{J}_n = -q n \mu_n \left[\nabla \left(V + A_{BN} \frac{\Delta E_g}{q} \right) \right] + k T \mu_n \nabla n. \quad (4.104)$$

Коэффициент λ_n в обобщенном соотношении Эйнштейна (4.91) определяется из уравнения (4.98) для плотности тока основных носителей заряда. Для условий невысокого уровня инжекции при допущении, что плотность тока \mathbf{J}_n значительно меньше каждого из слагаемых в правой части уравнения (4.98), следует

$$\lambda_n = \frac{n}{q} \frac{\partial (qV + \Delta E_{g1})}{\partial n} = - \frac{\partial (E_{c0} - \Delta E_{gn} A_{BN})}{q \partial (\ln(n))}. \quad (4.105)$$

Выражение, аналогичное (4.105), может быть получено и для λ_p . Выражение (4.105) выполняется только для основных носителей заряда в квазинейтральных областях приборов. В области p - n -перехода для случая ограничения несильного его легирования $C_{A,D}|_{p-n} \leq 10^{19} \text{ см}^{-3}$ и отсутствия вырождения $p, n \ll N_v N_c$, $\lambda = \varphi_T$, аналогично для неосновных носителей $\lambda = \varphi_T$.

С учетом обобщенного соотношения Эйнштейна уравнение для плотности электронного тока

$$\mathbf{J}_n = q \mu_n \left[\lambda_n \nabla n - n \nabla \left(A_{BN} \frac{\Delta E_g}{q} + V \right) \right] \quad (4.106)$$

в зависимости от того, как описываются параметры зонной структуры полупроводника ΔE_g и A_{BN} , представляет некоторый иерархический набор уравнений различного уровня сложности. Наиболее точным (и сложным) описанием в уравнении (4.106) эффектов высоко-го легирования может служить включение в него экспериментально идентифицированных параметров из электрических измерений $\Delta E_g(C, T)$ и $A_{BN}(C, T)$, зависящих от температуры. Обзор экспери-

ментальных зависимостей $\Delta E_g(C)$ приведен на рис. 4.10. Достаточно корректная и простая экспериментальная идентификация фактора асимметрии A_{BN} осуществима из оценки контактной разности потенциалов p - n -перехода φ_0 при измерениях обратносмещенных барьерных емкостей. Наиболее простым из иерархического набора уравнений (4.106) может служить общепринятое допущение $A_{BN}=0.5$ ($\Delta E_{gn}=\Delta E_{gp}$). Максимальное упрощение расчетных формул в предположении справедливости статистики Больцмана может быть проэкстраполировано на реальные сильнолегированные структуры. Использование упрощенной записи уравнений для плотностей тока носителей

$$J_n = nq\mu_n [\nabla (V + \Delta E_g/2q)] + kT\mu_n \nabla n, \quad (4.107)$$

$$J_p = -pq\mu_p [\nabla (V - \Delta E_g/2q)] + kT\mu_p \nabla p \quad (4.108)$$

с экспериментально идентифицированной величиной сужения запрещенной зоны ΔE_g , учитывающей влияние статистики Ферми—Дирака (вырождения) и квазибольцмановской статистики:

$$n = n_{ie} \exp [(V - \varphi_n)/\varphi_T], \quad (4.109)$$

$$p = n_{ie} \exp [(V - \varphi_p)/\varphi_T], \quad (4.110)$$

является достаточно корректным и обобщенным фактором описания ЭСЛ в ФТ-моделях полупроводниковых структур.

Таким образом, использование в ФСУ полупроводника соотношений для плотностей тока (4.103), (4.104) и статистики носителей для сильнолегированного полупроводника (4.109), (4.110) позволяет синтезировать ФТ-модель, удобную для численного расчета БТ с помощью соответствующих одномерных и двумерных программ на ЭВМ. Подчеркнем несомненную актуальность синтеза подобной ФТ-модели полупроводниковых структур, достаточно строго учитывающей искажения зонной структуры в сильнолегированных полупроводниках, совокупность ЭСЛ, модифицирующих механизмы переноса и комбинации носителей заряда, но по сложности и вычислительным затратам незначительно отличающейся от общепринятых, «стандартных» ФТ-моделей, использующих статистику Больцмана. Существенной методологической основой корректности рассматриваемой феноменологической ФТ-модели является надежная идентификация зависимостей основных электрофизических параметров модели ΔE_g , A_{BN} от концентрации и температуры. Другим подходом синтеза ФТ-модели является включение в ФСУ (4.1)—(4.5) дополнительной системы уравнений ЭСЛ с непосредственным использованием статистики Ферми—Дирака. При этом, например, параметры ΔE_g , A_{BN} , в свою очередь, становятся функциями базисных модельных переменных зонной структуры полупроводника. Однако, во-первых, корректное описание, идентификация фундаментальных переменных зонной структуры полупроводника весьма затруднительны; во-вторых, сама ФСУ становится чрезвычайно сложной и, несмотря на вынужденное ее упроще-

ние, существенные вычислительные затраты падают на долю решений дополнительной системы уравнений ЭСЛ, зачастую сравнимых по затратам с решениями ФСУ без ЭСЛ. С практической точки зрения полезна возможность точного численного моделирования БТ с несильнолегированным p - n -переходом, $\ll (1 \div 3) \cdot 10^{18} \text{ см}^{-3}$ в рамках «квазибольцмановской» статистики с единственным экспериментально идентифицированным параметром $-\Delta E_g$ из электрических измерений согласно модели искажений электрических зон. Для БТ с произвольным сильным легированием необходимо включение в модель концентрационных зависимостей ΔE_g и A_{BN} . Выводы относительно не столь сильного влияния фактора асимметрии A_{BN} (а также обобщенного коэффициента λ) при ФТ-расчете электрических характеристик БТ согласуются с многочисленными численными экспериментами для полупроводниковых приборов. Вероятна возможность корректного использования упрощенной записи ФСУ в ФТ-модели, где ЭСЛ учитываются в одном эффективном параметре n_{ie} через ΔE_g из электрических измерений в квазибольцмановском описании статистики и уравнений плотностей тока для получения высокой точности расчета полупроводниковых структур с произвольным легированием. Причем в общем случае параметр ΔE_g в ФТ-моделях может быть использован в качестве подстроечного, обеспечивающего необходимую адекватность расчета электрических характеристик приборов.

4.7. МОДЕЛИ ПРОЦЕССОВ РЕКОМБИНАЦИИ — ГЕНЕРАЦИИ

Известно, что, когда физическая система полупроводника отклоняется от состояния термодинамического равновесия ($pn \neq n_{ie}^2$), проявляются кинетические механизмы рекомбинации — генерации, посредством которых физическая система стремится к восстановлению нарушенного равновесия ($pn = n_{ie}^2$). На рис. 4.12 показаны схематически основные процессы рекомбинации — генерации, ответственные за кинетику восстановления термодинамического равновесия с соответствующими равновесными концентрациями носителей заряда.

На рис. 4.12, б показан механизм рекомбинации через дискретный энергетический уровень ловушки, расположенный в запрещенной зоне полупроводника, а на рис. 4.12, в — через два таких уровня. В механизмах рекомбинации через дискретный уровень ловушковой захватывается электрон, а затем дырка. На рис. 4.12, а проиллюстрирован механизм рекомбинации зона — зона, когда происходит непосредственная рекомбинация электрона с дыркой. В рамках механизма рекомбинации зона — зона закон сохранения энергии обеспечивается либо за счет испускания кванта света (излучательная рекомбинация), либо за счет передачи избыточной энергии взаимодействия другому свободному электрону или дырке (Оже-рекомбинация).

Рассмотрим механизм рекомбинации — генерации Шокли — Рида — Холла (ШРХ) через дискретный уровень ловушки, включающий четыре отдельных процесса: а) захват электрона — электрон из зоны проводимости захватывается незанятым уровнем ловушки, который становится занятым; б) захват дырки — электрон из занятого уровня ловушки перемещается в валентную зону и там нейтрализует дырку. «Ловушечный» уровень опять становится незанятым; в) эмиссия дырки — электрон из валентной зоны захватывается уровнем ловушки, дырка попадает в валентную зону и занимает ловушку; г) эмиссия электрона — электрон из занятого уровня ловушки перемещается в зону проводимости. «Ловушечный» уровень становится незанятым.

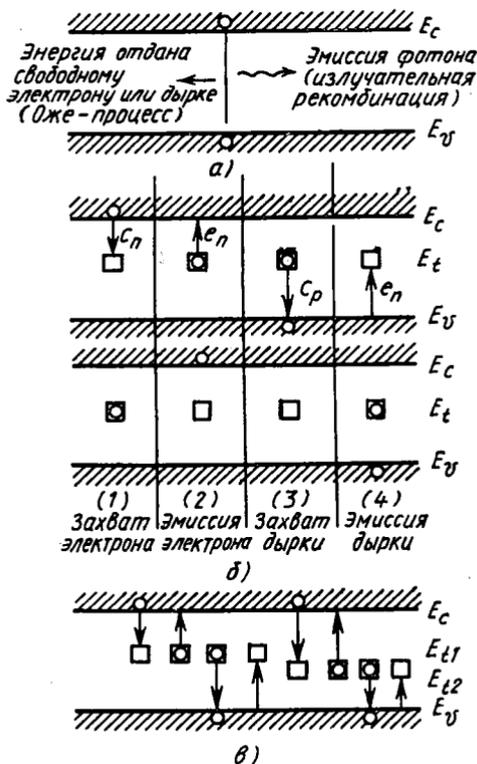


Рис. 4.12. Основные рекомбинационные процессы: а — Оже-рекомбинация зона — зона; б — ШРХ-рекомбинация через моноэнергетический уровень; в — рекомбинация с участием двух уровней

Роль ловушек захвата могут играть атомы доноров и акцепторов. Характерной особенностью ловушек захвата является то, что они взаимодействуют в основном только с одной зоной — проводимости или валентной. Перейдем к описанию механизма ШРХ — рекомбинации через ловушки с энергией E_t и концентрацией C_t . Поскольку полупроводник находится в неравновесном состоянии, распределение электронов и дырок по энергетическим состояниям и локальным уровням E_t характеризуется квазиуровнями Ферми F_n^* , F_p^* и F_t^* . Функция распределения f_t описывает часть занятых ловушек.

Скорости захвата в единице объема для электронов и дырок могут быть описаны выражениями

$$R_n = C_{cn} n (1 - f_t), \quad (4.111)$$

$$R_p = C_{cp} p f_t, \quad (4.112)$$

где C_{cn} , C_{cp} — скорости захвата электронов и дырок на электрон и дырку соответственно, когда все ловушки не заняты.

Аналогично скорости эмиссии в единице объема для электронов и дырок могут быть представлены равенствами

$$G_n = C_{en} f_t, \quad (4.113)$$

$$G_p = C_{ep} (1 - f_t), \quad (4.114)$$

где C_{en} , C_{ep} — скорости эмиссии электронов и дырок на электрон и дырку соответственно.

Полные скорости рекомбинации — генерации дырок и электронов равны, поскольку дырки и электроны всегда взаимодействуют парами:

$$R = R_n - G_n = R_p - G_p. \quad (4.115)$$

Соотношение (4.115), справедливое для статики, нарушается в ходе переходных процессов, когда происходит резкое изменение концентраций подвижных носителей заряда. В состоянии термодинамического равновесия скорости захвата электронов и дырок равны соответствующим скоростям эмиссии, что дает возможность выразить скорости эмиссии электронов и дырок таким образом:

$$G_{en} = C_{cn} n_0 \frac{1 - f_{t0}}{f_{t0}}, \quad (4.116)$$

$$C_{ep} = C_{cp} p_0 \frac{f_{t0}}{1 - f_{t0}}. \quad (4.117)$$

Здесь n_0 , p_0 , f_{t0} — значения переменных в состоянии термодинамического равновесия.

Согласно выражениям (4.115) — (4.117),

$$R = C_{cn} (n(1 - f_t) - n_1 f_t) = C_{cp} (p f_t - p_1 (1 - f_t)), \quad (4.118)$$

откуда следует

$$R = \frac{C_{cn} C_{cp} (np - n_1 p_1)}{C_{cn} (a + n_1) + C_{cp} (p + p_1)}. \quad (4.119)$$

Так как $np = n_{ie}^2 \exp \frac{E_{Fn}^* - E_{Fp}^*}{kT}$, то при совпадении энергий уровней Ферми с уровнем ловушки $E_t = E_{Fp}$, $n = n_1$ и при $E_t = E_{Fp}$, $p = p_1$, $E_t = 0,5(E_v - E_c)$

$$n_1 p_1 = n_{ie}^2, \quad (4.120)$$

скорость ШРХ-рекомбинации (4.119) с учетом (4.120) записывается как

$$R_{\text{ШРХ}} = \frac{np - n_{ie}^2}{\tau_n (n + n_{ie}) + \tau_p (p + n_{ie})}, \quad (4.121)$$

где $\tau_n = 1/C_{cn}$, $\tau_p = 1/C_{cp}$ — соответственно времена жизни электронов и дырок, зависящие от концентрации и температуры; $C_{cn} = \sigma_n v_t C$, $C_{cp} = \sigma_p v_t C$.

Рассмотрим механизмы Оже-рекомбинации — генерации, характерные при больших концентрациях полупроводника: а) захват электрона — электрон из зоны проводимости перемещается в валентную зону, передавая избыточную энергию другому электрону в зоне проводимости. Электрон рекомбинирует с дыркой; б) захват дырки — электрон из зоны проводимости перемещается в валентную зону, передавая избыточную энергию другой дырке в валентной зоне, которая удаляется от края валентной зоны. Электрон рекомбинирует с дыркой; в) эмиссия электрона — электрон из валентной зоны перемещается в зону проводимости за счет потребления энергии от высокоэнергетичного электрона в зоне проводимости; г) эмиссия дырки — электрон из валентной зоны перемещается в зону проводимости за счет потребления энергии от высокоэнергетичной дырки в валентной зоне.

Скорости Оже-рекомбинации — генерации для четырех рассмотренных процессов выразим следующим образом:

$$R_n^{\text{Оже}} = A_n n^2 p, \quad R_p^{\text{Оже}} = A_p n p^2, \quad (4.122)$$

$$G_n^{\text{Оже}} = C_{en}^{\text{Оже}} n, \quad G_p^{\text{Оже}} = C_{ep}^{\text{Оже}} p, \quad (4.123)$$

где A_n , A_p , $C_{en}^{\text{Оже}}$, $C_{ep}^{\text{Оже}}$ — коэффициенты Оже-рекомбинации — генерации.

Поскольку в состоянии термодинамического равновесия не существует актов рекомбинации — генерации, можно найти связь коэффициентов Оже-рекомбинации и генерации:

$$C_{en}^{\text{Оже}} = A_n n_{ie}^2, \quad C_{ep}^{\text{Оже}} = A_p n_{ie}^2. \quad (4.124)$$

Согласно соотношениям (4.122) — (4.124) полная скорость Оже-рекомбинации — генерации

$$R^{\text{Оже}} = R_n^{\text{Оже}} - G_n^{\text{Оже}} + R_p^{\text{Оже}} - G_p^{\text{Оже}} = (A_n n + A_p p)(np - n_{ie}^2). \quad (4.125)$$

Интегральная скорость рекомбинации с учетом механизма ШРХ и Оже-механизмов (4.121), (4.125) записывается в виде

$$R^{\text{Оже, ШРХ}} = (np - n_{ie}^2) \{A_n n + A_p p + 1/[\tau_p(p + n_{ie}) + \tau_n(n + n_{ie})]\}. \quad (4.126)$$

Использование формулы интегральной скорости рекомбинации (4.126) в ФСУ (4.2), (4.3) позволяет модифицировать уравнения непрерывности за счет описания рекомбинационных процессов с учетом ЭСЛ в полупроводнике [при соответствующем описании эффективной собственной концентрации n_{ie} (4.60)].

Наряду с рекомбинацией — генерацией в объеме полупроводника электроны и дырки в приборе могут рекомбинировать — генерировать на его поверхности. При этом скорость рекомбинации —

генерации на поверхности может значительно (особенно при недостаточной культуре технологического производства) превышать объемные характеристики процесса. Для численного моделирования полупроводниковых структур вполне удовлетворительно описание поверхностных процессов граничным условиям для Si—SiO₂ (4.43).

В выражении полной скорости генерации — рекомбинации полупроводника

$$R_{\Sigma} = R^{Oжe} + R^{ШРХ} + G_{Av} = R^{Oжe, ШРХ} + G_{Av}$$

скорость лавинной генерации задается соотношением

$$G_{Av} = L_n |J_n| + L_p |J_p|,$$

причем коэффициенты лавинной ионизации носителей заряда экспоненциально зависят от напряженности поля E:

$$L = L_{00} \exp\left(-\frac{b}{|E|}\right),$$

где L_{00} , b — экспериментально определенные параметры для электронов и дырок.

Однако обычно в численных расчетах полупроводниковых структур учитывать генерационную составляющую в выражении полной скорости рекомбинации — генерации нецелесообразно, поскольку для большинства структур характерны невысокие рабочие напряжения и отсутствие областей с большими электрическими полями.

4.8. МОДЕЛИ ЭЛЕКТРОФИЗИЧЕСКИХ ПАРАМЕТРОВ ДЛЯ ФИЗИКО-ТОПОЛОГИЧЕСКОГО РАСЧЕТА ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

В ФСУ (4.1) — (4.5) входят электрофизические параметры (ЭФП), которые сами по себе являются интегральным результатом сложных и недостаточно изученных для сильнолегированного кремния физических механизмов. Это подвижности дырок и электронов μ_p , μ_n , входящие в уравнения переноса (4.4), (4.5), (4.103), (4.104); параметры, входящие в скорость рекомбинации; времена жизни τ_p , τ_n , Оже-коэффициенты A_p , A_n (4.126), скорость поверхностной рекомбинации $s_{n,p}$ (4.43), а также параметры ΔE_g и A_{BN} , входящие непосредственно в уравнения переноса (4.103), (4.104) и через n_{ie} в уравнения (4.2), (4.3), (4.126). Параметры $\mu_{n,p}$, $\tau_{n,p}$, $A_{n,p}$, ΔE_g , A_{BN} в общем случае не постоянны, зависят от локальных значений концентраций примесей и подвижных носителей, напряженности электрических полей в структуре и плотности рабочего тока. Суть теоретико-экспериментальной идентификации указанных параметров заключается в установлении физических механизмов, закладываемых в электрофизическую модель параметра, и далее в обработке экспериментальных данных, полученных на специальных тестовых или приборных структурах с целью вывода эмпирических или

полуэмпирических соотношений, описывающих физические механизмы. ЭФП можно подразделить на две группы: 1) «квазифундаментальные» параметры, которые практически не зависят или слабо зависят от реального технологического процесса БТ: $\mu_{n,p}(C_D, C_A, n, p, E)$, $\Delta E_g(C_D, C_A, n, p)$, $A_{n,p}$, A_{BN} ; 2) параметры, зависящие от конкретного технологического процесса: $\tau_{n,p}$, $s_{n,p}$. Можно отметить некоторую условность подобного деления; например, параметры ΔE_g , $\mu_{n,p}$ зависят от искажений и напряжений, возникающих обычно в современной технологии после имплантации, но исчезающих при дальнейшем отжиге. При более глубоком изучении физико-технологических механизмов в полупроводнике, вероятно, всем ЭФП в ФТ-моделях можно придать фундаментальный «справочный» исчерпывающий характер, однако актуальной на современном этапе синтеза моделей и идентификации ЭФП для конкретных технологических процессов несомненна.

Модель подвижности основных носителей заряда. Подвижность основных носителей заряда определяется несколькими механизмами рассеяния. Относительно высокая подвижность слаболегированных областей возникает вследствие рассеяния на фонах и дефектах и несколько снижается из-за дополнительного кулоновского рассеяния на атомах ионизированной примеси. При небольших концентрациях легирующих примесей широко используют эмпирическое выражение, хорошо согласуемое с большинством экспериментальных данных по подвижности носителей:

$$\mu_{n,p}(C) = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + (C_T/C_D)^\alpha} \quad (4.127)$$

Здесь $C_T = C_A + C_D$ — полная концентрация легирующих примесей, причем для электронов $\mu_{\max} = 1360 \text{ см}^2 \cdot \text{В}^{-1} \cdot \text{С}^{-1}$, $\mu_{\min} = 92 \text{ см}^2 \cdot \text{В}^{-1} \cdot \text{С}^{-1}$, $C_0 = 1.3 \cdot 10^{17} \text{ см}^{-3}$, $\alpha = 0.91$, для дырок соответственно 520, 65, $2.4 \cdot 10^{17}$, 0.61.

Однако при высоких концентрациях носителей заряда из-за эффекта дополнительного рассеяния на подвижных носителях происходит некоторое снижение подвижности, примерно на 20—40% при плотностях тока $J \geq 3 \cdot 10^3 \text{ А/см}^2$.

Эффект рассеяния на подвижных носителях заряда можно ввести в соотношение (4.127). Такой подход обосновывается тем, что снижение подвижности от рассеяния на ионизированных примесях и подвижных носителях обусловлено механизмом кулоновского взаимодействия. Например, результаты экспериментов для значительных концентраций легирующих примесей в большинстве случаев описываются выражением (4.127), где $C_T = 0.34(C_A + C_D) + 0.66(n + p)$, с использованием приведенных выше коэффициентов.

Зависимость подвижности от напряженности электростатического поля описывается соотношением

$$\mu_{ст} = \frac{\mu_0}{[1 + (E/\mu_0 v_a)^\beta]^{1/\beta}}, \quad (4.128)$$

где μ_0 — подвижность при слабых полях; $\beta_n=2$; $\beta_p=1$; $v_\alpha = v_{\alpha 0} / [1 + 0.8 \exp(T(K)/600)]$, причем максимальные скорости носителей заряда $v_{\alpha 0} = 1.1 \cdot 10^7$ см/с, $v_{\alpha 0}(p) = 9.5 \cdot 10^6$ см/с не зависят от концентрации примесей.

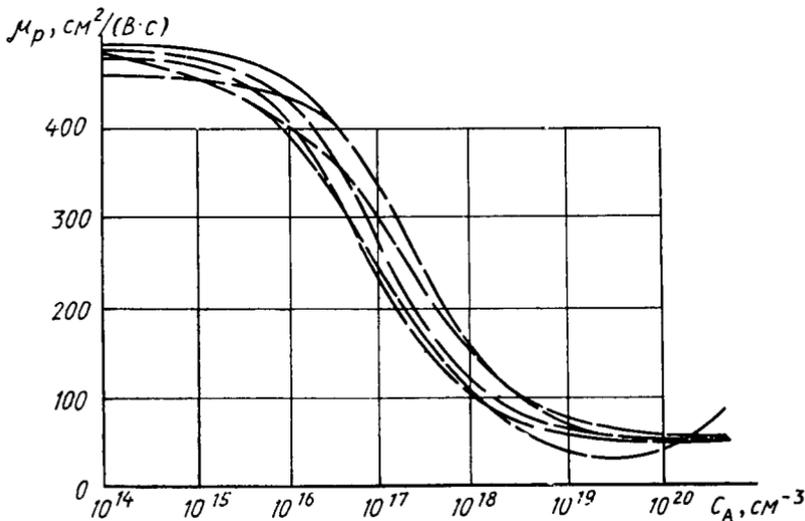


Рис. 4.13. Зависимости подвижности дырок от концентрации акцепторной примеси для ряда моделей электрофизических параметров

Для ФТ-расчета полупроводниковых структур с произвольными уровнями легирования в различных температурных режимах использовалось обобщенное соотношение для подвижности основных носителей с учетом рассеяния на ионизированной примеси, подвижных носителях заряда и полевого члена:

$$\mu^{-1} = \mu_{ct}^{-1} + \mu_{cs}^{-1} + (E/v_s)^3, \quad (4.129)$$

$$\mu_{ct} = \mu_1 T_n^{-0.57} + \frac{\mu_2 T_n^{-2.33}}{1 + [C_0 T_n^{-2.4} (C_D + C_A)]^{T_n^{-0.146}}}, \quad (4.130)$$

$$\mu_{cs} = [(pn)^{1/2} / (1.43 \cdot 10^{20})] \ln(1 + 4.5 \cdot 10^{11} (pn)^{-1/3}), \quad T_n = T(K)/300. \quad (4.131)$$

В выражениях (4.129)—(4.131) использованы параметры μ_1 , μ_2 , C_0 (см⁻³), α , β , v_s , которые имеют следующие значения: для электронов 88, 1268, $7.92 \cdot 10^{-18}$, 0.88, 1/2, $1.1 \cdot 10^7$, для дырок соответственно 54.3, 405, $427 \cdot 10^{-18}$, 0.88, 1, $1.03 \cdot 10^7$.

На рис. 4.13, 4.14 для сравнения приведены зависимости подвижности основных носителей заряда от соответствующих концентраций примесей, обнаруживающие примерно одинаковый ход кривых

(особенно для μ_n) в широком диапазоне концентрации для $T = 293$ К.

В большинстве ФТ-расчетов до недавнего времени в связи с отсутствием надежных экспериментальных данных полагалось примерное равенство подвижностей основных и неосновных носителей заряда.

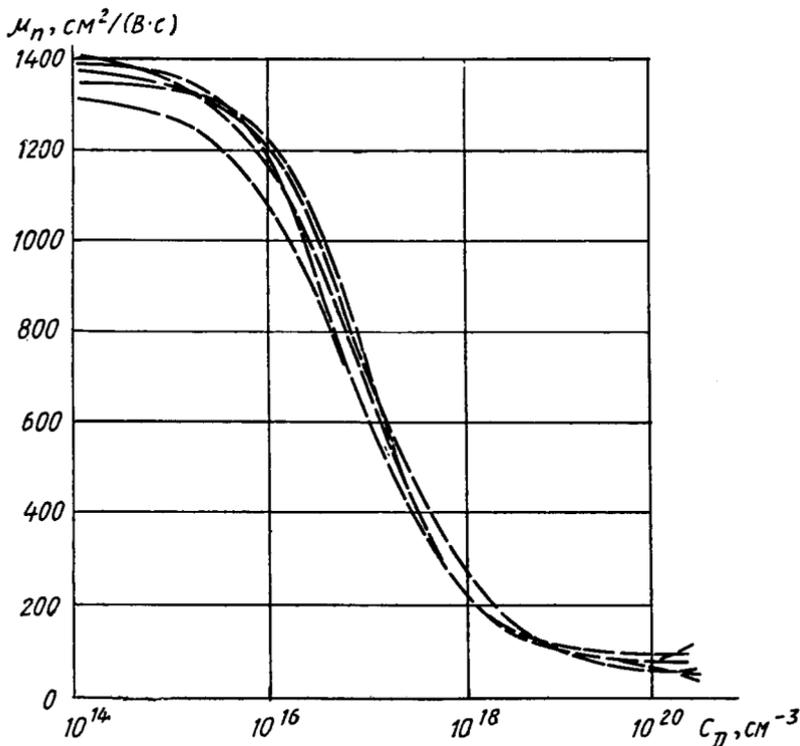


Рис. 4.14. Зависимости подвижности электронов от концентрации донорной примеси для ряда моделей электрофизических параметров

Модель подвижности неосновных носителей заряда. На рис. 4.15 представлены экспериментальные зависимости подвижности неосновных носителей заряда (дырок) и показано, что в области средней концентрации $10^{15} \leq C_D \leq 10^{16} \text{ cm}^{-3}$ она может быть несколько выше подвижности основных носителей, а в области очень высокой концентрации $C_D \geq 10^{20} \text{ cm}^{-3}$ — существенно ниже. Первое отличие объясняется эффектом экранирования основных носителей заряда и неодинаковым рассеянием носителей на примесных атомах. При ФТ-моделировании полупроводниковых структур обобщенным выражением для подвижности носителей (4.129) в члене рассеяния на ионизированной примеси можно подстроить (несколько уменьшить) коэффициенты μ_1, μ_2 с целью совпадения модельных

и экспериментальных зависимостей подвижностей неосновных носителей заряда.

Для объяснения отличия подвижностей основных и неосновных носителей заряда в полупроводнике с высокой концентрацией развита феноменологическая модель.

Суть уменьшения $\mu_p(C_D)$ при высоких концентрациях C_D заключается в: 1) увеличении эффективной массы m_p^* в условиях переноса дырок в сильнолегированном кремнии, более похожем на аморфный, нежели на кристаллический полупроводник; 2) уменьшении характеристического времени рассеяния за счет возрастания роли «ловушечных» центров в сильнолегированном кремнии, когда «хвостовые эффекты» функции распределения приводят в разупорядоченной кристаллической решетке к высокой локализации «ловушечных» центров, инициирующих интенсивный захват неосновных носителей заряда (рис. 4.16). Модель выделяет на первый план влияние ЭСЛ на движение неосновных носителей заряда при практически пренебрежимом влиянии ЭСЛ на процесс переноса заряда основными носителями.

Для данной концентрации $C_D = n$ в n -области полупроводника плотность захваченных «ловушечными» центрами дырок p_t относительно плотности дырок p определяется смещением $\Delta E_{v, bt}$ (рис. 4.16). Оценку эффективной подвижности неосновных носителей заряда (дырок) проведем следующим образом. Плотность дырочного тока для сильного легирования полупроводника описывается обобщенным соотношением (4.83) с включением эффективной подвижности дырок $\mu_{p, \text{эф}}$:

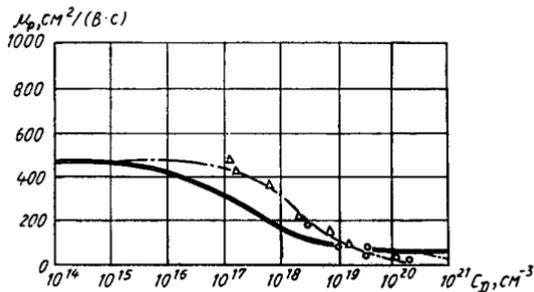


Рис. 4.15. Зависимости подвижности дырок от концентрации акцепторов (сплошная кривая) и доноров (штрихпунктир)

характеристического времени рассеяния за счет возрастания роли «ловушечных» центров в сильнолегированном кремнии, когда «хвостовые эффекты» функции распределения приводят в

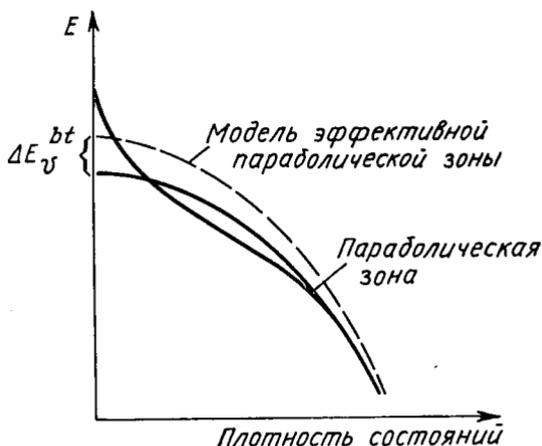


Рис. 4.16. Образование «хвостов» валентной зоны в сильнолегированном кремнии для описания зависимости захваченных «ловушечными» центрами дырок

$$J_p = p \mu_{\text{эф}}^p \nabla E_{Fp}, \quad (4.132)$$

где $\mu_{\text{эф}}^p$ определяется из выражения

$$p \mu_{\text{эф}}^p = p_i \mu_{pt} + (p - p_i) \mu_p(C_D), \quad (4.133)$$

в котором учитываются поток захватываемых дырок с малым коэффициентом подвижности μ_t и поток незахватываемых дырок с коэффициентом подвижности, идентичным для основных и неосновных носителей заряда $\mu_p(C_D) = \mu_n(C_D)$.

Для параболической зонной структуры неосновные носители заряда (дырки) невырождены и описываются с помощью статистики Больцмана:

$$\begin{aligned} p_t &= N_v \exp \left[\frac{E_{v\text{эф}} - E_{Fp}}{kT} \right] - N_v \exp \left[\frac{E_{v\text{эф}} - \Delta E_v^{bt} - E_{Fp}}{kT} \right] = \\ &= p \left[1 - \exp \left(-\frac{\Delta E_v^{bt}}{kT} \right) \right]. \end{aligned} \quad (4.134)$$

Подстановка выражения (4.134) в (4.133) дает возможность оценить $\mu_{\text{эф}}^p$:

$$\mu_{\text{эф}}^p = \left[1 - \exp \left(-\frac{\Delta E_v^{bt}}{kT} \right) \right] \mu_{pt} + \exp \left(-\frac{\Delta E_v^{bt}}{kT} \right) \mu_n(C_D). \quad (4.135)$$

При $C_D < 3 \cdot 10^{20} \text{ см}^{-3}$ $\mu_{\text{эф}}^p \gg \mu_{pt} \sim 10^{-2} \text{ см}^2/(\text{В} \cdot \text{с})$. При $T = 300 \text{ К}$ формула (4.135) упрощается и отношение коэффициентов подвижности неосновных и основных носителей заряда записывается в виде

$$\frac{\mu_{\text{эф}}^p}{\mu_n(C_D)} \approx \exp \left(-\frac{\Delta E_v}{kT} \right) \mu_n(C_D). \quad (4.136)$$

Диапазон действия данной феноменологической модели составляет $5 \cdot 10^{19} - 3 \cdot 10^{20} \text{ см}^{-3}$. Модель позволяет не только корректно связать коэффициенты подвижности основных и неосновных носителей заряда с помощью соотношений (4.129) — (4.131), (4.136), но и получить однозначное толкование результатов экспериментальной идентификации параметров $\Delta E_g^{\text{он}}$, $\Delta E_g^{\text{эл}}$, $\mu_p(C_D)$, $\mu_n(C_D)$, $\mu_{\text{эф}}^p$ в сильнолегированном полупроводнике. Феноменологическая модель позволяет также корректно привести в соответствие соотношения для коэффициентов диффузии основных и неосновных носителей $D_{n,p} = \frac{kT}{q} \mu_{n,p}$ и, кроме того, уточнить времена жизни $\tau_{n,p} \sim \sim L^2/D_{n,p}$ для экспериментов в сильнолегированном кремнии, основанных на измерении диффузионной длины L неосновных носителей заряда.

Модель сужения «электрической» ширины запрещенной зоны кремния. Для расчета электрических характеристик БТ в дальнейшем будут использованы значения «электрической» ширины запрещенной зоны $\Delta E_g^{\text{эл}}$ (просто ΔE_g), что отражено в рассмотренной

модели искажений энергетической зоны. Поскольку в различных экспериментах получен довольно значительный разброс величин ΔE_g , кроме инструментальных погрешностей важно выявить принципиальные причины физико-технологического разброса. Влияние разных статистик, применяемых при обработке экспериментальных данных ΔE_g , отражено в формуле (4.69) с соответствующим пересчетом значений ΔE_g , экспериментально определенных в электрических измерениях. К другим вопросам коррекции ΔE_g , вытекающим из условий опыта, следует отнести технологический фактор, связанный с искажением кристаллической решетки при сильном легировании полупроводника. Имеется два физико-технологических эффекта, связанных с искажениями кристаллической решетки при сильном легировании: перенасыщение легирующими атомами и нарушение, вызванное имплантацией примеси. Если напряжение решетки из-за перенасыщения легирующими атомами снимается в основном соответствующим режимом отжига, то искажение решетки после диффузии из ионно-имплантированных слоев не снимается стандартными процессами отжига. При этом сужение ширины запрещенной зоны в сильнолегированных искаженных решетках кремния ($C_{A,D} \geq 10^{20} \text{ см}^{-3}$) соответствовало величине $D_{\text{max}} = 32 \text{ мэВ}$. Использование величины $D(C_{A,D}, T)$, связанной с нерегулярностями в распределении результирующей примеси при соответствующих технологиях формирования сильнолегированных областей (без имплантации $D=0$, с имплантацией $20 \leq D \leq 32 \text{ мэВ}$), позволяет трансформировать выражение (4.62) в более общее:

$$\Delta E_g = \Delta E_g^{3\pi} + S - D(C, T), \quad (4.137)$$

где S определяется статистикой носителей, а D — рассмотренными нерегулярностями; величина D является функцией дозы имплантации.

Отметим, что составляющие S и D в (4.137) носят компенсирующий характер, однако вследствие значительной разности $S - D \approx \approx 10 \div 40 \text{ мэВ}$ (см. рис. 4.10) ими нельзя пренебрегать для выявления сути разброса имеющихся экспериментальных данных.

Для того чтобы установить некоторую аппроксимирующую функцию $\Delta E_g(C_{A,D})$, сделаем ряд замечаний. Наиболее достоверными результатами экспериментальной идентификации ΔE_g являются данные на структурах с тонкими диффузионными слоями, для которых наблюдается хорошее согласование с экспериментом (см. рис. 4.10); кроме того, для большинства из них имеется согласование времен жизни носителей заряда. Таким образом, эмпирическая аппроксимация ΔE_g

$$\Delta E_g = q(A + (A^2 + 0,5)^{1/2}) \text{ мэВ}, \quad (4.138)$$

где

$$A = \begin{cases} \ln p/10^{17}, & \text{Si } p\text{-типа,} \\ \ln n/1,5 \cdot 10^{17}, & \text{Si } n\text{-типа,} \end{cases}$$

скорректированная с учетом выражения (4.137), на сегодня отражает достигнутые результаты в экспериментальной идентификации важнейшего ЭФП.

Зависимости времен жизни электронов и дырок. Зависимости $\tau_{n,p}$, определенные разнообразными методами для широкого диапазона концентраций примеси, приведены на рис. 4.17. Так как интегральная скорость рекомбинации достаточно хорошо аппроксимируется двумя типами рекомбинации (ШРХ и Оже), на зависимости $\tau_p(C_D)$ можно выделить две модельные характеристики: $\tau_p \sim C_D^{-1}$ для $C_D < (1 \div 2) \cdot 10^{18} \text{ см}^{-3}$ и $\tau_p \sim C_D^{-2}$ для $C_D \geq (1 \div 2) \cdot 10^{18} \text{ см}^{-3}$, описываемые указанными механизмами рекомбинации. Для описания зависимости $\tau_p \sim C_D^{-\alpha}$, $\alpha = 0.3 \div 0.6$ предложена феноменологическая модель единого дефектного уровня в невырожденном кремнии, где определение плотности дефектов основано на законе действующих масс в состоянии термодинамического равновесия с преобладающей температурой образования дефектов. Данная модель описывается полуэмпирическими соотношениями

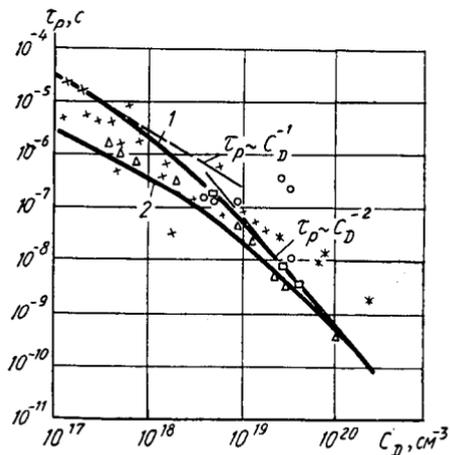


Рис. 4.17. Модельные (кривые 1, 2) и экспериментальные зависимости времен жизни дырок от концентрации донорной примеси ($T=300 \text{ К}$). Для кривой 2 плотность дефектов увеличена на порядок

концентрации примеси, приведены на рис. 4.17. Так как интегральная скорость рекомбинации достаточно хорошо аппроксимируется двумя типами рекомбинации (ШРХ и Оже), на зависимости $\tau_p(C_D)$ можно выделить две модельные характеристики: $\tau_p \sim C_D^{-1}$ для $C_D < (1 \div 2) \cdot 10^{18} \text{ см}^{-3}$ и $\tau_p \sim C_D^{-2}$ для $C_D \geq (1 \div 2) \cdot 10^{18} \text{ см}^{-3}$, описываемые указанными механизмами рекомбинации. Для описания зависимости $\tau_p \sim C_D^{-\alpha}$, $\alpha = 0.3 \div 0.6$ предложена феноменологическая модель единого дефектного уровня в невырожденном кремнии, где определение плотности дефектов основано на законе действующих масс в состоянии термодинамического равновесия с преобладающей температурой образования дефектов. Данная модель описывается полуэмпирическими соотношениями

где τ_{\max} , τ_{\min} — характеристические времена жизни; $\tau_{n,p}^0$ — времена жизни электронов и дырок, измеренные при концентрации C_r , для которых получено корректное теоретическое обоснование.

$$\tau_{n,p} = \tau_{\min} + \frac{\tau_{\max} - \tau_{\min}}{[1 + C_r/(C_A + C_D)]^\alpha}, \quad \tau_{n,p} = \frac{\tau_{n,p}^0}{[1 + C_r/(C_A + C_D)]^\alpha}, \quad (4.139)$$

Из рис. 4.17 видно, что значения коэффициентов $\alpha = 0.3 \div 0.5$ хорошо согласуются с экспериментальными данными $C < (1 \div 2) \times 10^{18} \text{ см}^{-3}$. Отметим, что времена жизни при ШРХ-рекомбинации существенно зависят от конкретного типа технологического процесса изготовления БИС, поэтому разброс времен жизни для одной и той же концентрации может возрастать в 5—10 раз при увеличении плотности дефектов на порядок (кривые 1, 2 на рис. 4.17). В то же время разброс значений $\tau_{n,p}$ при разности плотности дефектов на порядок в области высоких концентраций примесей, когда доминирует Оже-рекомбинация, составляет всего 5—30%.

где τ_{\max} , τ_{\min} — характеристические времена жизни; $\tau_{n,p}^0$ — времена жизни электронов и дырок, измеренные при концентрации C_r , для которых получено корректное теоретическое обоснование.

Экспериментально определенные Оже-коэффициенты лежат в довольно узком диапазоне: $A_n = (0.7 \div 3) \cdot 10^{-31} \text{ см}^6/\text{с}$, $A_p = (0.95 \div 1.2) \cdot 10^{-31} \text{ см}^6/\text{с}$; в качестве наиболее достоверных взяты значения $A_n = 2.8 \cdot 10^{-31} \text{ см}^6/\text{с}$, $A_p = 0.99 \cdot 10^{-31} \text{ см}^6/\text{с}$.

Поскольку зависимость $\tau_{n,p}(T)$ определяется в основном температурной зависимостью сечения захвата, справедливо соотношение

$$\tau_{n,p}^0 = \tau_{n,p}(300)(T/300)^{-2}. \quad (4.140)$$

Зависимости Оже-коэффициентов от температуры представляются в виде

$$A_{no}(T) \approx \text{const}, \quad A_{p0}(300)(T/300)^{0.2}. \quad (4.141)$$

4.9. МОДЕЛИ ЭЛЕКТРОФИЗИЧЕСКИХ ПРОЦЕССОВ В ПОЛИКРЕМНИИ

Модель подвижности основных носителей заряда. Одним из основных ЭФП поликремневого слоя является подвижность μ_{Si^*} , имеющая немонотонную зависимость от концентрации легирующей примеси. В модели подвижности основных носителей заряда в поликремнии сделаны следующие допущения: 1) поликремний состоит из идентичных зерен кубической формы, и ток создается практически одномерным переносом основных носителей заряда. Каждая ячейка поликремния с характерной длиной L состоит из кристаллического кубического зерна и разупорядоченной границы; 2) ток термоэмиссии в поликремнии пренебрежимо мал; 3) граница зерна описывается специфической зонной диаграммой с фиксированными шириной и хвостом зоны; 4) кремний в приграничной области является аморфным, и в нем доминируют прыжковый и диффузионный механизмы электропроводности. Модель подвижности основных носителей заряда, представленная системой алгебраических уравнений, дает хорошее согласование с экспериментальными результатами для поликремния p - и n -типов в широком диапазоне температур (рис. 4.18). Для подвижности неосновных носителей заряда в поликремнии на сегодня не создано модели, удовлетворительно описывающей экспериментальные зависимости. В качестве модели начального приближения примем примерное равенство подвижностей неосновных и основных носителей заряда в поликремнии: $\mu_p^{n+} \approx \mu_p^{p+}$.

Модель контакта кремний — поликремний ($\text{Si} - \text{Si}^*$). Модель контакта $\text{Si} - \text{Si}^*$ учитывает его влияние через эффективную скорость рекомбинации (ЭСР) на границе $\text{Si} - \text{Si}^*$ и дополнительное туннельное эмиттерное сопротивление $R_{тэ}$, которое представляет для неосновных носителей заряда сверхтонкий квазиокисный слой SiO между Si^* и Si . Полагается, что в слое Si^* имеется N одинаковых зерен с размером d_g и скорость рекомбинации на контакте Si^* -металл постоянна: $s_m = 10^6 \text{ см/с}$. Линейная зависимость между током и концентрацией в одной и той же точке позволяет оценить ЭСР s_p' на границе $\text{Si} - \text{Si}^*$ без учета SiO_x для оценки влияния

слоя Si* на поведение носителей заряда, инжектированных в эмиттер:

$$s'_p = a_g - \frac{b^2}{a_g + s_{эф}}, \quad (4.142)$$

где $s_{эф}$ — ЭСР на конце области первого зерна, зависящая от числа зерен.

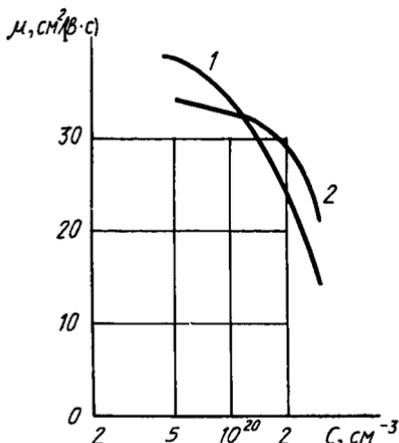


Рис. 4.18. Зависимость подвижности основных носителей заряда от концентрации в поликремнии (кривая 1 — эксперимент, кривая 2 — расчет)

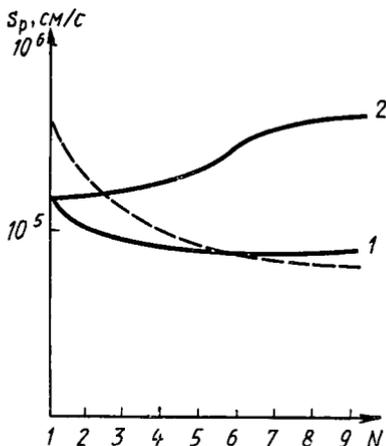


Рис. 4.19. Зависимости эффективной скорости поверхностной рекомбинации от количества зерен в поликремнии

При $N=1$ величина $s_{эф} = s_m$, при $N=2$, $s_{эф} = T_{\partial s} + S_{\partial s} - \frac{T_{\partial s}}{a_g + T_{gb} + S_{gb} - b_g^2(a_g + s_m)}$. Здесь s_{gb} — скорость рекомбинации на границе зерна; параметром T_{gb} , имеющим размерность скорости, обозначена величина $\frac{kT}{q} \frac{\mu_{gb}}{\delta}$; μ_{gb} — подвижность в границе

зерна; $\delta = 10 \text{ \AA}$ — толщина границы зерна; a_g, b_g — коэффициенты, зависящие от d_g , коэффициента диффузии D_p и диффузионной длины L_p . Таким образом, влияние Si*—контакта без учета границы раздела Si—Si* выражается в замене ЭСР s_m на s'_p (4.142).

Расчет s'_p для $N > 2$ и, следовательно, толщин слоя Si $L_{Si} = N(d_g + \delta)$ проводят циклически: при расчете s'_p ($N=3$) $s_{эф}$ полагают равным s'_p ($N=2$), при расчете s'_p ($N=4$) $s_{эф} = s'_p$ ($N=3$) и т. д.

Результаты такого расчета представлены на рис. 4.19. Видно, что s'_p уменьшается при увеличении толщины Si*—контакта и сечения захвата $c_p = 10^{-15} \text{ см}^2$ (кривая 1, c_p входит в выражение для s_{gb}) и, наоборот, увеличивается при $c_p = 10^{-14} \text{ см}^2$ (кривая 2). Расчет сделан для фиксированной толщины слоя $L_{Si*} = 1000 \text{ \AA}$. За-

зависимость $s_p'(N)$ для изменяющейся толщины слоя Si^* при $d_g = 230 \text{ \AA}$ изображена на рисунке пунктиром.

Учет влияния слоя SiO_2 в контакте $\text{Si}-\text{Si}^*$ реализуется расчетом туннельного тока через потенциальный барьер, который представляет для основных и неосновных носителей заряда квазиокисный изолированный слой шириной Δ , который можно создать перед осаждением Si^* на Si .

Таким образом, влияние границы раздела $\text{Si}-\text{Si}^*$ моделируется введением дополнительного сопротивления для основных носителей заряда, рассчитываемого из ВАХ туннельного тока, и изменением эффективной скорости поверхностной рекомбинации для неосновных носителей заряда.

4.10. МОДЕЛИ ПРОЦЕССОВ ТЕРМОГЕНЕРАЦИИ В СИЛЬНОЛЕГИРОВАННЫХ ПОЛУПРОВОДНИКОВЫХ СТРУКТУРАХ

Известен ряд подходов к моделированию процессов термогенерации. В простейшем случае термогенерация моделируется производением вектора напряженности эффективного электрического поля E на сумму векторов плотностей тока $J_n + J_p$ в полупроводниковой структуре:

$$H = E(J_n + J_p). \quad (4.143)$$

Однако простейшее выражение (4.143) не всегда является адекватной оценкой процесса термогенерации для ряда температурных полей. Корректной формой моделирования процесса термогенерации является представление, основанное на экспериментальном факте: полная мощность, рассеиваемая полупроводниковой структурой, эквивалентна Джоулеву нагреванию:

$$P = - \int_S (\varphi_n J_n + \varphi_p J_p) dS, \quad (4.144)$$

где φ_n , φ_p — квазипотенциалы Ферми электронов и дырок, S — полная поверхность структуры.

Главное допущение, сделанное при выводе уравнения (4.144), состоит в том, что все электроны и дырки суммарно вносят вклад в полную мощность рассеивания в структуре. Используя теорему Грина, можно преобразовать интеграл по поверхности (4.144) в интеграл по объему, причем

$$P = - \int_V (\varphi_n \text{div } J_n + \text{grad } \varphi_n J_n + \varphi_p \text{div } J_p + \text{grad } \varphi_p J_p) dV. \quad (4.145)$$

С учетом того, что подинтегральное выражение здесь соответствует термогенерации в единице объема, (4.145) записывается в виде

$$H = -J_n \text{grad } \varphi_n - J_p \text{grad } \varphi_p + qR(\varphi_p - \varphi_n). \quad (4.146)$$

Выражение (4.146) учитывает эффекты сильного легирования и высоких плотностей токов: $J \geq (0.6 \div 2) \cdot 10^5 \text{ A/cm}^2$.

МЕТОДЫ И АЛГОРИТМЫ ЧИСЛЕННОГО ФИЗИКО-ТОПОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

5.1. МЕТОДЫ ДИСКРЕТИЗАЦИИ ФСУ

Численное решение ФСУ полупроводника в двух- и трехмерном приближении для ФТ-моделирования реальной интегральной полупроводниковой структуры представляет собой сложную математическую проблему. Несмотря на большое количество предложенных методов для решения ФСУ и крайне интенсивный темп исследований за последние двадцать с лишним лет начиная с пионерской работы по численному моделированию БТ (Гуммель, 1964 г.), в целом необходимо констатировать, что численное моделирование приборов только сейчас стало стандартной инженерной практикой при исследовании структур с несложной геометрией и является скорее искусством при исследовании структур с реальной геометрией в широком режимном диапазоне. Выбор наилучшего численного метода зависит от типа моделируемых приборных структур с реальной геометрией, рабочих режимов, вида анализа (стационарного, нестационарного и др.), характеристик и предельных вычислительных ресурсов используемых ЭВМ. Несмотря на то что огромное разнообразие применяемых численных методов свидетельствует об отсутствии на сегодня единого универсального, максимально эффективного подхода для всех типов структур и видов анализа, рассмотрим общие положения численного моделирования, основные численные методы дискретизации и решения ФСУ, проведем сравнение численных методов, их возможностей, аспектов выбора, которые существуют при разработке развитого математического обеспечения численного ФТ-моделирования применительно к широкому классу полупроводниковых структур и в первую очередь БТ БИС.

Численное решение ФСУ в частных производных, определяющих работу приборных структур, состоит из двух основных этапов. На первом этапе пространственную область разбивают на сетку из отдельных точек или узлов. Применение какого-либо метода дискретизации к данной краевой задаче сводит ее к алгебраической задаче в пространстве с достаточно большой (но конечной) размерностью. На втором этапе определенными численными методами находят неизвестные переменные. Дискретизированные уравнения для такой задачи выводят из исходной системы ФСУ полупроводниковой структуры с помощью аппроксимации

производных или интегралов, в которых входят значения неизвестных функций только в узлах сетки, что можно сделать используя интерполяционные функции в окрестности узла.

В соответствии с общей процедурой дискретизации возможен ряд вариантов вывода разностных уравнений для краевой задачи. Наиболее широкое распространение получили метод конечных разностей (МКР) и метод конечных элементов (МКЭ).

МКР основан на локальной аппроксимации дифференциально-го оператора некоторым разностным оператором. В МКЭ искомое решение глобально аппроксимируется набором функций формы, которые служат пробными функциями. Для обоих методов дискретизации исходным пунктом служат ФСУ в дифференциальной или интегральной форме либо эквивалентная вариационная задача. Кроме того, обоим методам присуще разбиение всего объема прибора на подобласти и использование пробных функций.

Некоторое представление о взаимодействии МКР и МКЭ дает простой пример их применения к уравнению Пуассона. Из различных МКР выберем наиболее эффективный метод интегральных тождеств, в котором требуется интегральное представление, означающее, что

$$\oint_{\partial v} \nabla V dS = -\frac{1}{\epsilon} \int_v \rho dv \quad (5.1)$$

для любой подобласти с объемом v . Эти подобласти образуются при разбиении всей области на ячейки без перекрытий или промежутков, в результате чего каждая ячейка содержит один узел сетки. В одномерном случае для узла x_i сетки такая ячейка представляет собой интервал $[(x_{i-1} + x_i)/2, (x_i + x_{i+1})/2]$, где x_{i-1} , x_{i+1} — соседние узлы сетки. При этом разностные уравнения выводятся путем аппроксимации интеграла (5.1) по объему ячейки. Принимая распределение потенциала между соседними узлами сетки линейным:

$$V(x) = V_i + (V_{i+1} - V_i) \frac{x - x_i}{x_{i+1} - x_i}, \quad (5.2)$$

получим одномерное разностное уравнение

$$\frac{V_{i+1} - V_i}{x_{i+1} - x_i} - \frac{V_i - V_{i-1}}{x_i - x_{i-1}} = -\frac{1}{\epsilon} \int_{(x_{i-1} + x_i)/2}^{(x_i + x_{i+1})/2} \rho(x) dx. \quad (5.3)$$

Отметим, что окончательный результат зависит еще и от аппроксимации для $\rho(x)$. Считая плотность объемного заряда постоянной, получаем каноническое конечно-разностное представление уравнения Пуассона:

$$\frac{V_{i+1} - V_i}{x_{i+1} - x_i} - \frac{V_i - V_{i-1}}{x_i - x_{i-1}} = -\frac{1}{2\epsilon} (x_{i+1} - x_{i-1}) \rho_i. \quad (5.4)$$

В МКЭ область прибора аналогично разбивают на подобласти. В одномерном случае подобласть или элемент — это интервал

$[x_i, x_{i+1}]$ между двумя соседними узлами сетки. Для рассматриваемого примера такую же локальную аппроксимацию (5.2) дают функции формы вида

$$U_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}; & \forall (x) \in [x_{i-1}, x_i], \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & \forall (x) \in [x_i, x_{i+1}], \\ 0, & \forall (x) \in [x_{i-1}; x_{i+1}], \end{cases} \quad (5.5)$$

которые служат базисом для глобальной аппроксимации неизвестного потенциала

$$V(x) = \sum_i V_i U_i(x). \quad (5.6)$$

Теперь для уравнения Пуассона можно найти уравнение конечных элементов:

$$-\int \nabla U_i \nabla V dv = -\frac{1}{\varepsilon} \int U_i \rho dv, \quad (5.7)$$

где область интегрирования распространяется только на элементы, примыкающие к i -му узлу, поскольку на всех остальных участках U_i и ∇U_i равны нулю.

Вычисляя уравнения конечных элементов (5.7) с помощью одномерных аппроксимаций (5.5), находим

$$\frac{V_{i+1} - V_i}{x_{i+1} - x_i} - \frac{V_i - V_{i-1}}{x_i - x_{i-1}} = -\frac{1}{\varepsilon} \int_{x_{i-1}}^{x_{i+1}} U_i(x) \rho(x) dx. \quad (5.8)$$

Этот результат отличается от разностного уравнения (5.3) лишь способом учета плотности объемного заряда в интегралах. В уравнении (5.3) МКР $\rho(x)$ интегрируется непосредственно по всей области ячейки, тогда как в МКЭ требуется взвешенное интегрирование в большей области,

включающей соседние элементы. Если принять плотность объемного заряда в интервалах интегрирования постоянной, то оба метода (МКР и МКЭ) дадут идентичное разностное уравнение (5.4). Из представленного анализа уравнения Пуассона для одномерного случая можно сделать заключение о том, что при одинаковых аппроксимациях МКР и МКЭ дают практически одинаковые результаты и для многомерного приближения.

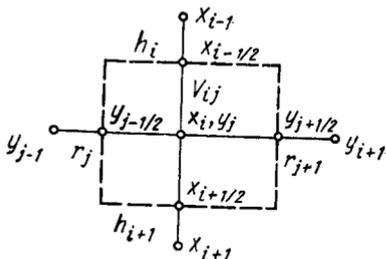


Рис. 5.1. Расположение узлов сетки

Разностная схема для уравнения Пуассона в двумерном приближении. Проиллюстрируем применение интегроинтерполяционного МКР для уравнения Пуассона в двумерном приближении. В области $V = \{(x, y), 0 < x < L_x, 0 < y < L_y\}$ (рис. 5.1) введем неравномерную сетку $\omega = \{(x_i, y_j), 0 \leq i \leq M_1, 0 \leq j \leq M_2\}$, где $x_0 = 0, y_0 = 0, x_{M_1} = L_x, y_{M_2} = L_y$,

$$\begin{aligned} x_{i+1} &= x_i + h_{i+1}, & y_{j+1} &= y_j + r_{j+1}, & i &= 0, 1, \dots, M_1 - 1, \\ & & & & j &= 0, 1, \dots, M_2 - 1; \end{aligned} \quad (5.9)$$

h_i, r_j — соответствующие неравномерные шаги сетки.

Помимо точек множества ω необходимо ввести также потоковые точки $(x_{i+1/2}, y_j), (x_i, y_{j+1/2})$, где

$$\begin{aligned} x_{i+1/2} &= x_i + h_{i+1/2}, & y_{j+1/2} &= y_j + r_{j+1/2}, \\ i &= 0, 1, \dots, M_1 - 1, & j &= 0, 1, \dots, M_2 - 1. \end{aligned} \quad (5.10)$$

Значения функций, определенных на сетке, обозначим следующим образом:

$$U(x_i, y_j) = U_{i,j}, \quad (5.11)$$

$$J(x_{i+1/2}, y_j) = J_{i+1/2,j}, \quad J(x_i, y_{j+1/2}) = J_{i,j+1/2}.$$

Интегрируя уравнение Пуассона по области

$$V_{ij} = \{(x, y), x_{i-1/2} \leq x \leq x_{i+1/2}, y_{j-1/2} \leq y \leq y_{j+1/2}\}, \quad (5.12)$$

$$\int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \Delta V \, dx \, dy = - \int_{x_{i+1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \rho(x, y) \, dx \, dy, \quad (5.12a)$$

$$\rho = p - n + C_D - C_A,$$

получим

$$\begin{aligned} & \int_{y_{j-1/2}}^{y_{j+1/2}} E_x(x_{i+1/2}, y) - E_x(x_{i-1/2}, y) \, dy + \\ & + \int_{x_{i-1/2}}^{x_{i+1/2}} (E_y(x, y_{j+1/2}) - E_y(x, y_{j-1/2})) \, dx = \\ & = \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \rho(x, y) \, dx \, dy, \end{aligned} \quad (5.13)$$

где $E_x(x, y), E_y(x, y)$ — соответственно x - и y -компоненты вектора напряженности электрического поля:

$$E_x(x, y) = - \frac{\partial V(x, y)}{\partial x}, \quad E_y(x, y) = \frac{\partial V(x, y)}{\partial y}. \quad (5.14)$$

Предполагая постоянство E_x , E_y и $\rho(x, y)$ в соответствующих интервалах:

$$\begin{aligned} E_y(x, y_{j+1/2}) &= E_{i,j+1/2}, \quad E_y(x, y_{j-1/2}) = E_{i,j-1/2}, \\ E_x(x_{i+1/2}, y) &= E_{i+1/2,j}, \quad E_x(x_{i-1/2}, y) = E_{i-1/2,j}, \\ \rho(x, y) &= \rho_{ij}, \end{aligned} \quad (5.15)$$

получим

$$\begin{aligned} [(E_x)_{i+1/2,j} - (E_x)_{i-1/2,j}] r_j^* + [(E_y)_{i,j+1/2} - (E_y)_{i,j-1/2}] h_i^* &= \\ &= Q_{ij} h_i^* r_j^*, \end{aligned} \quad (5.16)$$

где $h_i^* = (h_i + h_{i+1})/2$, $r_j^* = (r_j + r_{j+1})/2$.

Выразим теперь $E_{i+1/2,j}$ через значения $V(x, y)$ в узлах сетки. Для этого в первом из равенств (5.14), полагая $y = y_j$ и интегрируя по отрезку $[x_i, x_{i+1}]$

$$\int_{x_i}^{x_{i+1}} E_x(x, y_j) dx = -(V_{i-1,j} - V_{ij}), \quad (5.17)$$

получим с использованием соотношений (5.15)

$$(E_x)_{i+1/2,j} = -\frac{V_{i+1,j} - V_{ij}}{h_{i+1}}. \quad (5.18)$$

Аналогично находим $(E_y)_{i,j+1/2}$:

$$(E_y)_{i,j+1/2} = -\frac{V_{i,j+1} - V_{ij}}{r_{j+1}}. \quad (5.19)$$

Подставляя соотношения (5.18) и (5.19) в (5.16), получим

$$\begin{aligned} (\Delta^h V)_{ij} &= \frac{1}{h_{ij}^*} \left(\frac{V_{i+1,j} - V_{ij}}{h_{i+1}} - \frac{V_{ij} - V_{i-1,j}}{h_i} \right) + \\ &+ \frac{1}{r_j^*} \left(\frac{V_{i,j+1} - V_{ij}}{r_{j+1}} - \frac{V_{ij} - V_{i,j-1}}{r_j} \right) = n_{ij} - p_{ij} - (C_D)_{ij} + (C_A)_{ij}. \end{aligned} \quad (5.20)$$

Поскольку узел сетки (x_i, y_j) выбран произвольно, уравнение (5.20) справедливо во всех внутренних узлах сетки ω , т. е. при $0 < i < M_1$, $0 < j < M_2$. Если сетка равномерна и V — достаточно гладкая функция, то разностный оператор Δ^h аппроксимирует исходный дифференциальный оператор с погрешностью аппроксимации $O(h^2 + r^2)$.

Разностное уравнение типа (5.20) можно получить и другими МКР, МКЭ, применяя иные виды аппроксимаций. Обычно считают, что главное достоинство МКЭ проявляется при работе с треугольными сетками, так как треугольные элементы позволяют

легче выполнить локальные уточнения решения без чрезмерного увеличения общего числа узлов. Во многих случаях разница в количестве узлов сетки может оказаться существенно больше, чем в простейших примерах локального уточнения решения на прямоугольной и треугольной сетках (рис. 5.2). Кроме того, треуголь-

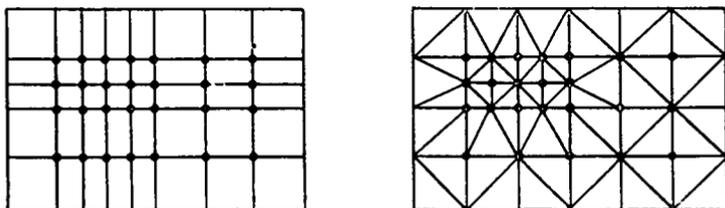


Рис. 5.2. Локальное уточнение решения на прямоугольной и треугольной сетках

ник разбиения позволяет точнее прямоугольного аппроксимировать нелинейные границы областей и линии раздела. С другой стороны, такие задачи можно решать и МКР, вводя в прямоугольники треугольные подобласти или в более общем случае используя непосредственно треугольную сетку. Если в последнем случае прибегнуть к методу интегральных тождеств, то ячейка, связанная с некоторым узлом, будет определяться перпендикулярами к серединам отрезков, соединяющих соседние узлы (рис. 5.3). Обычно же возможности треугольных сеток используются не полностью, поскольку здесь требуется дополнительный расход времени и памяти на построение сетки и хранение промежуточных результатов.

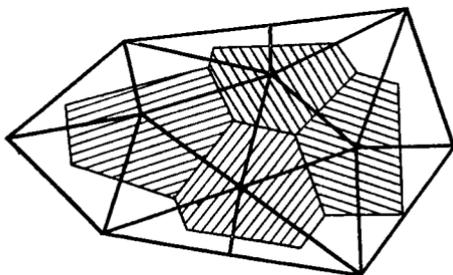


Рис. 5.3. Треугольная сетка и ячейки, построенные для ее узлов

Дискретизация уравнений непрерывности, разностные

аппроксимации уравнений плотности тока. Известно, что при замене соотношений (4.107), (4.108) центрально-разностной аппроксимацией возникают большие вычислительные трудности, связанные с обеспечением устойчивости решения, причем в общем случае для достижения приемлемой точности вычислений требуется более 50—200 узлов разностной сетки на один p - n -переход, вследствие чего концентрации подвижных носителей заряда в нем могут меняться на 5—7 порядков, имеют значительные градиенты и сильно зависят от внешнего напряжения. Следовательно, не менее важную роль, чем выбор «мелкости» сетки под геометрические условия задачи, играет согласование приближенных реше-

ний с физическим смыслом задачи численного расчета полупроводниковых структур, особенно когда свойства численных решений обусловлены присущими структуре внутренними физическими нелинейностями. Иными словами, применяемые пробные функции должны отражать прежде всего физические процессы. Поэтому в качестве пробных функций целесообразно выбирать приближенные аналитические решения исследуемой задачи.

Проанализируем уравнение непрерывности для дырок (4.2). Для стационарных условий с учетом выражения (4.5), выполнения «квазибольцмановской» статистики и эффекта сильного легирования [см. выражения (4.110)] уравнение непрерывности имеет вид

$$\frac{1}{q} \nabla J_p = -\nabla \left\{ \mu_p n_{ie} \left[\exp \left(\frac{\varphi_p - V}{\varphi_T} \right) \right] \nabla \varphi_p \right\} = -R. \quad (5.21)$$

Рассматривая уравнение (3.10) и выражение для квазиуровня Ферми φ_p аналогично выражениям для потенциала (5.1) и (5.2), в окрестности каждого узла сетки получаем экспоненциальные распределения плотности тока дырок, которые представляются неудачными аппроксимациями, поскольку пространственные изменения плотности тока незначительны. Следовательно, для описания поведения функции плотности тока требуется иной подход. Пренебрегая рекомбинацией в пределах одномерного интервала $[x_i, x_{i+1}]$ и интегрируя уравнение (5.21), определяем постоянную плотность тока для этого интервала:

$$\frac{J_{p\ i+1/2}}{q} = - \frac{\exp(\varphi_{pi+1}/\varphi_T) - \exp(\varphi_{pi}/\varphi_T)}{\int_{x_i}^{x_{i+1}} \frac{\exp(V/\varphi_T)}{\varphi_T \mu_p n_{ie}} dx}. \quad (5.22)$$

Поскольку выбираемый интервал $[x_i, x_{i+1}]$ мал, здесь вполне приемлемы допущения о линейном распределении потенциала V (5.2), а также постоянстве произведения подвижности на эффективную собственную концентрацию $(\mu_p n_{ie})_{i+1/2}$. Тогда из выражений (5.21) и (5.22) имеем

$$\begin{aligned} \exp(\varphi_p(x)/\varphi_T) &= \exp(\varphi_{pi}/\varphi_T) + \exp(\varphi_{-i+1}/\varphi_T) - \\ &- \exp\left(\frac{\varphi_{pi}}{\varphi_T}\right) \left[\frac{\exp(V(x)/\varphi_T) - \exp(V_i/\varphi_T)}{\exp(V_{i+1}/\varphi_T) - \exp(V_i/\varphi_T)} \right]. \end{aligned} \quad (5.23)$$

$$\frac{J_{p\ i+1/2}}{q} = (\mu_p n_{ie})_{i+1/2} \frac{(V_{i+1} - V_i)}{(x_{i+1} - x_i)} \left[\frac{\exp(\varphi_{pi+1}/\varphi_T) - \exp(\varphi_{pi}/\varphi_T)}{\exp(V_{i+1}/\varphi_T) - \exp(V_i/\varphi_T)} \right]. \quad (5.24)$$

Подставив приближенное решение для плотности тока (5.24) в формулу (5.21), получим разностное уравнение

$$\frac{1}{q} (J_{p\ i+1/2} - J_{p\ i-1/2}) = \frac{1}{2} (x_{i+1} - x_{i-1}) R(x_i), \quad (5.25)$$

обеспечивающее очень хорошую аппроксимацию одномерного уравнения непрерывности.

Описанная эффективная разностная схема обычно имеет более широкие пределы применимости по сравнению с другими «традиционными» (например, центрально-разностными) схемами, в которых плотности тока определяются из полиномиальных пробных функций для квазиуровня Ферми или для концентраций носителей заряда. При допущении $|V_{i+1} - V_i| \ll \varphi_T$ и $|\varphi_{pi} - \varphi_{pi-1}| \ll \varphi_T$ уравнение (5.24) существенно упрощается. Поскольку при выводе такие допущения не были сделаны, конечно-разностные аппроксимации (5.23), (5.24) значительно повышают точность численного моделирования или позволяют применять более грубые сетки в областях, где допущения, принятые при аналитическом вычислении интеграла в (5.22), нарушаются не слишком сильно.

Обобщенная конечно-разностная аппроксимация плотностей тока. Разностная аппроксимация (5.23), (5.24) была получена в упрощающем предположении $\Gamma_n = \Gamma_p = G_{Tn} = G_{Tp} = 0$, $\lambda_n = \lambda_p = \varphi_T$ без учета эффектов сильного легирования и искажения энергетических зон. Расширим возможности описанной эффективной конечно-разностной аппроксимации для одномерного обобщенного уравнения плотности тока (4.98), записанного в виде

$$J_n = q\mu_n \left(\lambda_n \frac{dn}{dx} - n \frac{d\Phi_n}{dx} \right), \quad (5.26)$$

где $\Phi_n = (\Gamma_n - \chi)/q + V$ — эффективный квазипотенциал, действующий на электроны в зоне проводимости.

Уравнение (5.26) в нормированном виде записывается таким образом:

$$\frac{dn}{dx} = \chi_n \frac{dF}{dx} = J \frac{\kappa}{\mu}, \quad (5.27)$$

где

$$\chi = (\lambda/\varphi_T)^{-1}, \quad F = \Phi_n/\varphi_T. \quad (5.28)$$

Линейное неоднородное дифференциальное уравнение первого порядка (5.27) имеет решение

$$n = \exp\left(\int x dF\right) \left[\int \frac{J}{\mu} \exp\left(-\int x \frac{dE}{d\xi} d\xi\right) dx + C_0 \right], \quad (5.29)$$

где C_0 — константа интегрирования.

Разобьем отрезок $[0, L]$ сеткой с шагом h_i , индексом i будем обозначать значения переменных в узле i , индексом $i+1/2$ — в интервале между узлами i и $i+1$. В общепринятом (рассмотренном выше) предположении о постоянстве J , μ , κ между узлами сетки и при использовании приближенного равенства $\int \exp(-f(x)) dx \approx \frac{\exp(-f(x))}{f'}$ получим следующие соотношения:

$$n_i = \exp\left(F_i x_{i+1/2}\right) \left(\frac{x_{i+1/2} J_{i+1/2}}{\mu_{i+1/2}} \frac{e^{-F_i x_{i+1/2}}}{\frac{F_{i+1} - F_i}{h_i} x_{i+1/2}} + C_0 \right), \quad (5.30)$$

$$n_{i+1} = \exp\left(F_{i+1} x_{i+1/2}\right) \left(\frac{x_{i+1/2} J_{i+1/2}}{\mu_{i+1/2}} \frac{e^{-F_i x_{i+1/2}}}{\frac{F_{i+1} - F_i}{n_i} x_{i+1/2}} + C_0 \right). \quad (5.31)$$

Из (5.30) получаются обобщенные эффективные разностные аппроксимации (5.23), (5.24), которые используются в ФТ-модели с учетом ЭСЛ при расчетах на ЭВМ полупроводниковых структур с произвольным легированием:

$$J_{n,i+1/2} = \mu_{n,i+1/2} \frac{DV_{i+1/2}^n}{h_i} \frac{n_{i+1} e^{DV_{i+1/2}^n x_{i+1/2}^n} - n_i}{e^{DV_{i+1/2}^n x_{i+1/2}^n} - 1}, \quad (5.32)$$

где

$$DV_{i+1/2}^n = V_{i+1} - V_i + (\Delta E_{gi} A_{BNi} - \Delta E_{gi+1} A_{BNi+1}). \quad (5.33)$$

Аналогичный вид имеет разностная аппроксимация для J_p :

$$J_{p,i+1/2} = -\mu_{p,i+1/2} \frac{DV_{i+1/2}^p}{h_i} \frac{p_{i+1} e^{DV_{i+1/2}^p x_{i+1/2}^p} - p_i}{e^{DV_{i+1/2}^p x_{i+1/2}^p} - 1}, \quad (5.34)$$

где

$$DV_{i+1/2}^p = V_{i+1} - V_i + (\Delta E_{gi}(1 - A_{BNi}) - \Delta E_{gi+1}(1 - A_{BNi+1})). \quad (5.35)$$

В рассматриваемых далее ФТ-программах одномерного, квази-двумерного и двумерного моделирования используются именно разностные аппроксимации (5.32)—(5.35). Это позволяет получить в принципе приемлемую точность расчета на сетках с относительно небольшим количеством узлов, несмотря на большие градиенты численных решений. Подчеркнем, что получение эффективной конечно-разностной аппроксимации (5.32)—(5.35) можно обобщить на многомерные случаи, поскольку в рассматриваемых ситуациях составляющие плотности тока вдоль линий сетки в промежутках между узлами принимаются постоянными и аппроксимируются выражениями типа (5.24). Такая эффективная конечно-разностная аппроксимация пригодна не только для прямоугольных, но и для треугольных сеток.

Известно, что ошибка аппроксимации уравнения Пуассона пропорциональна третьей производной потенциала, а уравнения непрерывности — первой и второй производным соответствующей плотности тока. Поэтому, чтобы выбрать сетку строго оптималь-

но, надо сначала решить соответствующие уравнения и найти нужные распределения переменных. Общий подход заключается в постепенном улучшении начально выбранной сетки после нахождения промежуточных решений.

5.2. СХЕМЫ ДИСКРЕТИЗАЦИИ ФСУ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

Нормировка ФСУ. При конструировании численных алгоритмов и проведении ФТ-расчетов удобно пользоваться ФСУ, записанной в безразмерном виде. В большинстве работ по численному моделированию полупроводниковых структур используют так называемые стандартные нормировки независимых переменных, систематизированные в табл. 5.1.

Таблица 5.1

Стандартная нормировка независимых переменных

Переменная	Символ	Нормировка
D_n, D_p	D_0	$1 \text{ см}^2 \cdot \text{с}^{-1}$
$\mu_n \mu_p$		D_0/V_0
R		$D_0 C_0/x_0^2$
t		x_0^2/D_0
x	x_0	$[\varepsilon kT/(q^2 n_i)]^{1/2}$
V	V_0	kT/q
n, p, c	C_0	n_i

Нормированная ФСУ записывается в виде

$$\operatorname{div} \operatorname{grad} V - (n - p - C) = 0, \quad (5.36)$$

$$\operatorname{div} (D_n \operatorname{grad} n - \mu_n n \operatorname{grad} V - R(V, n, p)) = \frac{\partial n}{\partial t}, \quad (5.37)$$

$$\operatorname{div} (D_p \operatorname{grad} p + \mu_p p \operatorname{grad} V - R(V, n, p)) = \frac{\partial p}{\partial t}, \quad (5.38)$$

$$J_n = \mu_n n \operatorname{grad} V - D_n \operatorname{grad} n, \quad (5.39)$$

$$J_p = \mu_p p \operatorname{grad} V + D_p \operatorname{grad} p. \quad (5.40)$$

Ввиду того что зависимые переменные в нормированной стандартным способом системе (5.36) — (5.40) не одного порядка и возможны различные вычислительные сложности при отладке численных алгоритмов решения ФСУ, предложена «оптимальная» процедура нормировки, использующая нормировки независимых переменных, систематизированные в табл. 5.2, 5.3.

Таблица 5.2

Оптимальная нормировка независимых переменных

Переменная	Символ	Нормировка
x	x_0	$\max(x-y), x, y \in D$
V	V_0	kT/q
n, p, C	C_0	$\max C(x), x \in D$
D_n, D_p	D	$\max(D_n, x), D_p(x), x \in D$
μ_n, μ_p		D_0/V_0
R		$D_0 C_0/x_0^2$
t		x_0^2/D_0

Таблица 5.3

Числовые значения стандартной и оптимальной нормировок

Переменная	Стандартная нормировка (см. табл. 5.2)	Оптимальная нормировка (см. табл. 5.2)
x	$4.09 \cdot 10^{-3}$ см	$3 \cdot 10^{-3}$ см
V	0.026В	0.026 В
n, p, C	10^{10} см ⁻³	10^{20} см ⁻³
D_n, D_p	1 см ² /с	36.2 см ² /с
μ_n, μ_p	38.7 см ² /(В·с)	1400 см ² /(В·с)
R	$6 \cdot 10^{14}$ см ⁻³ ·с ⁻¹	$4 \cdot 10^{26}$ см ⁻³ ·с ⁻¹
t	$1.67 \cdot 10^{-5}$ с	$2.5 \cdot 10^{-7}$ с

Оптимально нормированная ФСУ записывается в виде

$$\lambda^2 \operatorname{div} \operatorname{grad} V - (n - p - C) = 0, \quad (5.41)$$

$$\operatorname{div}(D_n \operatorname{grad} n - \mu_n n \operatorname{grad} V) - \dot{R}(V, n, p) = \frac{\partial n}{\partial t}, \quad (5.42)$$

$$\operatorname{div} \operatorname{div}(D_p \operatorname{grad} p + \mu_p p \operatorname{grad} V) - R(V, n, p) = \frac{\partial p}{\partial t}, \quad (5.43)$$

$$\lambda^2 = V_0 \epsilon / x_0^2 q C_0.$$

Уравнения для плотностей тока идентичны (5.39), (5.40).

Типы схем дискретизации. Пространственная дискретизация ФСУ полупроводниковой структуры дает для каждого узла сетки одно конечно-разностное уравнение Пуассона и два конечно-разностных уравнения непрерывности. Полученная система содержит обыкновенные дифференциальные уравнения во временной области, так как дискретизация в геометрическом пространстве уравнения непрерывности содержат производные концентрации носителей заряда по времени.

Записывая узловые значения искоемых функций в векторной форме, данную систему можно привести к виду

$$f_V(V, \varphi_p, \varphi_n, t) = 0, \quad (5.44)$$

$$\varphi_n(V, \varphi_p, \varphi_n, t) = 0, \quad (5.45)$$

$$\varphi_p(V, \varphi_p, \varphi_n, t) = 0, \quad (5.46)$$

где вектор-функции $f_V, \varphi_n, \varphi_p$ представляют собой пространственные конечно-разностные аппроксимации ФСУ.

Для интегрирования системы уравнений (5.44) — (5.46) требуются численные алгоритмы дискретизации во временной области и получение конечно-разностных аппроксимаций для производных.

Для дискретизации ФСУ во временной области существует несколько схем. Наиболее надежной и стабильной является чисто неявная схема:

$$F_1(V_{m+1}, n_{m+1}, p_{m+1}) = 0, \quad (5.47)$$

$$\frac{n_{m+1} - n_m}{\tau_{m+1}} - F_2(V_{m+1}, n_{m+1}, p_{m+1}) = 0, \quad (5.48)$$

$$\frac{p_{m+1} - p_m}{\tau_{m+1}} - F_3(V_{m+1}, n_{m+1}, p_{m+1}) = 0, \quad (5.49)$$

где $F_1(V, n, p), F_2(V, n, p), F_3(V, n, p)$ — стационарные члены соответственно уравнений Пуассона и непрерывности, составляющие ФСУ; m, τ — соответственно номер и величина шага по времени.

Однако она требует решения очень большой нелинейной системы уравнений на каждом шаге. Значительно большей простотой отличается полунеявная схема

$$\frac{n_{m+1} - n_m}{\tau_{m+1}} - F_2(V_m, n_{m+1}, p_m) = 0, \quad (5.50)$$

$$\frac{p_{m+1} - p_m}{\tau_{m+1}} - F_3(V_m, n_{m+1}, p_{m+1}) = 0, \quad (5.51)$$

$$\tau_{m+1}(V_{m+1} - V_m) - F_1(V_{m+1}, n_{m+1}, p_{m+1}) = 0, \quad (5.52)$$

где τ_{m+1} — стабилизирующий коэффициент, которая позволяет решать каждое уравнение по отдельности. Но у этой схемы возможны осцилляции решения.

Наконец, предложена третья схема, в которой последнее уравнение (5.52) из системы (5.50) — (5.52) заменено на

$$\nabla^2 \frac{V_{m+1} - V_m}{\tau_{m+1}} - F_2(V_{m+1}, n_{m+1}, p_{m+1}) + F_3(V_{m+1}, n_{m+1}, p_{m+1}) = 0. \quad (5.53)$$

Данная схема не имеет ограничений на величину τ , хотя не является полностью консервативной. Выбор той или иной схемы зависит прежде всего от ресурсов ЭВМ.

В МКР имеется уникальная возможность использования эффективной конечно-разностной аппроксимации типа (5.24), (5.25) и (5.30) — (5.35), резко снижающей размерность дискретизации и повышающей надежность схем дискретизации различного типа. Это следует иметь в виду при разработке численных методов решения ФСУ, в том числе универсальных, для перспективных субмикронных сильнолегированных сверхтонкослойных структур, с максимальной эффективностью используя знание физики и особенностей моделируемых структур, возможности успешного применения обоснованных физических аппроксимаций, допущений, упрощений сложного реального объекта.

5.3. ОБЩИЕ МЕТОДЫ РЕШЕНИЯ СИСТЕМ КОНЕЧНО-РАЗНОСТНЫХ УРАВНЕНИЙ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

При анализе переходных процессов пространственная дискретизация и неявное интегрирование во времени для каждого момента времени дают нелинейную систему уравнений вида

$$F_V(V, n, p) = 0,$$

$$F_n(V, n, p) = 0,$$

$$F_p(V, n, p) = 0,$$

в которой для простоты опущены аргумент t^m и верхние индексы. При анализе статических режимов система аналогичного вида получается для каждого набора внешних напряжений смещения. Поэтому в обоих случаях приходится решать систему из $3N$ нелинейных уравнений большой размерности, для которой, например в двумерном случае, количество точек сетки обычно составляет от 500 до 5000. Решение такой системы уравнений большой размерности — наиболее трудоемкая и ответственная процедура в процессе численного моделирования полупроводниковых структур. Следовательно, при решении многомерных задач крайне актуален выбор наиболее эффективных вычислительных алгоритмов.

Метод Ньютона. Существует два принципиально различных метода решения конечно-разностных ФСУ: Ньютона и Гуммеля. В методе Ньютона все три конечно-разностных уравнения решают одновременно, в последовательном методе Гуммеля уравнения ФСУ решают отдельно, поочередно используя значения переменных, полученных с предыдущих итераций.

Систему конечно-разностных уравнений ФСУ

$$(\Delta^h V)_{ij} = n_{ij} - p_{ij} - (C_D)_{ij} + (C_A)_{ij}, \quad (5.54)$$

$$(\Delta_n(V))_{ij} = R(p_{ij}, n_{ij}), \quad (5.55)$$

$$(\Delta_p(V))_{ij} = R(p_{ij}, n_{ij}), \quad 0 < i < M_1, \quad 0 < j < M_2, \quad (5.56)$$

где разностные операторы Δ^n , Λ_n , Λ_p определены равенствами (5.20), (5.25), (5.32) — (5.35), запишем в виде

$$(\Delta^h V)_{ij} - n_{ij} + p_{ij} = (C_D)_{ij} + (C_A)_{ij}, \quad (5.57)$$

$$F_{ij}(V_{l-1,j}, V_{ij}, V_{i+1,j}, V_{i,j+1}, n_{i-1,j}, \dots, n_{i,j+1}, p_{i-1,j}, \dots, p_{i,j+1}) = 0, \quad (5.58)$$

$$G_{ij}(V_{l-1,j}, V_{ij}, V_{i+1,j}, V_{i,j-1}, V_{i,j+1}, n_{i-1,j}, \dots, n_{i,j+1}, p_{i-1,j}, \dots, p_{i,j+1}) = 0. \quad (5.59)$$

Линеаризуя уравнения (5.58), (5.59) в окрестности известного l -го приближения, получаем, что для определения $(l+1)$ -го приближения необходимо решить систему линейных уравнений с соответствующими разностными краевыми условиями.

Таким образом, на каждом шаге в методе Ньютона необходимо одновременно решать линеаризованную систему линейных уравнений. После того как эта система уравнений решена, полагаем $l = l+1$ и переходим в рамках внешнего итерационного процесса к определению следующего приближения. Достоинством данного алгоритма является высокая (квадратичная) скорость сходимости внешнего итерационного процесса, не зависящая от уровня инжекции. Однако имеется ряд значительных вычислительных трудностей, возникающих при реализации данного алгоритма в многомерном, в частности в двумерном, приближении, связанных прежде всего с решением на ЭВМ системы линейных уравнений. Если при решении одномерной задачи матрица системы уравнений сильно разрежена и имеет простую структуру, что позволяет эффективно решать систему прямыми методами (например, методом матричной прогонки), то в двумерном случае из-за сложных систем уравнений большой размерности прежние методы решения уже неэффективны. Это приводит к тому, что определение $(l+1)$ -го приближения в методе Ньютона требует огромных вычислительных затрат, значительно больших, чем в других последовательных методах решения конечно-разностных уравнений ФСУ. Кроме того, непосредственная реализация метода Ньютона требует существенно больших затрат оперативной памяти ЭВМ по сравнению с методами реализации последовательных методов Гуммеля.

Хотя в методе Ньютона отсутствуют проблемы, связанные со сходимостью на высоком уровне инжекции, зато примерно в 7—10 раз возрастает объем оперативной памяти ЭВМ и в 10—20 раз — время решения систем уравнений в режимах низких и средних инжекций.

Метод Гуммеля. В простейшем варианте метода Гуммеля каждое из указанных конечно-разностных уравнений — Пуассона и непрерывности для дырок и электронов — решают независимо, путем их разделения и последовательного решения трех систем уравнений (5.54) — (5.56). На рис. 5.4 проиллюстрирован алгоритм

метода Гуммеля. Сначала решают уравнение Пуассона (5.54) в предположении заданных квазиуровней Ферми. Полученные на первом шаге значения потенциала подставляют в уравнение непрерывности, которое затем решают отдельно и последовательно. Дан-

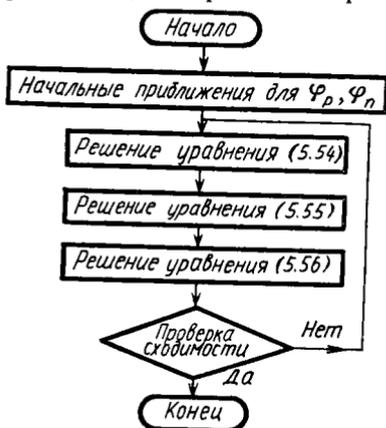


Рис. 5.4. Алгоритм физико-топологического расчета БТ Гуммеля

ный итерационный процесс последовательно повторяют до тех пор, пока для всех неизвестных переменных не будут получены совпадающие с заданной точностью последовательно рассчитанные значения потенциала. При многомерном моделировании полупроводниковых структур метод Гуммеля с последовательным алгоритмом (рис. 5.4) имеет преимущество, так как экономит память ЭВМ и хорошо сходится во всех случаях, когда три уравнения ФСУ слабо связаны между собой. Кроме того, при удачном выборе начального приближения результаты уже начальных циклов итерационного процесса

оказываются близкими к точному решению.

В методе Гуммеля, заложенном в программы двумерного ФТ-моделирования, используется следующий внешний итерационный процесс:

$$\Delta^h V_{ij}^{l+1} = n_{ij}^{l+1} - p_{ij}^{l+1} - C_{Dij} + C_{Aij} + (n_{ij}^{l+1} + p_{ij}^{l+1})(V_{ij}^{l+1} - V_{ij}^l), \quad (5.60)$$

$$\Lambda_n(V^l) n_{ij}^{l+1} - p_{ij}^l R_{ij}^l n_{ij}^{l+1} = -R_{ij}^l, \quad (5.61)$$

$$\Lambda_p(V^l) p_{ij}^{l+1} - n_{ij}^{l+1} R_{ij}^l p_{ij}^{l+1} = -R_{ij}^l. \quad (5.62)$$

Здесь Δ^h , Λ_n , Λ_p — разностные операторы, определенные равенствами (5.20), (5.25), (5.32) — (5.35); R_{ij}^l — дискретизированная запись выражения (4.126) для интегральной рекомбинации, стоящего в фигурных скобках; $l=0, 1, 2$ — номер итерации.

Соотношения (5.60) — (5.62) определяют внешний итерационный процесс. Отметим, что для решения каждой из систем линейных уравнений (5.60) — (5.62), т. е. для определения V^{l+1} , n^{l+1} , p^{l+1} , в двумерном случае необходимо использовать внутренний итерационный процесс (этот вопрос будет рассмотрен ниже). Внешний итерационный процесс заканчивается, если выполняется условие

$$\varepsilon < \varepsilon^* \quad \left\{ \begin{array}{l} |V_{ij}^{l+1} - V_{ij}^l|, |V_{ij}^{l+1}| \leq \varphi_T, \\ \varepsilon = \max_{ij} \varepsilon_{ij}, \quad \varepsilon_{ij} = \left\{ \begin{array}{l} \frac{|V_{ij}^{l+1} - V_{ij}^l|}{|V_{ij}^{l+1}|}, \quad |V_{ij}^{l+1}| > \varphi_T, \end{array} \right. \end{array} \right. \quad (5.63)$$

где ε^* — заданная в численном алгоритме точность.

Сравним методы Ньютона и Гуммеля с целью определения перспектив моделирования полупроводниковых структур. Сравнение этих методов на ряде модельных задач подтверждает, что в режиме высокой инжекции хорошую сходимость можно получить лишь с помощью метода Ньютона, а блочно-ньютоновский и гуммелевский методы являются более экономичными для расчетов при низкой и средней инжекции. Кроме того, в методе Ньютона предпочтительнее использовать в качестве переменных концентрации или экспоненты квазипотенциалов, но не квазиуровни Ферми: от выбора переменных зависит скорость сходимости метода Ньютона, если полученные системы линейных алгебраических уравнений (СЛАУ) решаются точно, прямым методом исключения Гаусса.

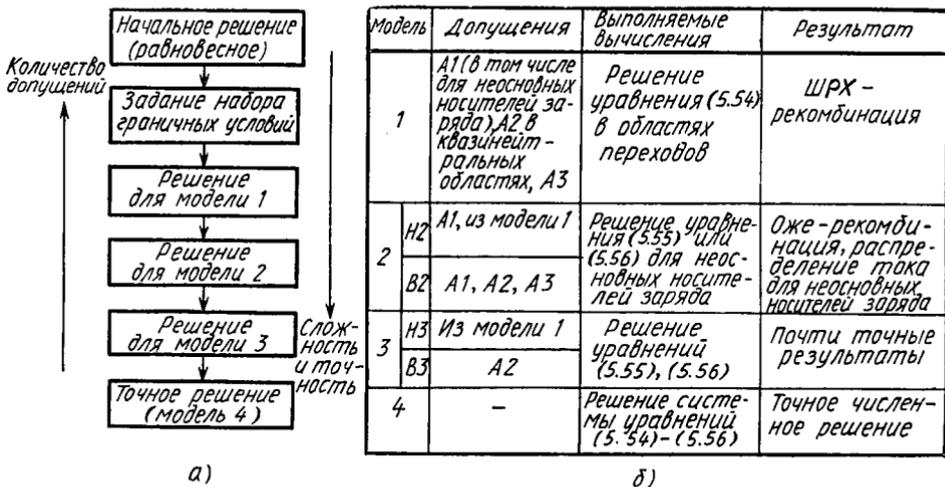


Рис. 55. Иерархия моделей (а); используемые допущения в пошаговом алгоритме ФТ-программы (б)

Вероятно, наибольшей эффективностью обладают программы, использующие оба метода: для расчета при низких инжекциях и начальное приближения при высоких — метод Гуммеля, для расчета при высоких инжекциях — метод Ньютона.

Другие методы. Отметим также большую эффективность других общих методов поэтапного решения ФСУ, когда последовательно решают системы уравнений от самой простой, с множеством ограничений (квазинейтральность, отсутствие рекомбинации и т. д.) до самой сложной, без ограничений, причем решение, полученное на предыдущем этапе, является начальным приближением для решения на последующем этапе.

В данном алгоритме численного моделирования сохраняются преимущества алгоритмов последовательного решения, при этом гарантируется сходимость и снижение затрат машинного времени примерно на порядок. В алгоритме поэтапного (пошагового) ре-

шения достигается требуемое итоговое решение для произвольного рабочего режима за четыре шага, выполняемых по иерархической схеме, представленной на рис. 5.5, а, б. Если определенные подсистемы из системы уравнений (5.54)—(5.56) используются в соответствующих областях полупроводниковой структуры, то допущения:

A1) постоянства квазиуровня Ферми для основных носителей заряда и приложенных напряжений;

A2) нейтральности объемного заряда

$$\rho = 0 \rightarrow V = f(\varphi_n, \varphi_p); \quad (5.64)$$

A3) пренебрежимо малого влияния рекомбинации на механизмы переноса носителей заряда: $R = 0$,

— учитывают взаимную развязку этих уравнений. Исключение исходных допущений позволяет шаг за шагом построить иерархический набор численных моделей, представленных на рис. 5.5, б. Поскольку справедливость указанных допущений зависит от режима работы приборной структуры, эту иерархию подразделяют на две ветви соответственно для низкого (Н) и высокого (В) уровней инжекции. Для средних уровней инжекции можно двигаться по любой из ветвей (Н или В), получая примерно равные затраты машинного времени и точность численных расчетов.

В простейшей модели этой иерархии нужно решить лишь одно из уравнений для каждой подобласти приборной структуры. На каждом иерархическом уровне решение такой неполной системы уравнений дает хорошие приближения соответствующих внутренних переменных. Поэтому результат численного моделирования на предыдущем уровне служит начальным приближением для следующего иерархического уровня, благодаря чему в каждом случае требуется количество итераций оказывается небольшим. Кроме того, при необходимости совместного решения более чем одного уравнения эти уравнения всегда остаются слабо связанными между собой и, следовательно, их можно решать с помощью эффективного последовательного алгоритма, успешно применяющего подпрограммы непосредственного решения систем с ленточными матрицами с обеспечением точности решения.

При точном численном режиме всех трех уравнений ФСУ такая развязка достигается путем обобщения условия квазинейтральности, что дает функцию $V = V(\varphi_n, \varphi_p, \rho)$, которая затем подставляется в оба дискретизированных уравнения переноса, при этом уравнение Пуассона можно рассматривать как средство дальнейшего определения ρ , а не V .

5.4. ИТЕРАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ ЛИНЕАРИЗОВАННЫХ УРАВНЕНИЙ ПУАССОНА И НЕПРЕРЫВНОСТИ

Сравнение методов. На каждом шаге внешнего итерационного процесса в общих методах Ньютона и Гуммеля необходимо решать одну и три СЛАУ большой размерности. Прямые методы решения

таких уравнений требуют огромных вычислительных затрат, поэтому при численном моделировании структур в многомерном приближении используют итерационные методы (для внутренних итерационных процессов).

Прежде чем перейти к обсуждению итерационных методов решения СЛАУ, рассмотрим, в каких системах переменных можно решать ФСУ: 1) V — потенциал, n — концентрация электронов, p — концентрация дырок; 2) V , $\varphi_n = V - \varphi_T \ln(n/n_{ie})$ — квазиуровень (квазипотенциал) Ферми для электронов, $\varphi_p = V + \varphi_T \ln(p/n_{ie})$ — квазиуровень (квазипотенциал) Ферми для дырок; 3) V , $\Phi_p \approx \exp(\varphi_n/\varphi_T)$, $\Phi_p = \exp(\varphi_p/\varphi_T)$ — экспоненты квазипотенциалов. Каждая из систем обладает рядом преимуществ и недостатков. Только для системы 3 матрица СЛАУ является симметричной и с диагональным преобладанием. Наименьший разброс переменных имеет система 2, наибольший — система 3, что не позволяет рассчитывать режимы с большими отрицательными напряжениями на электродах. С другой стороны, в системе переменных 2 уравнения имеют наибольшие нелинейности, что существенно затрудняет сходимость итерационного процесса. Видимо, эта система может быть использована только в нестационарном анализе, когда изменение переменных от итерации к итерации невелико. В статическом анализе, когда хорошее начальное приближение неизвестно, система 1 несколько лучше, чем система 3. Однако решение СЛАУ итерационными методами лучше всего проводить в системе переменных 3. Далее будут обсуждаться методы решения СЛАУ при анализе стационарных характеристик. В нестационарном анализе решение существенно упрощается, так как член $1/\tau$ дает сильное диагональное преобладание коэффициентов матрицы СЛАУ.

Матрицы СЛАУ Пуассона имеют сильное диагональное преобладание, и их числа обусловленности (отношение максимального собственного значения матрицы к минимальному) невелики. Вследствие малых чисел обусловленности нахождение систем разностных уравнений не вызывает затруднений практически любыми численными методами.

Определение решений СЛАУ непрерывности для электронов и дырок является значительно более трудной задачей. Коэффициенты этих уравнений зависят от потенциала электрического поля, который может существенно изменяться по полупроводниковой структуре. Данное обстоятельство и использование неравномерных сеток приводят к большим числам обусловленности СЛАУ. В связи с этим применение методов простой итерации и Зейделя, скорость сходимости которых обратно пропорциональна числу обусловленности, для решения СЛАУ непрерывности требует огромных вычислительных затрат.

Более высокую скорость сходимости имеют методы верхней релаксации и их модификации. Однако их эффективность в случае плохо обусловленных СЛАУ существенно зависит от выбора специальных итерационных параметров. Коэффициенты уравнений

непрерывности, а значит, и собственные значения матрицы СЛАУ сильно меняются в ходе итерационного процесса, особенно при значительных изменениях рабочих напряжений в приборных структурах. Поэтому вычисление оптимальных итерационных параметров весьма затруднительно и данные методы редко применяются для решения СЛАУ непрерывности.

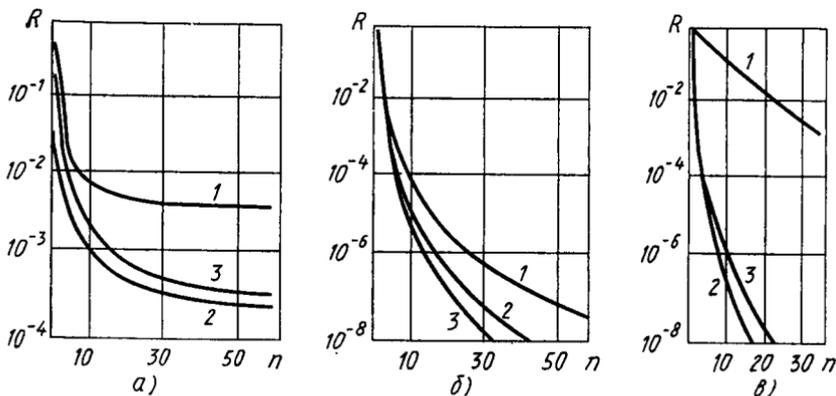


Рис. 5.6. Зависимости невязки решения от номера внутренней итерации, полученные для уравнений непрерывности дырок (а), электронов (б) и уравнения Пуассона (в):

1 — метод последовательной релаксации; 2 — метод неполного разложения; 3 — метод неполной факторизации

Как показывают численные эксперименты, метод неполной факторизации (МНФ) Булеева — Стоуна имеет достаточно высокую скорость сходимости только в случае сильно меняющихся коэффициентов матрицы СЛАУ. В этом методе необходимо подбирать итерационный параметр из интервала $[0, 1]$, причем неудачный выбор этого параметра ухудшает сходимость или приводит к расходимости итерационного процесса.

Релаксационный метод Федоренко (РМФ) обладает высокой скоростью сходимости, не зависящей от числа узлов сетки, и не требует выбора специальных итерационных параметров. Однако в этом методе требуются вспомогательные сетки. Автоматизировать же процедуру построения и использования сетки сложно, что ограничивает возможности применения такого метода для расчета ряда полупроводниковых структур.

Методы неполного разложения (МНР) и модификации МНР Холецкого с сопряженными градиентами (ICCG — Incomplete Cholesky — conjugate Gradient) (симметричный и асимметричный) высокoeffективны только для СЛАУ с симметричной положительно-определенной матрицей. МНР практически не требуют задания априорной информации для определения итерационных параметров, что является большим достоинством этих методов и их модификаций.

Численные эксперименты. Для БТ исследовалась эффективная разностная схема (5.60)—(5.62). В качестве внутренних итерационных процессов для решения уравнений Пуассона и непрерывности использовались МПВР, МНР, МНФ и РМФ.

Результаты сравнения с пересчетом затрат на одну итерацию последовательной верхней релаксации для пятой внешней итерации при $U_0 = -0,65$ В показаны на рис. 5.6. Экспериментально установлено, что уравнение Пуассона решается любым методом, но сходимость МПВР медленнее сходимости МНР или МНФ в 10 раз и практически МПВР не уменьшает

невязку решения уравнений непрерывности. При использовании остальных методов существенно уменьшалась невязка решений уравнений непрерывности на первых итерациях, но затем наступало насыщение. При достаточно большом номере внешней итерации невязка существенно уменьшалась только РМФ. Таким образом, последний метод наиболее эффективен для данного случая. Рассмотрим переход от системы переменных I, p, n, V к системе переменных II, Q, Φ, V , где $Q = p \varepsilon^{V/\varphi_T}$, $\Phi = n e^{-V/\varphi_T}$.

Система II имеет два недостатка:

1) сложность введения статистики Ферми — Дирака; 2) невозможность

расчета при больших внешних потенциалах, вызванная ограничением разрядной сетки ЭВМ; $U_{\max} < 1$ В для ЭВМ БЭСМ-6, $|U_{\max}| < 3$ В для ЭВМ ЕС. Однако матрицы систем уравнений непрерывности становятся симметричными с диагональным преобладанием. Экспериментально установлено, что уравнения непрерывности, записанные в системе переменных II, одинаково хорошо решаются любым итерационным методом. Окончательно выбран МНР, по скорости не уступающий МНФ, но не использующий ни одного итерационного параметра.

Таким образом, СЛАУ, записанные в переменных Q, Φ, V , решались поочередно МНР до достижения заданной сходимости по стабилизации V . На рис. 5.7 приведены зависимости от номера итерации l , полученные в различных численных экспериментах:

$$\varepsilon_{ix}^{l+1} = \max \left\{ \begin{array}{l} \left| \frac{V_{ij}^{l+1} - V_{ij}}{V_{ij}} \right|, \quad |V_{i,j}^{l+1}| > \varphi_T, \\ |V_{ij}^{l+1} - V_{ij}|, \quad |V_{i,j}^{l+1}| \leq \varphi_T. \end{array} \right.$$

Из графиков рис. 5.7 видно, что: 1) при решении ФСУ в системе переменных I достаточную точность позволяет достичь

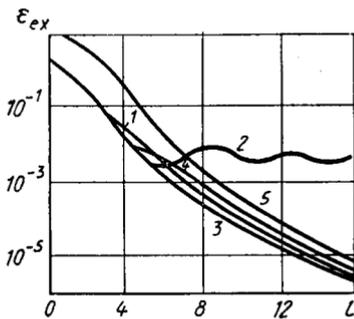


Рис. 5.7. Зависимость сходимости внешнего цикла по потенциалу от номера внешней итерации для различных численных экспериментов

только РМФ (кривая 1), а не МНР (кривая 2); 2) переход к системе II (кривые 3—5) с использованием МНФ позволяет достичь высокой скорости сходимости; 3) увеличение числа узлов сетки (кривые 1—3 — 374 узла, кривая 4 — 690 узлов, кривая 5 — 930 узлов) несущественно влияет на скорость сходимости МНР, по крайней мере до данного количества узлов; 4) усложнение приборной структуры влияет только на начальную невязку, но не на скорость сходимости. При 2—5 внутренних итерациях для уравнения Пуассона и 2—10 — для уравнений непрерывности бóльшая

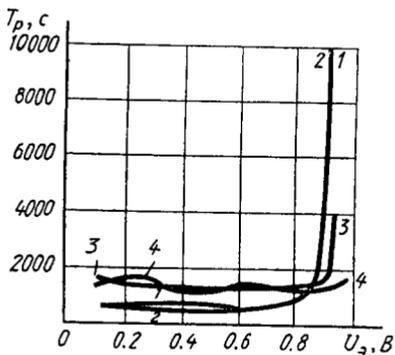


Рис. 5.8. Сравнение времени расчета ВАХ БТ при двумерном ФТ-моделировании для различных напряжений на эмиттерном переходе (низкого и высокого уровней инжекции)

часть вычислительных затрат приходится на вычисление коэффициентов матриц СЛАУ. Следовательно, дальнейшее уменьшение времени счета может быть достигнуто только в результате уменьшения числа внешних итераций. В качестве другого примера приведем результаты сравнения различных методов решения при моделировании БТ с существенно двумерным растеканием тока в области высокой инжекции: $U_э = 0.9$ В (плотность тока на краях эмиттера 10^5 А/см²). Сетка при численном моделировании содержала 1300 узлов. Сравнивались следующие методы решения 1) метод Гуммеля, уравнение Пуассона решается методом ИССГ, уравнения непрерывности — прямым методом Гаусса; 2) блочно-ньютоновский метод, переменные V , Φ_n , Φ_p и V , ϕ_n , ϕ_p ; 3) метод Ньютона, переменные V , Φ_n , Φ_p ; 4) то же для переменных V , n , p . Результаты времени расчета ВАХ приведены на рис. 5.8. Они подтверждают сделанный ранее вывод о том, что метод Гуммеля эффективнее при низкой инжекции, метод Ньютона — при высокой. Кроме того, видно, что решение СЛАУ эффективным прямым методом исключения Гаусса позволяет получать решение за приемлемое машинное время.

Перспективы прямых методов. В связи с увеличением объема памяти и быстродействия современных и перспективных ЭВМ в последнее время возрос интерес к прямым методам решения СЛАУ. Известно, что время решения СЛАУ методом исключения Гаусса увеличивается пропорционально $m^3 n^{1.5 \div 1.75}$, где m — количество независимых переменных, приходящихся на один узел; n — число узлов сетки. Кроме того, требуется гораздо бóльшая память, чем для реализации любого итерационного метода. Однако все эти недостатки перекрываются тем преимуществом, что метод Гаусса при правильном применении сразу дает точное (с точностью до ошибки представления чисел в ЭВМ) решение СЛАУ. Несмотря

на большие успехи в разработке и применении эффективных итерационных методов решения СЛАУ, они все же остаются достаточно капризными и неуниверсальными, о чем свидетельствуют противоречивые данные об их сравнительной эффективности.

5.5. ПРОГРАММЫ СТАЦИОНАРНОГО ФИЗИКО-ТОПОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

Программа стационарного одномерного ФТ-моделирования ТРАН 1. В программе ТРАН 1 для исключения постоянных коэффициентов и возможности работы с величинами, удобными для представления в ЭВМ, проводят специальную нормировку (по рассмотренным ранее правилам) и ФСУ записывается в виде

$$\frac{\partial}{\partial x} \left\{ \mu_n \left[n \frac{\partial}{\partial x} (V + \Delta E_g / 2q) - \frac{\partial n}{\partial x} \right] \right\} = -(pn - n_{ie}^2) RT, \quad (5.65)$$

$$\frac{\partial}{\partial x} \left\{ \mu_p \left[p \frac{\partial}{\partial x} (V - \Delta E_g / 2q) + \frac{\partial p}{\partial x} \right] \right\} = (pn - n_{ie}^2) RT, \quad (5.66)$$

$$\frac{\partial^2 V}{\partial x^2} = -(p - n + C_D - C_A). \quad (5.67)$$

В системе (5.65)–(5.67) все переменные и коэффициенты — нормированные и безразмерные, использовано соотношение Эйнштейна $D = \varphi_T \mu$, RT — член, стоящий в фигурных скобках в выражении для рекомбинации члена R . Вся область моделирования от $x_{\mathcal{E}}$ до $x_{\mathcal{K}}$ разбивается набором узлов с координатами x_i , $i = 1, 2, \dots, m$, причем $x_1 = x_{\mathcal{E}}$, $x_m = x_{\mathcal{K}}$. Каждое из уравнений (5.65)–(5.67) записывается на данной сетке в конечно-разностном виде с помощью определенной аппроксимации дифференциальных разностных операторов. Например, уравнение (5.67) в узле i записывается как

$$\frac{(V_{i+1} - V_i)/h_{i+1} - (V_i - V_{i+1})/h_i}{0,5(h_i + h_{i+1})} = -(p_i - n_i + C_{D_i} - C_{A_i}), \quad (5.68)$$

где индексом i помечены значения величин в i -м узле; $h_i = x_i - x_{i-1}$.

Аналогичные соотношения можно записать во всех узлах, кроме j -го и m -го, для которых записывается разностный аналог граничных условий:

$$U_{\mathcal{E}} = \left(V - \varphi_T \ln \left(\frac{n}{n_{ie}} \right) \right) \Big|_{x=x_{\mathcal{E}}}, \quad (5.69)$$

$$U_{\mathcal{K}} = \left(V - \varphi_T \ln \left(\frac{n}{n_{ie}} \right) \right) \Big|_{x=x_{\mathcal{K}}}. \quad (5.70)$$

Отметим, что в одномерной программе ТРАН 1 в некоторой точке $x_{\mathcal{B}}$ — «точке Гуммеля» вывода тока базы — используется специфическое граничное условие

$$\left(V + \varphi_T \ln \left(\frac{p}{n_{ie}} \right) \right) \Big|_{x=x_B} = 0. \quad (5.71)$$

Граничное условие (5.71) соответствует закреплению концентрации дырок и соответственно потенциала в точке $x=x_0$.

Таким образом, для (5.68) получаем систему из $m-2$ линейных алгебраических уравнений вида $A_i V_{i-1} - C_i V_i + B_i V_{i+1} = -F_i$, $i=2, 3, \dots, m-1$, с известными значениями V_1, V_2 . Аналогичные системы с более сложной аппроксимацией дифференциальных операторов (5.32)—(5.35) получаются и для уравнений (5.65), (5.66). Такие СЛАУ решаются просто и эффективно стандартным методом прогонки.

Поскольку в коэффициентах СЛАУ стоят величины, определяемые из других систем, решения проводят эффективным итерационным способом, а именно:

а) выбирают некоторое начальное значение величин V^0, n^0, p^0 , где верхний индекс — номер итерации;

б) при данном значении V решают уравнения (5.65), (5.66), получают значения n' и p' ;

в) подставляя n' и p' в уравнение (5.67), получают новое значение V' ;

г) сравнивают величины V^0 и V' . Если хотя бы в одном узле они отличаются более чем на заданную величину, то переходят к п. б), если нет, то итерационный процесс можно считать законченным.

Программа стационарного квазидвумерного ФТ-моделирования ТРАН 1.5. В квазидвумерном приближении в программе ТРАН 1.5 учитывается падение напряжения на сопротивлении активной области базы, зависящем от уровня инжекции, и снимаются ограничения закрепления концентрации дырок и потенциала в точке $x=x_B$ (5.71). При этом в программу ТРАН 1.5 вводится условие квазинейтральности

$$(p - n - C_A + C_D) \Big|_{x=x_B} = 0. \quad (5.72)$$

Для учета влияния падения напряжения на базовом сопротивлении в программе ТРАН 1.5 на каждой итерации вычисляется поверхностное сопротивление R_{SBA} :

$$R_{SBA} = \left[q \int_{x_{je}}^{x_{jc}} (\mu_n n + \mu_p p) dx \right]^{-1}. \quad (5.73)$$

Зависимость статического сопротивления базы с учетом эффектов вытеснения эмиттерного тока и модуляции проводимости базы вычисляют по формуле

$$R_{BA} = \frac{R_{SBA}}{a} \frac{\hbar}{l} \left[\ln \left(1 + \frac{a}{3} \right) \right], \quad a = R_{BA} I_B / \varphi_T, \quad (5.73a)$$

где I_B — рассчитываемый в программе ТРАН 1.5 ток в активной области базы.

С учетом соотношений (5.72), (5.73) граничные условия в точке $x=x_B$ (5.72) заменяются на

$$-R_{БA}I_{Б} = \left(V + \varphi_T \ln \frac{p}{n_{ie}} \right) \Big|_{x=x_B}. \quad (5.74)$$

Система уравнений (5.65)—(5.67) с соответствующими начальными и граничными условиями, включая (5.74), для программы ТРАН 1.5 дискретизируется, и СЛАУ, как и в программе ТРАН 1, решается методом прогонки. Итерационный процесс и алгоритм автоматического выбора шага по времени в программе ТРАН 1.5 аналогичны программе ТРАН 1.

Программы стационарного ФТ-моделирования ТРАН 1, ТРАН 1.5 на языке ФОРТРАН 4 для ЭВМ БЭСМ-6 позволяют: 1) проводить расчеты ВАХ $I_K, I_{Э}, I_{Б}^{-}(U_{Э})$, зависимости $B_N(I_K)$ при $U_{КБ} = \text{const}$ в схеме включения с общей базой или $U_{КЭ} = \text{const}$ в схеме включения с общим эмиттером; 2) вычислять электрические параметры БТ и электрофизические характеристики: электрические поля, потенциалы, квазипотенциалы Ферми, концентрации и др.; 3) «включать» и «выключать» модельные зависимости ЭФП, влияющие на характеристики приборов; 4) проводить указанные виды анализа в диапазоне температур $-30^{\circ}\text{C} \leq T \leq 130^{\circ}\text{C}$. ФТ-расчет БТ с помощью программ ТРАН 1, ТРАН 1.5 составляет 2—4 мин для 10 точек ВАХ для ЭВМ БЭСМ-6 в диапазоне плотностей тока $10^2—10^4$ А/см².

Программа стационарного двумерного ФТ-моделирования ТРАН 2. Программа ТРАН 2, написанная на языке ФОРТРАН 4 для ЭВМ БЭСМ-6, использующая внешний итерационный цикл (5.61)—(5.63), с приведенными в § 5.4 итерационными методами решения линеаризованных уравнений Пуассона и непрерывности позволяет проводить те же виды анализа, указанные в п. 1—4 для программ ТРАН 2, ТРАН 1.5, и рассчитывать одну точку ВАХ на разностной сетке 30×30 для низкого уровня инжекции за 5—7 мин, для высокого уровня инжекции — за 20—30 мин. Например, расчет восьми точек ВАХ занимает 100—120 мин машинного времени ЭВМ БЭСМ-6, $\epsilon_{ex} = 0.002$.

5.6. ПРОГРАММЫ НЕСТАЦИОНАРНОГО ДВУМЕРНОГО ФИЗИКО-ТОПОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ

Для расчета нестационарных характеристик БТ необходимо решение ФСУ полупроводника в двумерном приближении в общем виде:

$$\epsilon \epsilon_0 \Delta V = -g(p - n - C_D - C_A). \quad (5.75)$$

$$\text{div} [D_n \text{grad } n - \mu_n \text{grad} (V - \Delta E_g / 2q)] = R + \frac{\partial n}{\partial t}, \quad (5.76)$$

$$\text{div} (D_p \text{grad } p - \mu_p p \text{grad} (V + \Delta E_g / 2q)) = R + \frac{\partial p}{\partial t}, \quad (5.77)$$

где приняты стандартные обозначения, а граничные условия рассмотрены в § 4.2.

После аппроксимации нормированной системы ФСУ на неравномерной пространственной сетке с помощью чисто неявной схемы и замены переменных $\Phi = n \exp(-V/\varphi_T)$, $Q = p \exp(V/\varphi_T)$ получим следующий внешний итерационный процесс:

$$(\Delta^h V^{k+1})_{ij} = \Phi_{ij}^k \exp(V_{ij}^k) - Q_{ij}^k \exp(-V_{ij}^k) + \Phi_{ij}^k \exp(V_{ij}^k) + Q_{ij}^k \exp(-V_{ij}^k) (V_{ij}^{k+1} - V_{ij}^k), \quad (5.78)$$

$$(\Lambda_n (V^{k+1}) \Phi^{k+1})_{ij} - Q_{ij}^k \omega_{ij}^k \Phi_{ij}^{k+1} - \Phi_{ij}^{k+1} \frac{\exp(V_{ij}^{k+1})}{\tau^{l+1}} = \omega_{ij}^k - \Phi_{ij}^l \exp(V_{ij}^l) / \tau^{l+1}, \quad (5.79)$$

$$(\Lambda_p (V^{k+1}) Q^{k+1})_{ij} - \Phi_{ij}^{k+1} \omega_{ij}^k Q_{ij}^{k+1} \exp(-V_{ij}^{k+1}) / \tau^{l+1} = \omega_{ij}^k - Q_{ij}^l \exp(-V_{ij}^l) / \tau^{l+1}. \quad (5.80)$$

Здесь Δ^h , Λ_n , Λ_p — разностные операторы соответствующих дифференциальных уравнений (5.78) — (5.80); k — номер внешней итерации; l — номер слоя по времени; $\tilde{\omega}_{ij}^k$ — значение $\{A_{nn} + A_{pp} + \tau^l / [\tau_n(\rho + n_{ie}) + \tau_p(n + n_{ie})]\}$ на k -й итерации в узле (i, j) ; τ^l — шаг по времени на l -м слое.

Каждая из систем уравнений (5.78) — (5.80) решалась с помощью МНР до достижения стабилизации сходимости $\varepsilon_{in}^{k+1} = 0.04 \varepsilon_{ex}^k$, где ε_{ex}^k — сходимость внешнего итерационного процесса по стабилизации V . Максимальное число итераций МНР на каждой внешней итерации — 5, реально для достижения требуемой точности требовалось 2—3 итерации для систем (5.78), (5.79) и 3—5 итераций для системы (5.80). Начальное приближение (решение статической задачи) осуществлялось стандартными методами до $\varepsilon_{ex} = 10^{-6}$ с начальным шагом по времени $\tau^l = 10^{-5}$ нс.

Для выбора шага по времени применяют следующий алгоритм:

- 1) с момента времени t^l делают два шага по времени τ^l , полученное решение x_1 запоминается;
- 2) с того же момента времени t^l делают один двойной шаг, получается вектор решения x_2 ;
- 3) находят максимальное отличие между x_1 и x_2 : $Z_{12} = \max |x_1 - x_2|$.

Если $Z_{12} \leq \varepsilon^*$ — некоторое заданное число, то новый шаг по времени $\tau^{l+1} = \tau^l k (\varepsilon^* / Z_{12})^{1/2}$, где $k = 0.5 \div 0.9$ — коэффициент запаса. Если же $Z_{12} > \varepsilon^*$, то шаг по времени считают неудачным и с момента времени t^l попытку повторяют с шагом $\tau^l / 2$. Подобный алгоритм требует три дополнительных массива и некоторых затрат на «неэффективные» внешние итерации, однако он позволяет оптимально выбрать последовательность $\{\tau^l\}$.

Расчет переходного процесса базового и коллекторного токов на сетке 33×17 при $\varepsilon_{ex} = 0.001$ занимает около 8—10 ч машинного времени ЭВМ БЭСМ-6 и около 2,5—3 ч ЭВМ «Эльбрус 1-2К». Не-

смотря на то что схема (5.78)—(5.80) абсолютно устойчива к выбору шага, вследствие ее нелинейности требуется введение на каждом шаге по времени внешнего итерационного процесса для получения самосогласованного решения. При этом вычислительные затраты времени велики и не могут быть кардинально уменьшены за счет частичных модификаций нестационарных программ моделирования.

Для сокращения на порядок вычислительных затрат в модификационном варианте программы нестационарного ФТ-моделирования в двумерно приближении необходимо использовать полунейные абсолютно устойчивые конечно-разностные схемы, в которых потенциал электрического поля определяется не из уравнения Пуассона, а из уравнения непрерывности для полного тока электронов и дырок:

$$\operatorname{div}(\mathbf{I}_n - \mathbf{I}_p) = \frac{\partial}{\partial t} \Delta V. \quad (5.81)$$

Для всех схем строят специальные конечно-разностные аппроксимации для компонент электронного и дырочного токов типа модификаций аппроксимаций Шарфеттера — Гуммеля (5.32)—(5.35). В данных принципиально новых методах существенное сокращение времени расчета переходных процессов достигается за счет ликвидации внешнего итерационного процесса на каждом шаге по времени.

Известны три эффективные полунейные схемы (индексы i, j для краткости опущены):

$$1) \Lambda_n(V^l) n_n^{l+1} - p^l \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{n^{l+1} - n^l}{\tau^{l+1}}, \quad (5.82)$$

$$\Lambda_p(V^l) p^{l+1} - p^{l+1} \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{p^{l+1} - p^l}{\tau^{l+1}}, \quad (5.83)$$

$$\Lambda(n^{l+1}, p^{l+1}) V^{l+1} = F^{l+1} - \frac{\Delta^h V^{l+1} - \Delta^h V^l}{\tau^{l+1}}, \quad (5.84)$$

$$2) \Lambda_n(V^l) n^{l+1} - p^l \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{n^{l+1} - n^l}{\tau^{l+1}}, \quad (5.85)$$

$$\Lambda_p(V^l) p^{l+1} - p^{l+1} \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{p^{l+1} - p^l}{\tau^{l+1}}, \quad (5.86)$$

$$\Lambda(n^{l+1}, p^{l+1}) V^{l+1/2} = F^{l+1} - \frac{\Delta^h V^{l+1} - \Delta^h V^l}{\tau^{l+1}}, \quad (5.87)$$

$$\Delta^h V^{l+1} = n^{l+1} - p^{l+1} - (C_D + C_A) + \alpha(n^{l+1} + p^{l+1})(V^{l+1} - V^{l+1/2}); \quad (5.88)$$

$$3) \Lambda(n^l, p^l) V^{l+1} = \frac{F^l - \Delta^h V^{l+1} - n^l + p^l + (C_D + C_A)}{\tau^{l+1}}, \quad (5.89)$$

$$(\Delta_n(V^{l+1})n^{l+1}) - p^l \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{n^{l+1} - n^l}{\tau^{l+1}}, \quad (5.90)$$

$$(\Delta_p(V^{l+1})p^{l+1}) - p^{l+1} \omega^l n^{l+1} = -(n_{ie}^2) \omega^l + \frac{p^{l+1} - p^l}{\tau^{l+1}}, \quad (5.91)$$

где $\Delta(n^l, p^l) V^{l+1}$ — разностный аналог

$\nabla\{(\mu_n n^l + \mu_p p^l) \nabla V^{l+1}; F^l$ — разностный аналог

$\nabla\{(\mu_n \nabla n^l - \mu_p \nabla p^l) - (\mu_n n^l + \mu_p p^l) \nabla(\Delta E_g/2)\}$,

зависящие от используемых аппроксимаций для плотностей тока.

Схема 1 (5.82)—(5.84) является абсолютно устойчивой по отношению к выбору шага по времени τ при минимальных вычислительных затратах. Однако у нее есть существенный недостаток: на каждом шаге по времени конечно-разностное уравнение (5.84) со значительной погрешностью аппроксимирует уравнение Пуассона. Вследствие этого схема 1 обладает свойством накопления погрешности вычислений, что не позволяет проводить расчеты в нестационарном режиме с достаточно большими значениями τ .

Схема 2 (5.85)—(5.88) отличается от схемы 1 тем, что новое значение потенциала вычисляется в два приема: сначала из уравнения (5.87) вычисляется $V^{l+1/2}$, а затем из уравнения (5.89) — V^{l+1} . В уравнении (5.88) α — дополнительный итерационный параметр, $\alpha=0$ (τ). Преимуществом схемы 2 перед схемой 1 при такой же абсолютной устойчивости является то, что при выборе $\alpha=0$ (τ) уравнение Пуассона на каждом шаге решается с минимальной погрешностью 0 (τ^3) и не происходит накопления вычислительной погрешности.

В схеме 3 (5.89)—(5.91) самосогласованно на $(l+1)$ -м временном слое сначала определяется потенциал V^{l+1} через концентрации n^l, p^l , а затем концентрации n^{l+1}, p^{l+1} через потенциал V^{l+1} . Схема 3 также характеризуется абсолютной устойчивостью и минимальным накоплением вычислительной погрешности. Ее большие потенциальные возможности в полной мере реализуются при замене старых переменных (n, p, V) новыми переменными (Φ, Q, V), $n = \Phi \exp V$, $p = Q \exp(-V)$ и использовании соответствующей аппроксимации Шарфеттера — Гуммеля для плотностей тока в новых переменных в уравнениях (5.90), (5.91).

В этом случае все три СЛАУ схемы 3 имеют симметричные матрицы коэффициентов, причем матрицы коэффициентов СЛАУ (5.90), (5.91) обладают еще диагональным преобладанием. Поэтому при реализации программы моделирования можно использовать численные методы для СЛАУ с симметричными матрицами коэффициентов типа ИССГ. Схема 3 с новыми переменными взята за основу при реализации ФТ-программ. Время расчета переходных процессов с помощью ФТ-программы ТРАН 2Д составляет 20—60 мин на ЭВМ в зависимости от числа шагов по времени, размерности сетки и пространственного разбиения. Одномерный расчет занимает 3—5 мин.

ЭЛЕКТРИЧЕСКИЕ МОДЕЛИ ИНТЕГРАЛЬНЫХ ПРИБОРОВ ДЛЯ МАШИННОГО РАСЧЕТА БИС

6.1. ТРЕБОВАНИЯ К ЭЛЕКТРИЧЕСКИМ МОДЕЛЯМ. КЛАССИФИКАЦИЯ ЭЛЕКТРИЧЕСКИХ МОДЕЛЕЙ ИНТЕГРАЛЬНЫХ ПРИБОРНЫХ СТРУКТУР

Требования к электрическим моделям. Опыт машинного проектирования БИС показывает, что применение ЭВМ при расчете БИС, состоящих из нескольких сотен—тысяч активных приборных структур (например, БТ и МОПТ), наиболее эффективно при использовании достаточно точных электрических моделей приборных структур. Несмотря на огромный прогресс методов численного ФТ-моделирования, даже при внедрении в практику машинного проектирования соответствующего программного обеспечения и высокопроизводительных ЭВМ в ближайшей перспективе трудно ожидать реализацию возможности моделирования на уровне ФСУ полупроводника фрагментов БИС, состоящих из десятков—сотен приборных структур.

Под электрической моделью активных приборных структур БИС (диодов, БТ, МОПТ и др.) понимают систему уравнений электрической цепи (или адекватное математическое описание другого вида), позволяющих определить с требуемой точностью необходимые электрические характеристики как самих структур, так и фрагментов БИС на их основе в различных условиях функционирования структур и схем. Например, статические и динамические модели приборов БИС описываются системой электрических уравнений, связывающих токи и напряжения на выводах приборов с их внутренними параметрами при работе в стационарном и нестационарном режимах. Исходными параметрами электрических моделей интегральных структур являются прежде всего электрические параметры (например, обратные токи и емкости p - n -переходов, коэффициент передачи БТ, параметры различных ВАХ прибора), определяемые либо из внешних электрических измерений, либо из характеристик численных ФТ-расчетов. Искомыми выходными характеристиками электрических моделей являются токи и напряжения на выводах приборов.

Исходя из практических задач машинного проектирования и многоуровневого моделирования БИС эффективная электрическая модель интегральной приборной структуры должна удовлетворять

следующим основным требованиям: 1) обеспечивать необходимую адекватность функционирования приборов в широком диапазоне рабочих напряжений, токов, внешних воздействий (температуры, формы входных сигналов и др.), отображать с требуемой точностью разнообразные электрические характеристики в заданных произвольных технико-эксплуатационных режимах; 2) иметь однозначное, достаточно адекватное соответствие между электрическими параметрами модели или элементами эквивалентной схемы прибора и электрофизическими процессами в структуре в стационарном или нестационарном режиме; 3) включать некоторые специфические аппроксимации и упрощения для использования без излишней сложности электрических моделей в универсальных системах автоматизированного проектирования БИС; 4) обеспечивать возможность достижения компромисса точность — сложность расчета, минимизации вычислительных ресурсов за счет выбора достаточно простой и точной электрической модели из иерархического набора вложенных самосогласованных моделей, ориентированных на тот или иной диапазон режима работы в стационарных и нестационарных режимах.

Перечислим дополнительные требования, которым должны удовлетворять электрические модели актуальных приборных структур: 1) в качестве параметров модели целесообразно использовать электрические параметры структур, например коэффициенты передачи тока, емкости p - n -переходов, сопротивления областей для БТ, удельную крутизну, пороговые напряжения, коэффициент влияния подложки для МОПТ; 2) эквивалентная схема структур должна состоять из набора стандартных элементов электрических цепей, связывающих токи и напряжения (резисторов, конденсаторов, генераторов тока и напряжения и др.); 3) проблемы точности, «физичности», универсальности электрических моделей должны быть увязаны с приемлемыми вычислительными затратами и ресурсами современных и перспективных высокопроизводительных ЭВМ.

В зависимости от вида разрабатываемой схемы (нелинейная, линейная, низкочастотная, высокочастотная, сверхвысокочастотная и др.) разработчик должен синтезировать и применять те или иные электрические модели. При анализе нелинейных цифровых БИС необходимо применять модели для большого сигнала, адекватно описывающие характеристики приборов в широком диапазоне изменений рабочих напряжений и токов. Для линейных схем, работающих на малом сигнале, следует увеличить точность мало-сигнальной электрической модели вблизи рабочей точки. В высокочастотных и СВЧ-схемах требуется качественно учитывать внутреннюю инерционность, специфические инерционные особенности работы приборных структур, наличие многих электрических паразитных параметров, влияние которых не столь существенно в других частотных диапазонах.

Наиболее общими разновидностями электрических моделей яв-

ляются формальные и физические электрические модели. Формальные электрические модели обычно используют для аппроксимации характеристик приборных структур БИС, когда физика и специфика их функционирования изучены недостаточно полно или необходим только качественный анализ работы разрабатываемых схем. Как правило, формальные электрические модели содержат большое количество внутренних параметров и требуют большого количества электрических измерений, настроек экспериментальных и модельных зависимостей.

Электрические модели приборных структур делятся также на статические и динамические. Статические электрические модели отражают стационарное состояние прибора при неизменных внешних управляющих воздействиях и не учитывают переходные процессы и характеристики в приборе. Динамические модели дополнительно отражают переходные процессы, электрические характеристики, изменяющиеся во времени, определяющие поведение прибора при изменении во времени управляющих воздействий. Динамические модели подразделяются на малосигнальные, а также модели, работающие на большом сигнале. Кроме того, различают низкочастотные, высокочастотные и сверхвысокочастотные динамические модели.

Для полупроводниковых приборных структур критерий «малости» сигнала выполняется, если амплитуда входного сигнала не превышает нескольких фт. В электрических моделях большого сигнала часть параметров, например коэффициенты передачи тока, емкости p - n -переходов, сопротивления характерных областей, зависят от режима работы. В малосигнальных электрических моделях эти параметры считаются постоянными, равными своим значениям в рабочей точке, относительно которой воздействуют малые по амплитуде управляющие входные сигналы.

Биполярный транзистор является важнейшим полупроводниковым прибором как в дискретной электронике, так и в бурно развивающейся области БИС, СБИС, С³БИС. Уникальной особенностью БТ является способность усиливать и переключать электрические сигналы, поступающие на вход трехполюсника. Благодаря достижениям технологии, использованию новых физических и структурных принципов изготовления и режимного использования БТ в настоящее время значительно улучшены их надежность, усилительные и переключательные характеристики.

Особенностью поведения БТ является взаимодействие эмиттерного и коллекторного p - n -переходов, когда ток одного из переходов может эффективно управлять током другого перехода при расстоянии друг от друга меньше диффузионной длины носителей заряда. В основе работы БТ лежат принципы инжекции через один p - n -переход неосновных носителей заряда и собирания их другим переходом. Принципы усиления и преобразования сигналов отражены в самом названии БТ: transistor — transfer resistor — преобразователь резисторов.

Каждый из p - n -переходов БТ можно включить в прямом либо в обратном направлении. При напряжении на переходе, равном нулю, переход находится в состоянии термодинамического равновесия. В зависимости от напряжений, приложенных к p - n -переходам трехполюсника, различают три режима работы БТ:

1) отсечки — оба p - n -перехода включены в обратном направлении, закрыты, через БТ протекает минимально возможный ток, равный обратным токам переходов;

2) активный — один переход включен в прямом направлении, другой — в обратном, т. е. один переход открыт, другой — закрыт. Различают активный нормальный режим, когда открыт эмиттерный, а закрыт коллекторный переход, и активный инверсный режим, когда открыт коллекторный, а закрыт эмиттерный переход;

3) насыщения — оба перехода смещены в прямом направлении и открыты. Аналогично режимы нормального и инверсного насыщения отличаются принципом управления прямым смещением на базовом p - n -переходе при первоначальном обеспечении соответствующего обратного или фиксированного прямого смещения того или иного p - n -перехода.

Большинство электрических моделей БТ для большого сигнала, применяемых в программах автоматизированного расчета фрагментов БИС, синтезировано на основе секционного подхода разбиения их на области квазинейтральности и объемного заряда. Для каждой из областей используются аналитические решения одномерной ФСУ, а объединение в полную электрическую модель осуществляется сопряжением на p - n -переходах. Причем для односекционных классических моделей Эберса — Молла, Линвилла, Бьюфой — Спаркса, несмотря на различный уровень описания электрофизических процессов в характерных областях БТ, наблюдается полная самосогласованность и одинаковая погрешность расчета статических и динамических параметров БТ в режимах низких и средних уровней инжекции. Как следствие прогресса технологии БИС при создании БТ с микротопологией возникла необходимость разработки более точных и эффективных электрических моделей БТ на большом и малом сигналах в режимах малых времен переключения и высоких уровней инжекции для автоматизированного расчета быстродействующих схем.

Развитие электрических моделей БТ БИС и С³БИС с учетом их микротопологии и распределенности, специфики работы в режимах высоких плотностей токов, насыщения и квазинасыщения, в диапазоне изменения внешних температур не должно идти по пути излишней сложности и детализации описания работы, приводить к неоправданно большим затратам машинного времени соответствующих расчетов схем на ЭВМ. Известно, что использование сложных многомерных распределенных моделей БТ (не учитывая к тому же многие эффекты высокого уровня инжекции) для расчета быстродействующих схем требует значительных вычислительных ресурсов. Эффективные распределенные модели БТ

с двумя-тремя секциями как компромиссный вариант достижения приемлемой точности — сложности расчета схем синтезировались в основном на основе простейших (полуэмпирических) предпосылок и допущений и не обеспечивали удовлетворительной погрешности моделирования БИС, в первую очередь субнаносекундного диапазона. В качестве узловой проблемы отметим необходимость синтеза эффективных распределенных моделей БТ с одной-двумя секциями с учетом влияния микротопологии и эффектов высокого уровня инжекции применительно к расчету субнаносекундных фрагментов БИС в широком режимном диапазоне, обеспечивающих погрешность электронного моделирования, равную 10—20%.

Кроме большой актуальности разработки точных иерархических электрических моделей БТ, ориентированных на расчет быстродействующих и сверхбыстродействующих фрагментов БИС, наиболее остро встает вопрос об экспериментальной и машинной идентификации электрических параметров электрических моделей БТ. Несмотря на неоспоримую значимость и необходимость экспериментальных методов идентификации и параметров БТ, все большую актуальность приобретают методы и средства решения проблемы машинной идентификации параметров приборных структур как на основе настройки модельных и экспериментальных зависимостей объекта, так и на основе специфических методов многоуровневого моделирования и идентификации параметров на ЭВМ.

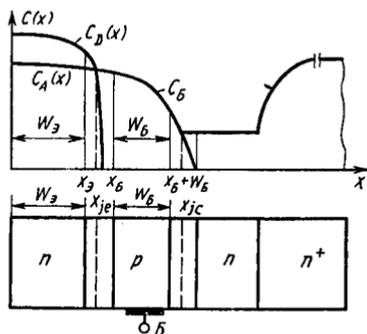


Рис. 6.1. Профиль распределения примесей и характерные области одномерного БТ

6.2. ОБОБЩЕННЫЕ ВЫРАЖЕНИЯ ДЛЯ ТОКОВ И КОЭФФИЦИЕНТА ПЕРЕДАЧИ ТОКА ИНТЕГРАЛЬНЫХ БТ

Обобщенные выражения для токов. Задача синтеза электрической модели БТ заключается в установлении фундаментальных связей между электрофизическими внутренними параметрами БТ и его выходными электрическими характеристиками. Чтобы получить обобщенное соотношение для токов коллектора и базы интегрального БТ типа *n-p-n* с профилем и характерными областями в одномерном приближении (рис. 6.1), используем соотношения для плотностей токов и концентраций, выведенные в гл. 4:

$$J_n = q\mu_n n \frac{\partial \varphi_n}{\partial x}, \quad (6.1)$$

$$J_p = -q\mu_p p \frac{\partial \varphi_p}{\partial x}, \quad (6.2)$$

$$n = n_{ie} \exp[(V - \varphi_n)/\varphi_T], \quad (6.3)$$

$$p = n_{ie} \exp[(\varphi_p - V)/\varphi_T]. \quad (6.4)$$

В общем случае при выводе обобщенных соотношений для токов необходимо использовать также уравнения непрерывности, которые в одномерном приближении к интегральной форме записываются в виде

$$J_n(x_2) - J_n(x_1) = \int_{x_1}^{x_2} q |R| dx, \quad (6.5)$$

$$J_p(x_2) - J_p(x_1) = - \int_{x_1}^{x_2} q |R| dx. \quad (6.6)$$

Производную по координате x от произведения

$$pn = n_{ie}^2 \exp \frac{1}{\varphi_T} (\varphi_p - \varphi_n) \approx n_{i0} \exp \frac{\Delta E_g}{kT} \quad (6.7)$$

можно представить в виде

$$\begin{aligned} \frac{d}{dx} (pn) &= \frac{n_{ie}^2}{\varphi_T} \exp(\varphi_p - \varphi_n) \left(\frac{\partial \varphi_p}{\partial x} - \frac{\partial \varphi_n}{\partial x} \right) = \\ &= \frac{np}{\varphi_T} \left(\frac{\partial \varphi_p}{\partial x} - \frac{\partial \varphi_n}{\partial x} \right). \end{aligned} \quad (6.8)$$

Подставляя (6.1), (6.2) в (6.8), интегралы в произвольной области БТ от x_1 до x_2 правой и левой частей соотношения (6.8) можно выразить следующим образом:

$$\int_{x_1}^{x_2} \frac{np}{\varphi_T} \left(\frac{\partial \varphi_p}{\partial x} - \frac{\partial \varphi_n}{\partial x} \right) dx = \int_{x_1}^{x_2} \frac{np}{kT} \left[-\frac{J_p}{\mu_p p} + \frac{J_n}{\mu_n n} \right] dx, \quad (6.9)$$

$$\int_{x_1}^{x_2} \frac{d}{dx} (pn) dx = (pn)|_{x_2} - (pn)|_{x_1}. \quad (6.10)$$

Проводя интегрирование выражения (6.8) по нейтральной области эмиттера, т. е. от 0 до $x_{\text{Э}}$, изменяя пределы интегрирования в (6.9), (6.10) от 0 до $x_{\text{Э}}$ и используя уравнение (6.6), а также граничные условия в точках $x = x_{\text{Э}}$, $x = 0$ на поверхности эмиттера

$$\frac{n(x_{\text{Э}}) p(x_{\text{Э}})}{n_{ie}^2(x_{\text{Э}})} = \exp \frac{U_{\text{Э}}}{\varphi_T}, \quad (6.11)$$

$$p(0) - p_{n0}(0) = p(0) = \frac{J_p(0)}{qS_{\text{ЭК}}}, \quad (6.12)$$

где $s_{\text{ЭК}}$ — скорость рекомбинации на поверхности эмиттерного контакта, для тока дырок $I_p(x_{\text{Э}}) \approx J_p(x_{\text{Э}}) S_{\text{Э}}$ ($S_{\text{Э}}$ — площадь активной области эмиттера) получим выражение

$$I_p(x_{\text{Э}}) = S_{\text{Э}} q D_{p0} n_{i0}^2 \exp \frac{U_{\text{Э}}}{Q_{\text{Э}} \varphi_{\text{Ф}}}. \quad (6.13)$$

В (6.13) $Q_{\text{Э}} \varphi_{\text{Ф}}$ — эффективный параметр нейтральной области эмиттера:

$$Q_{\text{Э}} \varphi_{\text{Ф}} = \frac{n(0) D_{p0}}{s_{\text{ЭК}}} \left(\frac{n_{i0}}{n_{i0}(0)} \right)^2 \left(1 - \frac{1}{J_p(x_{\text{Э}})} \int_0^{x_{\text{Э}}} q |R| dx \int_0^{x_{\text{Э}}} \left(\frac{D_{p0}}{D_p(x)} \right) \times \right. \\ \left. \times \left(\frac{n_{i0}}{n_{ie}(x)} \right)^2 n(x) \left(1 - \frac{1}{J_p(x_{\text{Э}})} \int_x^{x_{\text{Э}}} q |R| dx \right) dx, \quad (6.14)$$

где D_{p0} , n_{i0} — коэффициент диффузии дырок для низкой концентрации примесей и концентрация свободных носителей заряда; $n(x)$, $D_p(x)$, $n_{ie}(x)$ — электрофизические параметры, зависящие от координаты.

Отметим, что уравнения (6.13), (6.14) выведены при соотношении $n(x) \gg p(x)$ для подвижных носителей заряда. Для низких уровней инжекции концентрации электронов $n(0)$, $n(x)$ в (6.14) можно представить как концентрации донорной примеси $n(0) \approx C_D(0)$; $n(x) = C_D(x)$.

Суть первого граничного условия на эмиттерном переходе (6.11) раскрывается с помощью соотношения (6.7) при допущении постоянства квазипотенциала (квазиуровня) Ферми неосновных носителей заряда в эмиттерной области, причем

$$U_{\text{ЭБ}} = U_{\text{Э}} = \varphi_p(0) - \varphi_n(0). \quad (6.15)$$

Суть второго граничного условия на эмиттерном контакте заключается в конечной скорости поверхностной рекомбинации $s_{\text{ЭК}}$, не равной бесконечности: $s_{\text{ЭК}} \neq \infty$. При допущении $s_{\text{ЭК}} \rightarrow \infty$ граничное условие (6.12) упрощается до стандартного условия идеального контакта $p(0) \approx 0$.

Перейдем к рассмотрению обобщенного выражения тока коллектора БТ. Для этого необходимо выполнить интегрирование (6.8) по области базы, т. е. от $x = x_{\text{Б}}$ до $x = x_{\text{Б}} + W_{\text{Б}}$, с использованием (6.9), (6.10), (6.5). Граничными условиями интегрирования в области квазинейтральной базы служат выражения

$$\frac{n(x_{\text{Б}}) p(x_{\text{Б}})}{n_{ie}^2(x_{\text{Б}})} = \exp \frac{u_{\text{Э}}}{\varphi_{\text{Г}}}, \quad (6.16)$$

$$\frac{n(x_{\text{Б}} + W_{\text{Б}}) p(x_{\text{Б}} + W_{\text{Б}})}{n_{ie}^2(x_{\text{Б}} + W_{\text{Б}})} = \exp \frac{U_{\text{К}}}{\varphi_{\text{Г}}}, \quad (6.17)$$

являющиеся следствием допущения постоянства квазипотенциала (квазиуровня) Ферми неосновных носителей заряда в базовой области при выводе соотношений для напряжений на эмиттерном $U_{\text{Э}}$ и коллекторном $U_{\text{К}}$ p - n -переходах:

$$U_{\text{Э}} = U_{\text{ЭБ}} = \varphi_p(x_{\text{Б}}) - \varphi_n(x_{\text{Б}}), \quad (6.18)$$

$$U_{\text{К}} = U_{\text{КБ}} = \varphi_p(x_{\text{Б}} + W_{\text{Б}}) - \varphi_n(x_{\text{Б}} + W_{\text{Б}}). \quad (6.19)$$

После выполнения интегрирования с учетом граничных условий на переходах получим ВАХ тока электронов, практически равного току коллектора БТ:

$$I_{\text{К}} \approx I_n(x_{\text{Б}}) = - \frac{q D_{n0} n_{i0}^2 S_{\text{Э}} \left[\exp \frac{U_{\text{Э}}}{\varphi_T} - \exp \frac{U_{\text{К}}}{\varphi_T} \right]}{Q_{\text{Б эф}}}, \quad (6.20)$$

где $Q_{\text{Б эф}}$ — эффективный параметр квазинейтральной области базы:

$$Q_{\text{Б эф}} = \int_{x_{\text{Б}}}^{x_{\text{Б}} + W_{\text{Б}}} \frac{D_{n0}}{D_n(x)} \left(\frac{n_{i0}}{n_{ie}(x)} \right)^2 p(x) \times \\ \times \left(1 + \frac{1}{J_n(x_{\text{Б}})} \int_{x_{\text{Б}}}^x q |R| dx \left(1 - \frac{J_p(x) \mu_n(x) n(x)}{J_n(x) \mu_p(x) p(x)} \right) \right) dx. \quad (6.21)$$

Обычно в квазинейтральной базе выполняется условие $J_p(x) \times n(x) \mu_n(x) \ll J_n(x) p(x) \mu_p(x)$ и выражение $Q_{\text{Б эф}}$ (6.21) упрощается:

$$Q_{\text{Б эф}} = \int_{x_{\text{Б}}}^{x_{\text{Б}} + W_{\text{Б}}} \left(\frac{D_{n0}}{D_n(x)} \right) \left(\frac{n_{i0}}{n_{ie}(x)} \right)^2 p(x) \left(1 + \frac{1}{J_n(x_{\text{Б}})} \int_{x_{\text{Б}}}^x q |R| dx \right) dx. \quad (6.22)$$

При пренебрежимо малом влиянии рекомбинации в базе выражение (6.21) имеет еще более простой вид:

$$Q_{\text{Б эф2}} = \int_{x_{\text{Б}}}^{x_{\text{Б}} + W_{\text{Б}}} \frac{D_{n0}}{D_n(x)} \left(\frac{n_{i0}}{n_{ie}(x)} \right)^2 p(x) dx, \quad (6.23)$$

что позволяет записать результирующий ток коллектора БТ

$$I_{\text{К}} = - \frac{q S_{\text{Э}} \varphi_T \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - \exp \frac{U_{\text{К}}}{\varphi_T} \right)}{\int_{x_{\text{Б}}}^{x_{\text{Б}} + W_{\text{Б}}} \frac{p(x)}{\mu_n(x) n_{ie}^2(x)} dx}. \quad (6.24)$$

Для низких уровней инжекции в формуле (6.24) концентрация дырок может быть заменена на концентрацию акцепторной примеси $\rho(x) \approx C_A(x)$.

Коэффициент передачи тока одномерного БТ. Обозначив γ эффективность эмиттера, α_r — коэффициент переноса, коэффициент передачи тока БТ, включенного по схеме с общим эмиттером, можно записать в виде

$$B = \gamma \alpha_r / (1 - \gamma \alpha_r). \quad (6.25)$$

С использованием выражений (6.13), (6.14) и (6.20), (6.21) эффективность эмиттера в активном нормальном режиме определяется следующим образом:

$$\gamma = \frac{I_n(x_B)}{I_n(x_B) + I_p(x_E)} = \frac{1}{1 + \frac{D_{p0}}{D_{n0}} \frac{Q_{B\text{эф}}}{Q_{E\text{эф}}}}, \quad (6.26)$$

а коэффициент передачи тока БТ

$$\alpha_r = \frac{J_n(x_B + W_B)}{J_n(x_B)} = 1 - \frac{1}{|J_n(x_B)|} \int_{x_B}^{x_B + W_B} q |R| dx. \quad (6.27)$$

В общем случае при подстановке выражений (6.26), (6.27) в (6.25) коэффициент передачи тока

$$B = 1 - |J_n(x_B)|^{-1} \int_{x_B}^{x_B + W_B} q |R| dx \times \left[\frac{D_{p0} Q_{B\text{эф}}}{D_{n0} Q_{E\text{эф}}} + |J_n(x_B)|^{-1} \int_{x_B}^{x_B + W_B} q |R| dx \right]^{-1}. \quad (6.28)$$

Несмотря на некоторую громоздкость обобщенных выражений для тока коллектора I_K (6.20), тока базы $I_B \approx I_p(x_E)$ (6.13) и коэффициента передачи тока B (6.28), данные соотношения в целом справедливы для произвольных профилей распределения примесей, уровней инжекции, электрических режимов БТ и являются своего рода эквивалентом численного ФТ-моделирования БТ в одномерном приближении. Однако проблемы выбора границ интегрирования в данных выражениях (рис. 6.1), зависящих от уровня инжекции и профилей распределения примесей, описания электрофизических параметров D , R , n_{ie} через координаты и концентрации, довольно серьезны и сложны для аналитических методов расчетов характеристики БТ и не могут быть конкурентоспособными для численного ФТ-моделирования (даже в одномерном приближении) при требовании минимальной погрешности расчетов. Однако для качественных оценок и расчетов рассмотренные обобщенные выражения для

токов коллектора, базы и коэффициента передачи тока могут быть весьма полезны при рассмотрении БТ в упрощенных ситуациях.

Предположим, что скорости объемной рекомбинации в областях базы и эмиттера можно представить как

$$R = -\Delta n(x)/\tau_{nБ}, \quad (6.29)$$

$$R = -\Delta p(x)/\tau_{pЭ}, \quad (6.30)$$

где $\tau_{nБ}$, $\tau_{pЭ}$ — средние времена жизни неосновных носителей заряда (электронов, дырок) в базе и эмиттере БТ; $\Delta n(x)$, $\Delta p(x)$ — концентрации избыточных неосновных носителей в базе и эмиттере соответственно.

При сделанных допущениях (6.29), (6.30), а также при выводе формулы (6.23) коэффициент передачи тока B (6.28) может быть записан в виде

$$B = \frac{1 - \tau_{Б}/\tau_{nБ}}{\frac{D_{p0}Q_{БЭФ2}}{D_{n0}Q_{ЭЭФ2}} + \frac{\tau_{Б}}{\tau_{nБ}}}, \quad (6.31)$$

$$\text{где } Q_{БЭФ2} = \int_{x_{Б}}^{x_{Б}+W_{Б}} \left(\frac{D_{n0}}{D_n(x)}\right) \left(\frac{n_{i0}}{n_{ie}(x)}\right)^2 p(x) \left(1 - \frac{t_{Б}(x)}{\tau_{nБ}}\right) dx, \quad (6.32)$$

$$Q_{ЭЭФ2} = \frac{n(0)D_{p0}}{s_{ЭК}} \left(1 - \frac{t_{Э}}{\tau_{pЭ}}\right) + \int_0^{x_{Э}} \frac{D_{p0}}{D_p(x)} \left(\frac{n_{i0}}{n_{ie}(x)}\right)^2 n(x) \left(1 - \frac{t_{Э}(x)}{\tau_{pЭ}}\right) dx. \quad (6.33)$$

Эффективные времена переноса основных носителей заряда через области базы и эмиттера в (6.32), (6.33) определяются из общей зарядовой концепции ($Q \approx I\tau$):

$$\tau_{Б} = \frac{q}{|J_n(x_{Б})|} \int_{x_{Б}}^{x_{Б}+W_{Б}} \Delta n(x) dx = \frac{Q_{nБ}}{|J_n(x_{Б})|}, \quad (6.34)$$

$$\tau_{Э} = \frac{q}{|J_p(x_{Э})|} \int_0^{x_{Э}} \Delta p(x) dx = \frac{Q_{pЭ}}{|J_p(x_{Э})|}, \quad (6.35)$$

Для получения полезных оценочных соотношений для B , $\tau_{Б}$, $\tau_{Э}$ примем допущения равномерно легированного профиля БТ (рис. 6.1) с концентрациями в эмиттере $C_{Э}$, базе $C_{Б}$, соответствующими средними коэффициентами диффузии $D_{nБ}$, $D_{pЭ}$, временами жизни $\tau_{nЭ}$, $\tau_{pЭ}$ (диффузионными длинами неосновных носителей заряда в эмиттере и базе соответственно $L_{pЭ} = (D_{pЭ} \tau_{pЭ})^{1/2}$, $L_{nБ} = (D_{nБ} \tau_{nБ})^{1/2}$).

Тогда для бесконечной скорости рекомбинации на эмиттерном контакте ($s_{\text{ЭК}} \rightarrow \infty$) и тонких слоев эмиттера и базы $L_{p\varepsilon} \gg W_{\varepsilon}$, $L_{nБ} \gg W_{\text{Б}}$ максимальный коэффициент передачи (6.31) запишется так:

$$B_{\text{max1}} = \frac{D_{nБ}}{D_{p\varepsilon}} \frac{C_{\varepsilon}}{C_{\text{Б}}} \frac{W_{\varepsilon}}{W_{\text{Б}}} \exp\left(-\frac{\Delta E_{gБ\varepsilon}}{kT}\right), \quad (6.36)$$

где

$$\exp\frac{\Delta E_{gБ\varepsilon}}{kT} = \left(\frac{n_{i\varepsilon}}{n_{iБ}}\right)^2 \simeq \left(\frac{n_{i\varepsilon}}{n_{i0}}\right)^2, \quad (6.37)$$

$\Delta E_{gБ\varepsilon} = \Delta E_{g\varepsilon} - \Delta E_{gБ}$ — сужение ширины запрещенной зоны в сильнолегированном эмиттере относительно зоны в базе; обычно $\Delta E_{gБ\varepsilon} \approx \Delta E_g$ при $n_{i\varepsilon} \approx n_{i\varepsilon}$, $n_{i0} \approx n_{iБ}$.

В рассматриваемом случае

$$\tau_{\varepsilon 1} = W_{\varepsilon}^2 / 2D_{p\varepsilon}, \quad \tau_{\text{Б1}} = W_{\text{Б}}^2 / 2D_{nБ}. \quad (6.38)$$

Для конечной скорости рекомбинации на эмиттерном контакте и тонких слоев эмиттера и базы $W_{\varepsilon} / \tau_{p\varepsilon} \leq s_{\text{ЭК}} < \infty$, $L_{nБ} \gg W_{\text{Б}}$, $L_{p\varepsilon} \gg W_{\varepsilon}$, т. е. для БТ с «прозрачным» эмиттером, из выражений (6.31) — (6.33) имеем

$$B_{\text{max2}} = \left(\frac{D_{nБ}}{s_{\text{ЭК}} W_{\text{Б}}}\right) \frac{C_{\varepsilon}}{C_{\text{Б}}} \exp\left(-\frac{\Delta E_{gБ\varepsilon}}{kT}\right), \quad (6.39)$$

$$\tau_{\varepsilon 2} = \frac{W_{\varepsilon}^2}{2D_{p\varepsilon}} + \frac{W_{\varepsilon}}{s_{\text{ЭК}}}, \quad \tau_{\text{Б2}} = \frac{W_{\text{Б}}^2}{2D_{nБ}}. \quad (6.40)$$

Для малой скорости рекомбинации на эмиттерном контакте и «толстой» структуры

$$B_{\text{max3}} = \left(\frac{D_{nБ}}{D_{p\varepsilon}}\right) \left(\frac{C_{\varepsilon} W_{\varepsilon}}{C_{\text{Б}} W_{\text{Б}}}\right) \left(\frac{L_{p\varepsilon}}{W_{\varepsilon}}\right)^2. \quad (6.41)$$

Для БТ БИС практически важно рассмотрение первого и второго случаев тонкослойных БТ, для которых $B_{\text{max2}} \gg B_{\text{max1}}$.

Для обеспечения максимально возможного коэффициента B необходимо соблюдение условия $C_{\varepsilon} W_{\varepsilon} \gg C_{\text{Б}} W_{\text{Б}}$, т. е. большой разницы между граничными концентрациями примесей в эмиттерном и базовом слоях и сверхтонкой ширины базы ($W_{\varepsilon} \gg W_{\text{Б}}$). Поэтому в принципе при изготовлении высококачественных БТ эмиттерный слой стремятся легировать значительно сильнее базового и в то же время обеспечить минимальную ширину базового слоя. Однако сильное легирование эмиттера связано со значительным сужением ширины запрещенной зоны ΔE_g , что приводит к некоторой потере усилительных свойств стандартных БТ. При разработке БТ сверхскоростных БИС возрастают требования к сильнолегированности базы $C_{\text{Б}}$ для уменьшения сопротивления базы. При этом происходит частичная компенсация уменьшения $\Delta E_{gБ\varepsilon}$ за счет уменьшения $\Delta E_{gБ}$, когда $\Delta E_{gБ}$ (БТ С³БИС) $\ll \Delta E_g$ (стандартные БТ), что в це-

лом является благоприятным фактором возможности увеличения коэффициента B тонкослойных сильнолегированных БТ. Уменьшение эффективной скорости рекомбинации $S_{\text{ЭК}}$, как следует из выражения (6.39), также приводит к увеличению B_{max} БТ. Это достигается с помощью физико-технологических принципов реализации БТ с «прозрачным» эмиттером и поликремниевым эмиттерным контактом, где $10^3 \leq S_{\text{ЭК}} \leq 5 \cdot 10^4$ см/с.

Приведенные выражения для максимальных коэффициентов передачи тока (6.36), (6.39), (6.41), несмотря на большое количество упрощающих допущений, сделанных при их выводе, полезны для инженерных оценок предельных вариантов конкурирующих технологий изготовления БТ, поскольку позволяют оценить «идеальный» параметр гипотетического БТ без потенциальной деградации усиления в области малых и больших токов в реальных электрических режимах работы БТ в БИС. Для оценок реальных потерь усиления и быстродействия БТ в произвольных электрических режимах требуется не только детальное раскрытие идеальных компонент базового и коллекторного токов одномерной модели БТ, но и количественное описание эффектов микротопологии и распределенности (двумерных и трехмерных), высокого уровня инжекции и ряда других электрофизических процессов в БТ.

6.3. ИНТЕГРАЛЬНОЕ СООТНОШЕНИЕ УПРАВЛЕНИЯ ЗАРЯДОМ ДЛЯ СИНТЕЗА ЭЛЕКТРИЧЕСКИХ МОДЕЛЕЙ БТ

Суть интегрального зарядового соотношения. Соотношение для тока электронов (6.24), выведенное при допущениях: 1) $J_p(x) \mu_n(x) \times n(x) \ll J_n(x) \mu_p(x) p(x)$; 2) $\int q|R|dx \rightarrow 0$, — может быть переписано с заменой пределов интегрирования знаменателя

(рис. 6.1):

$$\int_{x_B}^{x_B + W_B} \text{ на } \int_0^{x_K}$$

$$I_K = \frac{q\varphi_T \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - \exp \frac{U_K}{\varphi_T} \right)}{\int_0^{x_K} S_{\text{Э}}(p, x) \frac{p(x)}{\mu_n(x) n_{ie}^2(x)} dx} \quad (6.42)$$

Интегрирование по всему объему БТ от эмиттерного контакта до n^+ -коллектора — эквивалента коллекторного контакта — для большинства используемых режимов практически не изменяет суммарного тока коллектора. Отличие в значениях (6.24), (6.42), которое будет обсуждено ниже, связано с эффектами расширения базы, насыщения и квазинасыщения, когда нельзя пренебрегать зарядом дырок в области коллектора, примыкающей к металлургической границе коллекторного перехода. Влияние концентрации дырок в

выражениях (6.24), (6.42) наиболее сильно в области базы, не столь заметно в эмиттере и пренебрежимо мало в n^+ -коллекторе и даже в эпитаксиальном n -слое в активном нормальном режиме. Вследствие того что протекающие токи создают заметные падения напряжения на омических областях БТ, предполагается, что напряжения $U_{\text{Э}}$, $U_{\text{К}}$ прикладываются непосредственно к эмиттерному и коллекторному переходам, учет падений напряжений приводит к модификациям $U_{\text{ЭК}} = U_{\text{ЭК}} - I r$. Выражение (6.42) носит название интегрального соотношения управления зарядом БТ и является фундаментальным для синтеза электрических моделей БТ при произвольных уровнях инжекции и профилей распределения примесей. Включение в подынтегральное выражение переменной функции площади эмиттера (активной области) БТ $S_{\text{Э}}(p, x)$ в общем случае также позволяет отразить модуляцию эффективной площади эмиттера вследствие эффекта вытеснения тока, приводящего к уменьшению $S_{\text{Э}}(p, x) \leq S_{\text{Э}}$ при высоких уровнях инжекции.

В целях дальнейшего практического использования выражение (6.42) может быть переписано для произвольного электрического режима БТ:

$$I_{\text{К}} = \frac{C_1}{Q_p} \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - \exp \frac{U_{\text{К}}}{\varphi_T} \right), \quad (6.43)$$

где

$$C_1 = q^2 S_{\text{Э}}^2 \varphi_T (\mu_n n_{ie}^2) = q^2 S_{\text{Э}}^2 (D_n n_{ie}^2), \quad (6.44)$$

$$Q_p = q S_{\text{Э}} \int_0^{x_{\text{К}}} p(x) dx. \quad (6.45)$$

Соотношения (6.43)—(6.45) служат основой электрической модели управления зарядом Гуммеля—Пуна, причем в первоначальной версии полагалось

$$C_1 = \text{const}, \quad \int_0^{x_{\text{К}}} p(x) dx = \int_{x_{\text{К}}}^{x_{\text{Б}} + W_{\text{Б}}} p(x) dx. \quad (6.46)$$

Справедливость равенств (6.46), как увидим далее, весьма относительна для ряда электрических режимов и профилей БТ, и канонизация (6.46) для всех рассматриваемых случаев является источником некоторой погрешности синтезируемых электрических моделей. Однако огромным преимуществом записи интегрального зарядового соотношения в виде (6.43), (6.44) по сравнению с обобщенными соотношениями (6.20)—(6.24), (6.42) является непосредственное включение в него заряда дырок в виде интеграла (6.45).

Поскольку инерционные особенности поведения БТ в переходных процессах или ВЧ- и СВЧ-режимах малого сигнала определяются изменением во времени интегрального заряда в БТ Q_p (6.45), интегральное зарядовое соотношение (6.43) является очень удоб-

ной, эффективной формой описания инерционных характеристик БТ и выявления структурной связи между стационарными и нестационарными характеристиками БТ в произвольных электрических режимах, включая режим высоких уровней инжекции (в базе, коллекторе). Благодаря четкой детализации интегрального заряда Q_p (6.45) в синтезируемых электрических моделях удается в достаточно простой и компактной форме учесть эффекты высоких уровней инжекции: расширения базы в области коллектора (эффект Кирка) и эмиттера, модуляции проводимости базы (эффект Вэбстера), модуляции ширины базы (эффект Эрли). Наряду с эффектом вытеснения эмиттерного тока, описываемым также соотношением (6.42), названные эффекты ответственны за деградацию усиления и быстродействия БТ в режиме высоких плотностей тока, необходимом для реализации максимального потенциального быстродействия элементов и фрагментов БИС на БТ.

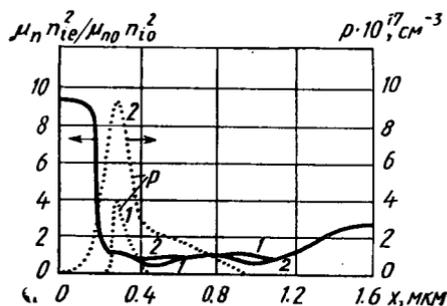


Рис. 6.2. Зависимости нормированного произведения подвижности электронов на собственную концентрацию носителей заряда, а также концентрации дырок от координаты для различных уровней инжекции: $J = 10^3$ А/см² (кривая 1); $J = 4 \times 10^4$ А/см² (кривая 2)

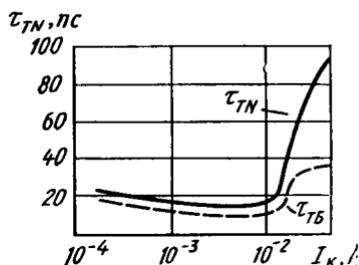


Рис. 6.3. Зависимости составляющих времени задержки эмиттер — коллектор БТ в активном нормальном режиме от тока коллектора

Проверка интегрального зарядового соотношения с помощью ФТ-моделирования БТ. Обобщая все экспериментальные и машинные ФТ-исследования различных типов БТ в одномерном приближении, можно констатировать, что описание характеристик и параметров БТ на основе зарядовых соотношений (6.20) — (6.24), (6.42) и строгого численного ФТ-моделирования практически адекватны. Погрешность аналитического расчета (0,5—5%) обусловлена упрощающими допущениями, сделанными при переходе от выражения (6.20) к (6.22) — (6.24), (6.42). В то же время основным источником погрешности зарядового соотношения (6.43) относительно (6.42) является использование упрощающих допущений в (6.45). В первую очередь погрешность обусловлена увеличением среднего значения $\mu_n n_{ie}^2$ с ростом тока относительно уровня

$C_1/q^2 S_{\Sigma}^2 \varphi_T$ в (6.44) при низких уровнях инжекции. С другой стороны, погрешность обусловлена сильным легированием областей БТ, когда для сильнолегированных областей эмиттера и эмиттерного перехода отношение $\mu_n(x) n_{ie}^2(x) / \mu_{n0} n_{i0}^2 > 1$ (рис. 6.2); под $\mu_{n0}, n_{i0} = \text{const}$ понимают параметры низкого уровня легирования полупроводника. Токовая зависимость C_1 (6.44) является результатом увеличения n_{ie} и уменьшения μ_n с повышением уровня легирования. Распределение параметра — нормированного произведения $\mu_n n_{ie}^2 / \mu_{n0} n_{i0}$ и концентрации дырок в объеме БТ — совмещено для исследования влияния эффектов высокого уровня инжекции (высокой плотности режимного тока, равной примерно $4 \cdot 10^4$ А/см²). Для рассматриваемого высокого уровня инжекции (кривая 2) увеличивается заряд дырок, накапливаемый в областях квазинейтрального эмиттера и эмиттерного p - n -перехода. Так как $n_{i\Omega} \gg n_{iB}$, тем самым значительно компенсируется уменьшение подвижности основных носителей заряда при увеличении концентрации донорной примеси в эмиттере. Данный факт приводит к некоторому увеличению параметра C_1 с ростом режимного тока. Накопление дырок в области эпитаксиального коллектора также увеличивается в режиме высокого уровня инжекции. Однако противоположно эмиттерной области произведение $\mu_n n_{ie}^2$ в эпитаксиальном слое отличается не так сильно от данного произведения в базе, чтобы давать заметный вклад в токовую зависимость параметра зарядового соотношения C_1 (6.44).

Использование в зарядовом соотношении (6.43) вместо Q_p (6.45) параметра $Q_{pB} = \int_{x_B}^{x_B+W_B} p(x) dx$ отражено в некотором изменении зависимостей $\tau_{TN} = dQ_p/dI$ (рис. 6.3) для достаточно тонкослойного ($W_{\Sigma} = W_B = 0.2$ мкм) и быстродействующего ($f_{\tau} \approx 5 \div 9$ ГГц) БТ.

Параметры τ_{TN} и τ_{TB} определяются соответственно от соотношений $\tau_{TN} = dQ_p/dI$, $\tau_{TB} = dQ_{pB}/dI$. В режиме низких уровней инжекции $Q_p = Q_{pB}$, поэтому различия в параметрах τ_{TN} , τ_{TB} обусловлены только зарядами в эмиттерном переходе. Однако уже в режиме средних уровней инжекции $Q_{pB} < Q_p$ влияние подвижных носителей заряда в области эмиттера и коллектора становится ощутимым, коэффициент B начинает заметно уменьшаться. При высоких уровнях инжекции вклад заряда дырок в эмиттерной и коллекторной эпитаксиальных областях оказывается все более доминирующим, что приводит как к более заметным изменениям параметров B , τ_{TN}/τ_{TB} , так и к существенной погрешности аппроксимации (6.43) — (6.46). Однако даже для высоких плотностей тока $(1-5) \cdot 10^4$ А/см² погрешность аппроксимации (6.43) — (6.45) не превышает 10—15%.

6.4. ОСНОВЫ ПОСТРОЕНИЯ ЭЛЕКТРИЧЕСКИХ ИЕРАРХИЧЕСКИХ МОДЕЛЕЙ БТ С УПРАВЛЕНИЕМ ЗАРЯДОМ

Принцип построения. Электрические модели БТ с управлением зарядом основаны на уравнениях (6.42), (6.43), связывающих напряжения на p - n -переходах БТ, коллекторный ток и полный заряд дырок БТ типа n - p - n . В частном случае в модели можно использовать заряд дырок в базе Q_{pB} , однако заметим, что Q_{pB} сложнее идентифицируется с помощью как экспериментальных, так и численных ФТ-исследований ввиду трудности определения границ интегрирования x_B и $x_B + W_B$. Задача моделирования инерционных и высокоинжекционных особенностей БТ сводится практически к моделированию полного заряда в БТ (в частном случае в базе). Сначала рассмотрим процессы в активном нормальном режиме, где полный заряд в БТ определяется суммой

$$Q_p = Q_{p0} + Q_{\Delta} + Q_K + Q_f. \quad (6.47)$$

В формуле (6.47) Q_{p0} определяет заряд дырок в состоянии термодинамического равновесия $U_{\Delta} = U_K = 0$, зависящий только от распределения примесей в базе, Q_f отражает увеличение Q_p с возрастанием режимного тока, т. е. накопленный избыточный заряд дырок, который равен по абсолютной величине концентации неосновных носителей заряда во всех областях БТ.

С зарядом Q_f связано время переноса в активном нормальном режиме τ_{TN} :

$$Q_f = \int_0^{I_K} \tau_{TN} dI, \quad (6.48)$$

учитывающее в общем случае перенос носителей заряда не только через базу (τ_{TB}), но также через эмиттер, эмиттерный, коллекторный переходы и «индуцированную, расширенную» базу в собственно эпитаксиальный слой БТ: $\tau_{TN} \geq \tau_{TB}$. Согласно определению (6.48), параметр τ_{TN} является малосигнальным и может быть определен через параметр f_T , т. е. граничную частоту БТ, на которой модуль коэффициента передачи равен единице; $\tau_{TN} = 1/(2\pi f_T)$ (при экстраполяции в область больших режимных токов без потери усиления и быстродействия БТ). Заряды Q_{Δ} и Q_K в (6.47) определяют заряды ионов в эмиттерном и коллекторном p - n -переходах соответственно при $U_K = 0$ и $U_{\Delta} = 0$:

$$Q_{\Delta} = \int_0^{U_{\Delta}} C_{je} dU_{\Delta}, \quad (6.49)$$

$$Q_K = \int_0^{U_K} C_{jc} dU_K, \quad (6.50)$$

где C_{je} , C_{jc} — барьерные емкости p - n -переходов.

Для однозначности описания Q_{Σ} , Q_K , Q_f отметим, что при прямом смещении в области p - n -перехода кроме заряда ионов имеется заряд подвижных носителей, который будем приписывать не Q_{Σ} , Q_K , а Q_f . Заряд подвижных носителей накапливается не только в базе БТ (Q_{f3}), но также в эмиттере (Q_{f1}), эмиттерном переходе (Q_{f2} — так называемый заряд нейтральной емкости эмиттерного прямосмещенного p - n -перехода), коллекторном переходе (Q_{f4}), коллекторном эпитаксиальном слое (Q_{f5}).

Полный заряд подвижных носителей может быть представлен в виде суммы с соответствующим описанием отдельных зарядов произведением тока коллектора на характеристическое время доминирующего процесса в области.

При этом заряд подвижных носителей в эмиттере

$$Q_{f1} = \tau_{\Sigma} I_K, \quad (6.51)$$

где τ_{Σ} — время переноса носителей в эмиттере (6.35), зависящее от толщины эмиттера и коэффициента диффузии D_p .

Заряд подвижных носителей в эмиттерном p - n -переходе

$$Q_{f2} = \tau_{pn\Sigma} I_K, \quad (6.52)$$

где $\tau_{pn\Sigma}$ — время накопления подвижных носителей в эмиттерном p - n -переходе, зависящее от градиента распределения и концентрации примесей.

Заряд подвижных носителей в базе

$$Q_{f3} = \tau_{TB} I_K, \quad (6.53)$$

где τ_{TB} — время переноса носителей в базе (6.34), зависящее от толщины базы и коэффициента диффузии D_n .

Заряд подвижных носителей в p - n -переходе

$$Q_{f4} = \tau_{pnK} I_K, \quad (6.54)$$

где τ_{pnK} — время накопления подвижных носителей в коллекторном p - n -переходе.

Наконец, заряд в коллекторном эпитаксиальном слое

$$Q_{f5} = \tau_K I_K. \quad (6.55)$$

Отметим, что если для описания емкости C_{jc} (6.50) используется стандартное аналитическое вольт-фарадное соотношение, то значение величины Q_K в формуле (6.50) может быть несколько меньше из-за падения напряжения на достаточно высокоомном эпитаксиальном слое от протекания тока коллектора. Дополнительный заряд ΔQ_K , соответствующий этому падению напряжения, как и заряд подвижных носителей Q_{f5} , пропорционален току I_K . Следова-

тельно, заряд ΔQ_K — Q_{f5} может быть добавлен к сумме $\sum_{i=1}^4 Q_{fi}$, при-

водя к росту общего заряда Q_f , а также к увеличению постоянной времени $\tau_{TN}^* \approx \tau_{TN} + \tau_K$, где $\tau_K \sim C_{jc} R_K$, R_K — сопротивление тела

коллектора. Полный заряд $\sum_{i=1}^5 Q_{fi} = Q_{DЭ}$ в активном нормальном режиме можно приписать диффузионной емкости эмиттера C_{de} (величины $U_{Э}$, I_K изменяются, напряжение U_K постоянно):

$$Q_{DЭ} = \sum_{i=1}^5 (\tau_{Э} + \tau_{pnЭ} + \tau_{Б} + \tau_{pnК} + \tau_{К}) I_K = \tau_{ЭК} I_K, \quad (6.56)$$

где $\tau_{ЭК}$ — время задержки эмиттер — коллектор.

Аналогично, полный заряд подвижных носителей $Q_{DK} = \sum_{i=6}^{10} Q_{fi}$, накапливаемый в различных областях БТ в активном инверсном режиме, можно приписать диффузионной емкости коллектора C_{dc} :

$$Q_{DK} = \sum_{i=6}^{10} Q_{fi} = Q_{f6} + Q_{f7} + Q_{f8} + Q_{f9} + Q_{f10} =$$

$$= (\tau_{IK} + \tau_{IpnК} + \tau_{ITБ} + \tau_{IpnЭ} + \tau_{fЭ}) I_{Э} = \tau_{КЭ} I_{Э}. \quad (6.57)$$

Здесь $Q_{f6} = \tau_{IK} I_{Э}$ — заряд, накапливаемый в коллекторе; τ_{IK} — время накопления в коллекторе; Q_{f7} — заряд, накапливаемый в коллекторном $p-n$ -переходе; $\tau_{IpnК}$ — время накопления в коллекторном $p-n$ -переходе; Q_{f8} — заряд, накапливаемый в базе; $\tau_{ITБ}$ — время переноса заряда в базе в активном инверсном режиме; Q_{f9} — заряд, накапливаемый в эмиттерном $p-n$ -переходе; $\tau_{IpnЭ}$ — время накопления в эмиттерном $p-n$ -переходе; Q_{f10} — заряд, накапливаемый в эмиттере; $\tau_{fЭ}$ — время накопления заряда в эмиттере. Аналогично времени задержки $\tau_{ЭК}$ в (6.56) величина $\tau_{КЭ}$ представляет собой время задержки коллектор — эмиттер в инверсном нормальном режиме в (6.57).

Соответствующие постоянные времени, входящие в выражения зарядов подвижных носителей Q_{fi} ($i=1 \div 10$), могут быть рассчитаны в электрофизических и геометрических параметрах БТ, что целесообразно при аналитическом расчете и оптимизации БТ, например, по критерию качества исполнения или быстродействия в той или иной схеме включения. Однако, как показывает опыт численного схемотехнического ФТ-моделирования БТ и схем на их основе, этот этап в целом не является основным при машинных расчетах и оптимизации БТ и схем по тем же критериям качества исполнения или быстродействия.

Выражение для «доминирующего» коллекторного тока (6.43) можно представить в виде

$$I_K = I_F - I_I, \quad (6.58)$$

где

$$I_F = \frac{I_s Q_{p0} \left(\exp \frac{U_{Э}}{\varphi_T} - 1 \right)}{Q_p}, \quad (6.59)$$

$$I_I = \frac{I_s Q_{p0} \left(\exp \frac{U_K}{\varphi_T} - 1 \right)}{Q_p}. \quad (6.60)$$

Ток насыщения ВАХ $I_K(U_K) - I_s$,

$$I_s Q_{p0} \approx q^2 S_{\Delta} n_i^2 D = C_1. \quad (6.61)$$

Подставляя (6.59), (6.60) в выражение для заряда (6.47), получим квадратное уравнение относительно Q_p , причем его решение с учетом соотношений для подвижных зарядов (6.56), (6.57) имеет вид

$$Q_p = \frac{Q_{p0} + Q_{\Delta} + Q_K}{2} + \left\{ \left(\frac{Q_{p0} + Q_{\Delta} + Q_K}{2} \right)^2 + I_s Q_{p0} \left[\tau_{rN} \left(\exp \frac{U_{\Delta}}{\varphi_T} - 1 \right) + \tau_{rI} \left(\exp \frac{U_K}{\varphi_T} - 1 \right) \right] \right\}^{1/2}. \quad (6.62)$$

В каноническом способе моделирования зарядов Q_{Δ} , Q_K по формулам (6.49), (6.50) необходимо проводить оценку интегралов на основании априорного описания барьерных емкостей через три-четыре электрических параметра. Однако существует более простой, но эффективный способ описания зарядов Q_{Δ} , Q_K и токов в формулах (6.58) — (6.60). Воспользуемся эквивалентной упрощенной формулой описания отношений зарядов Q_{Δ}/Q_{p0} , Q_K/Q_{p0} . Тогда можно записать

$$\frac{Q_{\Delta}}{Q_{p0}} = \frac{\int_0^{U_{\Delta}} C_{Je}(U) dU}{Q_{i0}} = \frac{\bar{C}_{Je} U_{\Delta}}{Q_{i0}} = \frac{U_{\Delta}}{V_{A\Delta}} = \frac{\mu_{\Delta} U_{\Delta}}{\varphi_T}, \quad (6.63)$$

$$\frac{Q_K}{Q_{p0}} = \frac{\int_0^{U_K} C_{Jc}(U) dU}{Q_{p0}} = \frac{\bar{C}_{Jc} U_K}{Q_{p0}} = \frac{U_K}{V_{AK}} = \frac{\mu_K U_K}{\varphi_T}, \quad (6.64)$$

где \bar{C}_{Je} , \bar{C}_{Jc} — эквивалентные барьерные емкости эмиттера и коллектора; μ_{Δ} , μ_K — коэффициенты обратной связи по напряжению; $V_{A\Delta} = \frac{\varphi_T}{\mu_{\Delta}}$, $V_{AK} = \frac{\varphi_T}{\mu_K}$ — напряжения Эрли.

С учетом выражений (6.63) (6.64) доминирующий ток коллектора I_K можно представить в виде

$$I_K = \frac{I_s}{g_B} \left(\exp \frac{U_{\Delta}}{\varphi_T} - \exp \frac{U_K}{\varphi_T} \right), \quad (6.65)$$

где

$$\frac{Q_B}{Q_{i0}} = g_B = 1 + \left(\frac{U_{\Delta}}{V_{A\Delta}} + \frac{U_K}{V_{AK}} \right) \frac{1}{2} +$$

$$+ \left\{ \left[\left(1 + \frac{U_{\text{Э}}}{V_{\text{АЭ}}} + \frac{U_{\text{К}}}{V_{\text{АК}}} \right) \frac{1}{2} \right]^2 + \frac{I_s \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right)}{I_N} + \frac{I_s \left(\exp \frac{U_{\text{К}}}{\varphi_T} - 1 \right)}{I_I} \right\}^{1/2}. \quad (6.66)$$

В (6.66) I_N , I_I — критические токи начала эффектов высокого уровня инжекции в активном нормальном и инверсном нормальном режимах:

$$I_N = Q_{p0}/\tau_{TN}, \quad I_I = Q_{p0}/\tau_{TI}. \quad (6.67)$$

Таким образом, имеется три типа уравнений, описывающих (по мере возрастания точности и сложности) доминирующий коллекторный ток через заряд дырок в БТ типа n - p - n : 1) упрощенные соотношения (6.65), (6.66) через набор стационарных параметров соответствующих токов и напряжений; 2) соотношения (6.58), (6.61) через барьерные и диффузионные емкости или времена переноса τ_{TN} , τ_{TI} ; 3) общие соотношения (6.43)—(6.45) через полный заряд в БТ (6.45), (6.47) (или его отдельные компоненты Q_{p0} , $Q_{\text{Э}}$, $Q_{\text{К}}$, Q_f). Типы моделей 2, 3 при идентичном описании зарядов $Q_{\text{Э}}$, $Q_{\text{К}}$, Q_f можно считать эквивалентными.

Наиболее привлекательной с точки зрения многоуровневого моделирования представляется концепция определения Q_{p0} из данных технологического расчета БТ, а Q_p — Q_{p0} (или отдельных компонентов $Q_{\text{Э}}$, $Q_{\text{К}}$, Q_f) — с помощью ФТ-моделирования в одно-, двух- и трехмерном моделировании. Использование рассчитанных зарядов (Q_p , Q_{p0} , $Q_{\text{Э}}$, $Q_{\text{К}}$, Q_f) с помощью стационарного ФТ-моделирования БТ в качестве базовых переменных для синтезируемых электрических моделей является основным звеном дальнейшей оценки параметров эквивалентной схемы БТ: емкостей (барьерных и диффузионных), ВАХ $I_{\text{К}}(U_{\text{Э}}, U_{\text{К}})$ и др.

С другой стороны, возможен синтез электрической модели БТ на основе первичных экспериментальных данных: I_s , зависимостей C_{je} , C_{jc} от напряжения и постоянных времени τ_{TN} , τ_{TI} от тока для эффективного описания зарядов $Q_{\text{Э}}$, $Q_{\text{К}}$ интегральными соотношениями (6.49), (6.50) и заряда Q_f интегральным соотношением (6.48) (рис. 6.3). Упрощение интегралов (6.48)—(6.50) путем усреднения, например, барьерных емкостей C_{je} , C_{jc} или описания возрастания τ_{TN} с ростом тока простыми аналитическими функциями $\tau_T(I)$ позволяет упростить синтез электрических моделей БТ в достаточно узком режимном диапазоне.

Наконец, имеется еще одна возможность синтеза электрических моделей БТ: расчеты с помощью технологических и ФТ-моделей основных электрических характеристик и параметров БТ и настройки ФТ- и аналитических зависимостей с целью выделения необходимых параметров электрической модели.

6.5. ОПИСАНИЕ ФИЗИЧЕСКИХ ЭФФЕКТОВ В БТ ЭЛЕКТРИЧЕСКОЙ МОДЕЛЬЮ С УПРАВЛЕНИЕМ ЗАРЯДОМ

Влияние зарядов $Q_{\text{Э}}$, $Q_{\text{К}}$ на ВАХ и коэффициент передачи тока. Рассмотрим влияние зарядов $Q_{\text{Э}}$ ($g_{\text{Э}} = Q_{\text{Э}}/Q_{\text{P0}}$) и $Q_{\text{К}}$ ($g_{\text{К}} = Q_{\text{К}}/Q_{\text{P0}}$) на ход зависимости $I_{\text{К}}$ ($U_{\text{Э}}$, $U_{\text{К}}$).

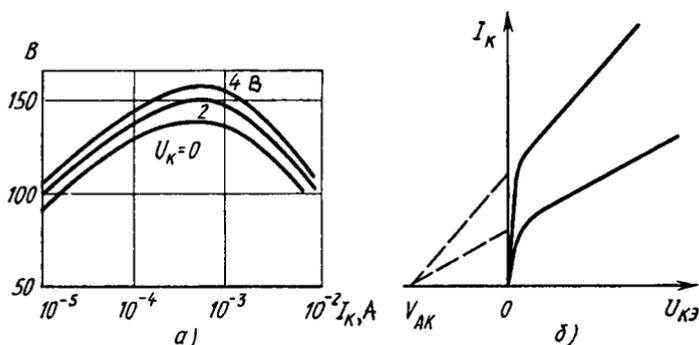


Рис. 6.4. Зависимости коэффициента передачи тока БТ от тока коллектора для различных напряжений на коллекторном переходе (а); определение напряжения Эрли из выходных ВАХ БТ (б)

При пренебрежимо малом влиянии других составляющих заряда $g_i \rightarrow 0$ в активном нормальном режиме для низкого уровня инжекции, согласно (6.64), (6.65), получим

$$I_{\text{К}} = \frac{I_{\text{с}}}{1 + g_{\text{Э}}} \left(\exp \frac{U_{\text{Э}}}{\varphi_{\text{T}}} - 1 \right), \quad (6.68)$$

$$\frac{\partial I_{\text{К}}}{\partial U_{\text{Э}}} = \frac{q I_{\text{К}}}{kT} \left(1 - \varphi_{\text{T}} \frac{C_{\text{je}}}{1 + g_{\text{Э}}} \frac{U_{\text{Э}}}{Q_{\text{P0}}} \right). \quad (6.69)$$

Наклон зависимости $I_{\text{К}}(U_{\text{Э}})$ в области малых токов характеризуется коэффициентом $n_{\text{Э}}$ φ_{T} , $n_{\text{Э}} \geq 1$. Согласно определению,

$$\frac{1}{n_{\text{Э}}} = \frac{\varphi_{\text{T}}}{I_{\text{К}}} \frac{\partial I_{\text{К}}}{\partial U_{\text{Э}}} = 1 - \frac{\varphi_{\text{T}} C_{\text{je}} U_{\text{Э}}}{(1 + g_{\text{Э}}) Q_{\text{P0}}} \quad (6.70)$$

и с учетом соотношений (6.69), (6.63) коэффициент $n_{\text{Э}}$, характеризующий отклонение зависимости $I_{\text{К}}(U_{\text{Э}})$ от идеальной с наклоном $n_{\text{Э}} = 1$, равен

$$n_{\text{Э}} \approx \frac{1}{1 - \varphi_{\text{T}} / (U_{\text{Э}} + U_{\text{АЭ}})}. \quad (6.71)$$

Аналогично при пренебрежимо малом влиянии других составляющих заряда влияние заряда в коллекторе $Q_{\text{К}}(g_{\text{К}})$ проявляется в ходе зависимости $I_{\text{К}}(U_{\text{Э}})$ (рис. 6.4, а)

$$I_K = \frac{I_S}{1 + g_K} \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right) = \frac{I_S}{1 + \frac{U_K}{U_{\text{АК}}}} \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right) \quad (6.72)$$

смещением коллекторного тока напряжением на коллекторном переходе (в активном нормальном режиме с ростом $|U_K|$ ток коллектора увеличивается). Этот же эффект проявляется в увеличении дифференциальной проводимости выходной характеристики $I_K(U_{\text{КЭ}})$ для БТ в схеме с общим эмиттером

$$g_0 = \left. \frac{\partial I_K}{\partial U_{\text{КЭ}}} \right|_{U_{\text{Э}} = \text{const}} = I_K(U_{\text{Э}}^*) V_{\text{АК}}, \quad (6.73)$$

приводящей к увеличению I_K с ростом $U_{\text{КЭ}}$ при фиксированном напряжении эмиттера (рис. 6.4, б).

Эффект модуляции ширины базы (эффект Эрли). Этот эффект — зависимость $B(I_K, U_{\text{КЭ}})$ — при произвольном уровне инжекции проявляется в следующем. При изменении напряжения обратнo-смещенного коллекторного перехода изменяется ширина p -перехода (в основном в сторону области базы), при этом изменение ширины базы влияет на коэффициенты переноса инжекции БТ. В данном случае $|U_{\text{КБ}}| \uparrow$, $W_{p\text{н}} \uparrow$, $W_{\text{Б}} \downarrow$, $\downarrow \tau_{\text{ТБ}} = W_{\text{Б}}^2 / \eta D_n$, $(\tau_{\text{ТБ}} / \tau) \downarrow$, $\uparrow \alpha_{\text{Т}} = 1 - \tau_{\text{Т}} / \tau$, $\uparrow B \approx \alpha_{\text{Т}} \gamma$.

Однако большинство современных БТ с тонкой базой и минимальными рекомбинационными потерями относится к типу БТ, определяемых эффективностью эмиттера. В этом случае $I_K \approx I_n$ и интеграл от концентрации характеризует степень влияния модуляции толщины базы на зависимость $I_K(U_K)$.

Дифференцируя по U_K выражение для коллекторного тока (6.43), получим

$$\frac{1}{I_K r_K} = - \frac{1}{Q_p} \left(\frac{dQ_p}{dU_K} \right), \quad r_K = \left. \frac{dU_K}{dI_K} \right|_{U_{\text{Э}} = \text{const}}. \quad (6.74)$$

Тогда, согласно соотношению (6.74),

$$1 / (I_K r_K) = C_{j\text{с}} / Q_p, \quad I_K r_K = Q_p / C_{j\text{с}} = V_{\text{АК}}. \quad (6.75)$$

Предполагая $V_{\text{АК}} \approx \text{const}$, выражение (6.75) можно переписать в виде

$$- \left(\frac{1}{Q_p} \right) \frac{dQ_p}{dU_{\text{КБ}}} = \frac{1}{V_{\text{АК}}} = \text{const} \quad (6.76)$$

с решением

$$Q_p = Q_{p0} (U_K = 0) \exp(-U_K / U_{\text{АК}}). \quad (6.77)$$

Подставляя (6.77) в формулу для коллекторного тока (6.43), получим общее выражение для тока коллектора, учитывающее эффект Эрли:

$$I_K = I_{K0} \exp \left\{ \frac{U_{\text{Э}}}{\varphi_T} + \frac{U_K}{V_{\text{АК}}} \right\} = I_{K0} \exp \left\{ \frac{U_{\text{Э}} + \frac{U_{\text{КБ}}}{\mu_K}}{\varphi_T} \right\},$$

$$I_{K0} = I_s, \quad \mu_K = \frac{V_{\text{АК}}}{\varphi_T} = - \left. \frac{\partial U_K}{\partial U_{\text{Э}}} \right|_{I_K}. \quad (6.78)$$

Выражение (6.78) может быть использовано для объяснения зависимости $B_0(U_K)$, роста коэффициента передачи тока B с увеличением напряжения $|U_K|$ (рис. 6.4, а):

$$B_0(U_K) = B_0(U_K = 0) \exp \frac{U_K}{V_{\text{АК}}} = B_0(U_K = 0) \exp \frac{U_K}{\mu_K \varphi_T}. \quad (6.79)$$

Для небольших значений $U_K \ll V_{\text{АК}}$ соотношение (6.79) упрощается:

$$B_0(U_K) = B_0(U_K = 0) (1 - U_K/V_{\text{АК}}). \quad (6.80)$$

Выражения (6.79), (6.80) позволяют по параметру $V_{\text{АК}}$ БТ прогнозировать изменение коэффициента передачи B в широком диапазоне изменения напряжения U_K (рис. 6.4, а, б).

Эффект накопления зарядов в базе. Рассмотрим влияние накопления подвижных носителей заряда в базе БТ Q_f ($g_f = Q_f/Q_{p0}$), при допущении, что нет их накопления в области коллектора ($Q_f \approx Q_{fb}$) и можно пренебречь влиянием зарядов $Q_{\text{Э}}, Q_{\text{К}}$. Тогда из соотношения (6.65) следует

$$g_b \approx \sqrt{g_{fb}}; \quad I_K = \frac{I_s \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right)}{\left(I_s \exp \frac{U_{\text{Э}}}{\varphi_T} I_N \right)^{1/2}} = \sqrt{I_s I_N} \exp \frac{U_{\text{Э}}}{2\varphi_T}, \quad (6.81)$$

откуда можно определить характеристический ток I_N — «ток колена» по пересечению двух асимптот зависимости $I_K(U_{\text{Э}})$ с наклоном φ_T и $2\varphi_T$ (рис. 6.5). Отметим, что за интегральное изменение зависимости $I_K(U_{\text{Э}})$ ответственны подвижные носители заряда, накопленные как в базе Q_{fb} , так и в коллекторе БТ ($Q_f - Q_{fb}$) (см. рис. 6.3). Не касаясь области малых токов, из рассмотрения типовых зависимостей коэффициента передачи тока и постоянных времени накопления можно сделать обобщающий вывод, что заряд Q_{fb} заметно влияет на изменение коэффициента B при токах коллектора, больших первого критического тока I_N ($I_K \geq I_N$), и слабо влияет на поведение кривой $\tau_{N}(I_K)$. В то же время заряд Q_{fK} , накопленный в коллекторе в условиях расширения активной базы (эффект Кирка) $W_{B\text{Эф}} \geq W_B$ для высоких уровней инжекции (см. рис. 6.1), сильно влияет как на зависимость $B(I_K)$, так и на зависимость

$\tau_{ЭК}$ (I_K) (рис. 6.6). Отметим, что в области средних токов, согласно полной токовой зависимости, коэффициент передачи тока

$$B = B_{\max} \frac{Q_{p0}}{Q_{p0} + Q_{\Sigma} + Q_K + Q_f} \quad (6.82)$$

близок к значениям B_{\max} , определяемым соотношениями (6.39), (6.41) для различных структурно-технологических решений БТ, и слабо зависит от зарядов $Q_{\Sigma} + Q_K$ (при фиксированном напряжении U_{Σ}^* , увеличении обратного смещения $|U_K|$ расширяется коллекторный p - n -переход, уменьшается заряд Q_f и увеличивается коэффициент B). При $I_K \geq I_N$ спад B становится более заметным, а при $I_K > I_{кр}$, когда эффективная база начинает расширяться в

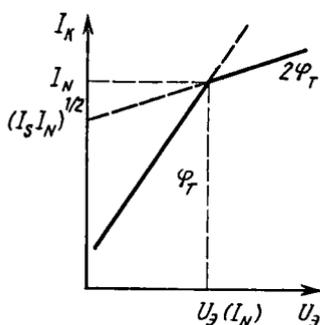


Рис. 6.5. Определение критического тока высокого уровня инжекции в базе БТ

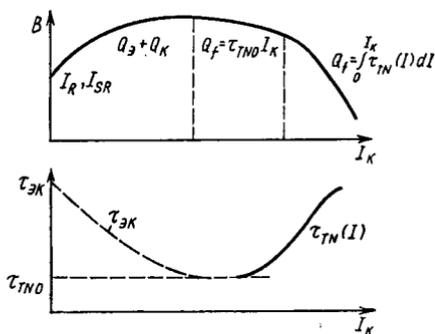


Рис. 6.6. Характерные области изменений зависимостей коэффициента передачи тока и времени задержки эмиттер — коллектор БТ от тока коллектора

относительно низкоомную коллекторную область, степень уменьшения коэффициента B с ростом тока значительно усиливается. Так, при $I_N \leq I_K \leq I_{кр}$

$$B = B_{\max}^* \frac{Q_{p0}}{Q_{p0} + Q_p} = \frac{B_{\max}}{1 + Q_p/Q_{p0}} \approx \frac{B_{\max}}{1 + I_K/I_N} \quad (6.83)$$

Параметр $\tau_{ТБ} = \frac{dQ_B}{dI_B}$ при аппроксимативном экспоненциальном распределении акцепторной примеси в базе $C_A(x) \approx C_A(x_1) \exp \times \left(-\eta \frac{x_1 - x_2}{W}\right)$, $\eta = \frac{C_A(x_1)}{C_A(x_2)}$, равный $\tau_{ТБ0} \approx \frac{W_B^2}{2D_p} \frac{2(\eta - 1 - e^{-\eta})}{\eta^2}$, при высоком уровне инжекции в базе становится независимым от η и I_K :

$$\tau_{ТБ}(I_K > I_N) \approx W_B / (4D_p) \quad (6.84)$$

Рост и ограничение $\tau_{TN} \approx \tau_{TB}$ при $I_N \leq I_K \leq I_{кр}$ иллюстрируются рис. 6.3. Однако для тонкослойных сильнолегированных БТ сверхскоростных БИС рассмотренный эффект модуляции проводимости базы, приводящий к отклонению зависимости $I_K(U_Э)$ (наклон $2\varphi_T$), спаду $V \sim I_K^{-1}$ и некоторому увеличению τ_{TN} , не столь заметен при сильном легировании базы по сравнению с эффектом расширения базы и накопления подвижных носителей заряда в коллекторе высококачественных БТ типа $n^+-p-n-n^+$. На рис. 6.7, а, б представлены расчетные ФТ-зависимости концентрации дырок и электронов от координаты в тонкослойном БТ, иллюстрирующие эффекты модуляции проводимости и расширения базы, отражающие факт доминирования последнего эффекта для достаточно толстых низколегированных слоев коллектора. Необходимо отметить независимость друг от друга критических токов $I_N, I_{кр}$.

Эффект расширения базы. Влияние накопления зарядов в коллекторе. Одними из основных эффектов в области высоких плотностей токов, приводящих к уменьшению коэффициента передачи V и увеличению постоянной времени τ_{TN} , являются эффекты расширения базы и квазинасыщения (или полного насыщения). Эффект расширения квазинейтральной базы в области коллектора может проявлять себя при условии, когда падение напряжения от протекания коллекторного тока на объемном сопротивлении коллектора вызывает резкое уменьшение обратного смещения на переходе коллектор—база (и смещение его в прямом направлении) в области активного БТ (под дном эмиттера).

Критический ток этого эффекта $I_{кр1}$ для небольших напряжений коллектора U_K оценивают из условия

$$I_{кр1} R_K = U_K + \varphi_{0K}; \quad (6.85)$$

$$J_{кр1} = \frac{I_K}{S_Э} = q C_K \mu_n \frac{U_K + \varphi_{0K}}{W_K}, \quad (6.86)$$

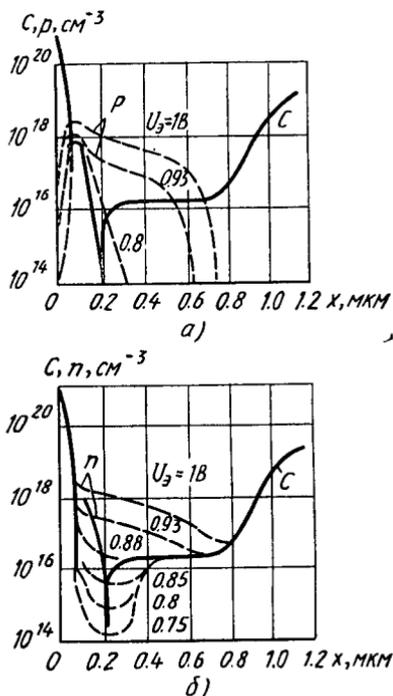


Рис. 6.7. Зависимости концентрации дырок (а) и электронов (б) в тонкослойном БТ от координаты для различных напряжений на эмиттерном переходе (сплошной линией показан профиль примеси в БТ)

где C_K, μ_n — соответственно концентрация донорной примеси и подвижность электронов в коллекторе.

При коллекторных токах больше критического ($I_K > I_{Kp1} S_{\Sigma}$) происходят расширение базы за счет эпитаксиального слоя с достаточно низкой концентрацией доноров C_K , а также значительное накопление подвижных носителей заряда в области коллектора, примыкающей к p - n -переходу.

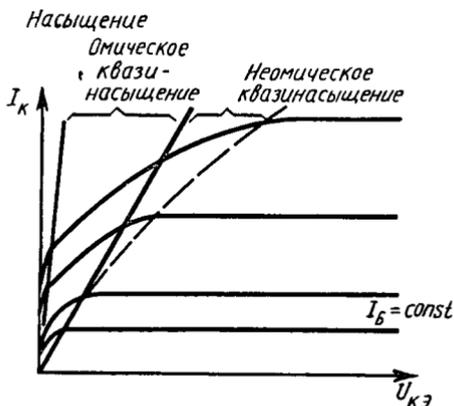


Рис. 6.8. Типовая выходная ВАХ БТ с общим эмиттером с выделением характерных областей насыщения и квазинасыщения

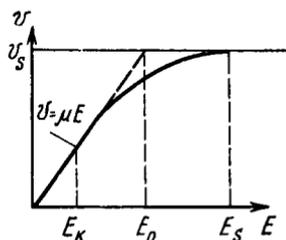


Рис. 6.9. Зависимость скорости подвижных носителей заряда напряженности от электрического поля

Область коллектора в данной ситуации проявляет себя как продолжение области базы с сильной концентрацией подвижных носителей заряда, наблюдается так называемое эффективное приращение базовой области. При низких плотностях тока область достаточно высокого электрического поля существует на границе активной базы и эпитаксиальной коллекторной области, причем концентрация электронов в эпитаксиальном коллекторе приблизительно равна концентрации доноров C_K . При $I_K > I_{Kp1}$ область высокого поля смещается в сторону границы эпитаксиального n -слоя и скрытого n^+ -слоя. В условиях квазинасыщения и высокой плотности тока концентрация электронов определяется постоянным градиентом. Таким образом, перемещение области высокого поля в глубь коллектора, а также постоянный градиент концентрации электронов позволяют рассматривать эффективную базу как состоящую из квазинейтральной базы при низких плотностях тока шириной W_B и дополнительного слоя $\Delta W_B = W_K (1 - I_{Kp1} / I_K)$, $I_K > I_{Kp1}$.

При больших значениях приложенного обратного напряжения эффект расширения базы может наступать без сопутствующего ему явления омического квазинасыщения, что отражено на выход-

ной ВАХ (рис. 6.8), при нарушении условий (6.85), (6.86). Это происходит потому, что в неомической области квазинасыщения при больших напряжениях U_K в эпитаксиальных БТ со слоем эпитаксии шириной W_K возникает большое электрическое поле $E \approx \approx |U_K|/W_K$, вследствие чего носители заряда в данной области становятся «теплыми» или даже «горячими». В предельном случае больших режимных напряжений U_K и (или) малых толщин W_K наблюдается дрейф носителей со скоростью насыщения v_s . Качественный характер изменения скорости носителей в коллекторной области отражен на рис. 6.9. Даже в условиях отсутствия эффекта квазинасыщения при некоторой плотности тока носителей, перемещающихся в коллекторе со скоростью насыщения v_s , в p - n -переходе плотность заряда подвижных носителей становится сравнимой с плотностью зарядов акцепторной и донорной примесей. При этом начиная с некоторой критической плотности тока $J_{кр2}$, $J_K > > J_{кр2}$

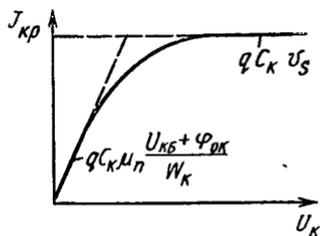


Рис. 6.10. Зависимость критической плотности тока расширения базы в коллектор от напряжения коллектор — база

$$J_{кр2} \approx qv_s C_K \quad (6.87)$$

происходит расширение базы в коллектор.

Если за критическую напряженность электрического поля принять величину $E_0 \approx v_s/\mu_n$, $E < E_0 < E_s$ (рис. 6.9), то можно уточнить условия начала эффекта расширения базы в коллектор для различных напряжений поля в коллекторе и привести им в соответствие критические плотности тока $J_{кр1}$, $J_{кр2}$. Тогда при $U_K + \varphi_{0K} \leq E_0 W_K$

$$J_{кр1} = qC_K \mu_n \frac{U_{KB} + \varphi_{0K}}{W_K}, \quad (6.88)$$

при $U_K + \varphi_{0K} > E_0 W_K$

$$J_{кр2} = qv_s C_K,$$

что позволяет построить качественную зависимость $J_{кр}(U_K)$ (рис. 6.10), причем условие $J_{кр} = qv_s C_K$ выполняется при больших полях в коллекторе: $E > (0.9 \div 2) \cdot 10^4$ В/см ($v_s \approx 10^7$ см/с).

Из рис. 6.7 видно, что для больших смещений U_3 концентрации дырок и электронов превосходят концентрацию примесей в эпитаксиальном слое, причем в эффективной (расширенной) базе происходит значительное избыточное накопление электронов (и соответственно дырок согласно закону электронейтральности полу-

проводника). Выражение для избыточного заряда электронов на единицу площади можно представить в виде

$$Q_{fK} \approx \tau_{TB} J_K, \quad (6.89)$$

где τ_{TB} — время переноса в эффективной базе в условиях высокого уровня инжекции (6.84).

С учетом Q_{fK} коллекторный ток

$$I_K = \frac{I_s Q_{p0}}{Q_{p0} + Q_{fK}} \exp \frac{U_{\Delta}}{\varphi_T}. \quad (6.90)$$

Выражение (6.90) можно записать в виде

$$I_K \left(1 + \frac{\tau_{TB} I_K}{Q_{p0}} \right) = I_s \exp \frac{U_{\Delta}}{\varphi_T}. \quad (6.91)$$

Тогда зависимость коэффициента передачи тока можно будет определить следующим образом:

$$B \approx \frac{B_{\max}}{1 + Q_K/Q_{p0}} \approx \frac{B_{\max}}{1 + I_K/I_{Kp}} \sim I_K^{-1}. \quad (6.92)$$

При «чистом» расширении базы для больших токов (плотностей токов), согласно (6.92), спад коэффициента B происходит пропорционально I_K^{-1} . Однако в условиях одновременного проявления эффектов модуляции проводимости базы и расширения базы спад этого коэффициента более интенсивен (см. рис. 6.6) и из соотношений (6.83), (6.92) при $I_K > I_{Kp}$ следует

$$\begin{aligned} B(I_K \geq I_{Kp}) &= \frac{B(I_K > I_N)}{1 + I_K/I_{Kp}} = \left[\frac{B}{1 + I_K/I_N} \right] \left/ \left(1 + \frac{I_K}{I_{Kp}} \right) \right. \approx \\ &\approx \frac{B_{\max} I_N I_{Kp}}{I_K^2}, \end{aligned} \quad (6.93)$$

т. е. уменьшение коэффициента B с ростом тока приобретает новую зависимость $B \sim I_K^{-2}$.

Сделаем оценку величины $\tau_{TN} \approx \tau_{TB}$ при расширении базы. При $I_K > I_{Kp}$ ширина эффективной базы

$$W_{B\text{эф}} = W_B + W_K (1 - I_{Kp}/I_K). \quad (6.94)$$

Поскольку в режиме высокого уровня инжекции выполняется соотношение

$$\tau_{TB}(W_B, I > I_N) = \frac{W_B^2}{4D_n}, \quad \tau_{TB}^*(W_{B\text{эф}}, I_K > I_{Kp}) = \frac{W_{B\text{эф}}}{4D_n}, \quad (6.95)$$

из (6.95) следует

$$\tau_{TB}^* = \tau_{TB} (W_{B\text{эф}}/W_B)^2 = \tau_{TB} \theta_1, \quad (6.96)$$

где

$$\theta_1 = \left[1 + \frac{W_K}{W_B} \left(1 - \frac{I_{Kp}}{I_K} \right) \right]^2. \quad (6.97)$$

На рис. 6.11 приведены расчетные зависимости $\tau_{тБ}(I_K)/\tau_{тБ}(I_{кр})$ от нормированного тока коллектора $I_K/I_{кр}$ при различных значениях параметра W_K/W_B . Чем меньше значение параметра W_K/W_B , тем быстрее и при меньших коллекторных токах происходит насыщение зависимости $\tau_{тБ}(I_K)/\tau_{тБ}(I_{кр})$. Ограничение роста $\tau_{тБ}(I_K)$ с увеличением тока объясняется расширением эффективной базовой области до максимальной границы $W_B^*_{эф} = W + W_K$, остающейся неизменной при дальнейшем увеличении тока I_K , что характерно в первую очередь для интегральных БТ с малой толщиной эпитаксиального слоя коллектора.

Для устранения численной неустойчивости в области малых токов I_K в выражение для θ_1 в канонической модели Гуммеля—Пуна вводят дополнительный токковый параметр I_2 , при котором фактор θ_1 , характеризующий рост времени переноса базы при расширении базы, записывается в виде

$$\theta_2 = \left\{ 1 + \frac{W_K}{4W_B} \times \frac{\left\{ [(I_K + I_{кр})^2 + I_2^2]^{1/2} - (I_K + I_{кр}) \right\}^2}{I_K^2 + I_2^2} \right\}^2 \quad (6.98)$$

Для того чтобы отразить не только эффект расширения базы, но и эффекты вытеснения эмиттерного тока, накопления носителей в пассивной базе, а также взаимовлияние эффектов, во многих случаях целесообразно «настраивать» реальные экспериментальные зависимости $\tau_{тБ} \approx \tau_{тВ}$ с теоретической зависимостью (6.98) с помощью фактора θ_3 , в котором используется параметр настройки n_p :

$$\theta_3 = \left\{ 1 + \frac{W_K}{4W_B} \cdot \frac{\left\{ [(I_K + I_{кр})^2 + I_2^2]^{1/2} - (I_K + I_{кр}) \right\}^2}{I_K^2 + I_2^2} \right\}^{n_p} \quad (6.99)$$

Двумерный эффект растекания коллекторного тока. Если при малых и средних уровнях инжекции поток электронов из эмиттера в коллектор, составляющих суть «доминирующего» тока коллектора, был практически одномерным, то при высоких уровнях инжекции, проявляющихся в базе и коллекторе БТ, происходит

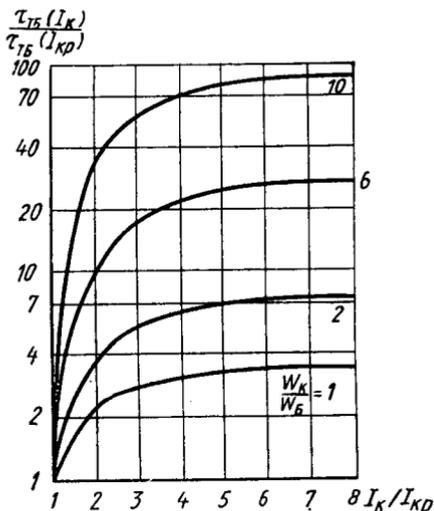
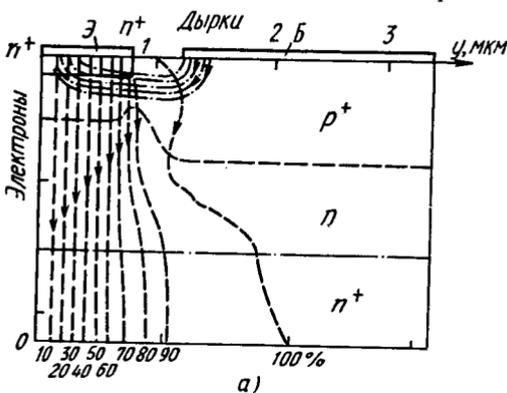


Рис. 6.11. Зависимости нормированной составляющей времени задержки эмиттер — коллектор БТ от нормированного тока коллектора при одномерном рассмотрении эффекта расширения базы

существенное нарушение одномерности. Этот факт иллюстрируется ФТ-расчетами тока электронов в двумерном приближении (рис. 6.12) при высокой плотности коллекторного тока $I \approx 4 \cdot 10^4$ А/см². Для минимизации эффекта паразитной боковой инжекции была



использована диэлектрическая стенка на боковой части эмиттерного перехода и отношение ширины эмиттера к его длине составляло всего 0.04. Это позволило, во-первых, выявить накопление подвижных носителей заряда в области пассивной базы за пределами активной области БТ (под дном эмиттера), во-вторых, составить представление о тангенциальной составляющей

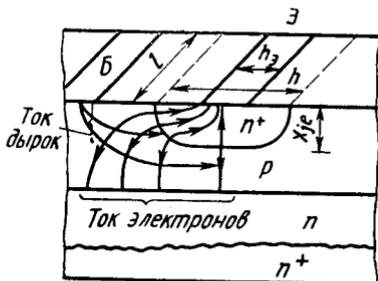
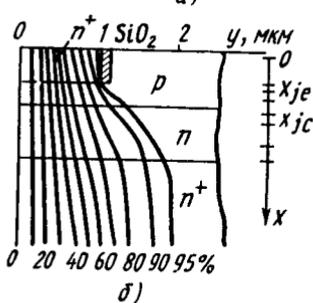


Рис. 6.12. Двумерное распределение потоков электронов и дырок в БТ:
а — в типовом тонкослойном БТ; б — в БТ с изолированной боковой областью эмиттерного перехода

Рис. 6.13. Выделение токов дырок и электронов БТ при реальной конфигурации эмиттерного перехода

движения электронов в пассивной базе с эффективной шириной $W_{Бэф} > W_{Б}$. Поэтому, с одной стороны, в электрической модели БТ необходимо отразить двумерное растекание тока коллектора и учесть диффузионную составляющую тока базы I_L , а также поток дырок, рекомбинирующих с электронами в пассивной базе (рис. 6.13). С другой стороны, в электрической модели БТ необходимо отразить эффект увеличения времени переноса базы и коллектора при высоких токах.

С точки зрения трансформации электрической модели периферийный (двумерный) ток электронов можно отнести к внутреннему БТ, оставляя плотность тока коллектора через БТ неизменной, т. е. при описании

$$I_K \approx \frac{C_1}{Q_{p0} + Q_f(2D)}, \quad Q_f(2D) > Q_f(1D) \quad (6.100)$$

следует несколько увеличить топологическую площадь эмиттера (ширину эмиттерной полоски).

Согласно концепции, учитывающей для высокой плотности тока в коллекторе двумерное растекание неосновных носителей заряда в базе БТ с тангенциальной составляющей скорости переноса, увеличение площади коллекторного перехода, собирающего носители относительно площади инжектирующего перехода, эффективную ширину базы можно определить из выражений

$$S_K = \gamma_K S_{\Delta} I_K / I_{Kp1}, \quad S_{\Delta} = hl,$$

$$\gamma_K = h / (2W_B), \quad (6.101)$$

$$W_{B\text{эф}} = [W_B^2 + 1/4 (I_K / I_{Kp1} - 1)^2 (\gamma_K h)^2]^{1/2} \quad (6.102)$$

при токах и напряжениях коллектора больше критических: $I_K \geq I_{Kp1}$, $|U_K| > |U_{Kp1}|$. Фактор степени увеличения площади коллектора γ_K может служить в качестве подстроечного параметра. Тогда для модели эффективного увеличения площади коллекторного перехода при двумерном растекании неосновных носителей заряда в активной базе БТ время переноса базы можно представить соотношением

$$\tau_{TB}(I_K) = \tau_{TB}(I_{Kp1}) [1 + \gamma_K^2 (I_K / I_{Kp1} - 1)^2 / 4] = \tau_{TB}(I_{Kp1}) \Gamma. \quad (6.103)$$

На рис. 6.14 приведены зависимости $\tau_{TB}(I_K) / \tau_{TB}(I_{Kp1})$ от нормированного тока коллектора при различных значениях $h / (2W_B)$. Зависимость $\tau_{TB}(I_K)$ не обнаруживает ограничения для больших токов в отличие от рис. 6.11, интерпретирующего эффект расширения базы в одномерном приближении.

Эффект накопления неосновных носителей заряда в эмиттере. В ряде случаев при пренебрежимо малом влиянии эффекта расширения базы в эмиттер эффект модуляции проводимости базы может проявляться в накоплении носителей заряда в эмиттерной области БТ. При этом наряду с потерей быстродействия в базе, увеличением τ_{TB} при $I_K > I_N$ в (6.84) возможна значительная потеря быстродействия БТ за счет возрастания постоянной времени

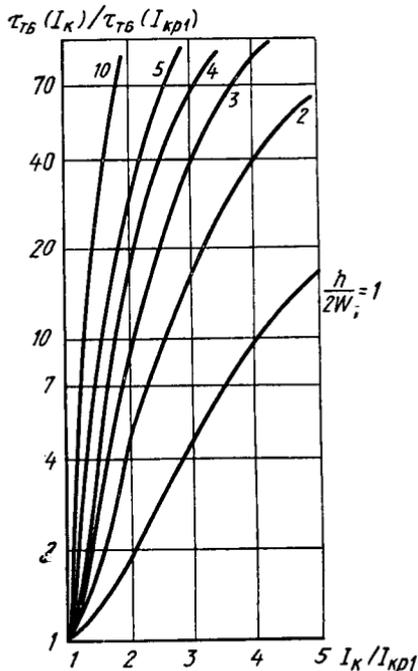
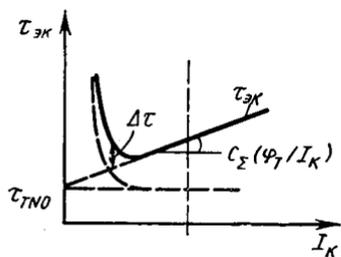


Рис. 6.14. Зависимости нормированной составляющей времени задержки эмиттер — коллектор БТ от нормированного тока коллектора при двумерном рассмотрении эффекта высокого уровня инжекции в коллекторе

накопления в эмиттере τ_{Σ} . Для некоторых конструкций БТ в определенных условиях (при больших плотностях тока) τ_{Σ} может даже стать доминирующей постоянной времени $\tau_{\Sigma} \approx \tau_{TN}$ и определять практически реальное быстроедействие БТ.

Будем исходить из того факта, что заряды, накопленные в эмиттере при низких Q_{Σ} (Н) и высоких Q_{Σ} (В) уровнях инжекции, определяются соотношениями согласно выражению (6.81):



$$Q_{\Sigma}(H) = \text{const} \frac{I_K}{I_S}, \quad I_K < I_N, \quad (6.104)$$

$$Q_{\Sigma}(B) = \text{const} \frac{I_K}{(I_S I_N)^{1/2}} = \frac{I_K^2}{I_S I_N}, \quad I_K \geq I_N. \quad (6.105)$$

Рис. 6.15. Выделение составляющих из зависимости времени задержки эмиттер — коллектор БТ от коллекторного тока

Для зарядов, накопленных в эмиттере в режимах низких и высоких уровней инжекции, выполняются также соотношения

$$Q_{\Sigma}(H) = \tau_{\Sigma}(H) I_K(H), \quad Q_{\Sigma}(B) = \tau_{\Sigma}(B) I_K(B). \quad (6.106)$$

При этом суммарный заряд, накопленный в эмиттере, с учетом выражений (6.104)—(6.106) можно записать в виде

$$Q_{\Sigma} = Q_{\Sigma}(H) + Q_{\Sigma}(B) = Q_{\Sigma}(H)(1 + I_K/I_N) = \tau_{\Sigma} I_K (1 + I_K/I_N). \quad (6.107)$$

Дифференцируя суммарный заряд (6.107) по времени, находим эффективную постоянную времени накопления носителей заряда в эмиттере

$$\tau_{\Sigma \text{эф}} = \frac{dQ_{\Sigma}}{dI_K} = \tau_{\Sigma} \left(1 + \frac{2I_K}{I_N} \right), \quad (6.108)$$

которую удобно представить в виде

$$\tau_{\Sigma \text{эф}} = \begin{cases} \tau_{\Sigma}, & I_K < I_N, \\ \tau_{\Sigma} (1 + 2I_K/I_N), & I_K \geq I_N, \end{cases} \quad (6.109)$$

где τ_{Σ} в зависимости от конструкции БТ определяется соотношениями (6.35), (6.38), (6.40).

Для малых критических токов I_N (начало эффекта модуляции проводимости базы) и больших режимных токов, согласно линейной зависимости (6.109), при $I_K \geq I_N$ может оказаться доминирующей постоянной времени $\tau_{TN} \sim \tau_{\Sigma}$ при минимальном влиянии других составляющих τ_i для ряда профилей, конструкций и режимов использования БТ.

Тогда при пренебрежимо малых других постоянных времени постоянную времени накопления можно представить в виде

$$\tau_{\tau N} = \begin{cases} \tau_{\tau N0} = \tau_{\text{Э}} + \tau_{\tau \text{Б}0} + \tau_{\text{К}}, & I_{\text{К}} < I_{\text{Н}}, U_{\text{К}} < U_{\text{кр}}, \\ \tau_{\tau \text{Б}} + \tau_{\text{Э}} \left(1 + \frac{2I_{\text{К}}}{I_{\text{Н}}}\right), & I_{\text{К}} \geq I_{\text{Н}}, U_{\text{К}} < U_{\text{кр}}, \\ \tau_{\text{Э}} \left(1 + \frac{2I_{\text{К}}}{I_{\text{Н}}}\right) + \tau_{\tau \text{Б}}(1 + \theta_{\text{Э}}), & I_{\text{К}} > I_{\text{кр}}, U_{\text{К}} < U_{\text{кр}}, \\ \tau_{\text{Э}} \left(1 + \frac{2I_{\text{К}}}{I_{\text{Н}}}\right) + \tau_{\tau \text{Б}}\Gamma, & I_{\text{К}} > I_{\text{кр}}, U_{\text{К}} \geq U_{\text{кр}}. \end{cases} \quad (6.110)$$

Зависимость $\tau_{\tau N}(I_{\text{К}})$ и другие ее составляющие удобно выделять из машинных и экспериментальных зависимостей полного времени задержки эмиттер—коллектор $\tau_{\text{ЭК}}(I_{\text{К}})$, $\tau_{\text{ЭК}} = 1/(2\pi f_{\tau})$, когда по оси токов откладывается величина $1/I_{\text{К}}$ (рис. 6.15). В большинстве случаев наклон $\tau_{\text{ЭК}}$ от $1/I_{\text{К}}$ в области малых токов определяется составляющей $\frac{\varphi_{\tau}}{I_{\text{К}}} C_{\Sigma}$:

$$C_{\Sigma} = \frac{1}{\varphi_{\tau}} \frac{d\tau_{\text{ЭК}}}{d(1/I_{\text{К}})}, \quad C_{\Sigma} = C_{j\text{е}} + C_{j\text{к}} + C_{\text{х}}, \quad \frac{\varphi_{\tau}}{I_{\text{Э}}} C_{\Sigma} = \tau(C_{\Sigma}), \quad (6.111)$$

$C_{j\text{е}}, C_{j\text{к}}$ — барьерные емкости эмиттерного и коллекторного переходов; $C_{\text{х}} = C_{\text{pe}} + C_{\text{pc}}$, $C_{\text{х}}$ — паразитная емкость, равная сумме конструктивных емкостей контактов, причем асимптота для больших токов ($1/I_{\text{К}} \rightarrow 0$) позволяет определить $\tau_{\tau N0}$ (6.110). Рост $\tau_{\text{ЭК}} = 1/(2\pi f_{\tau})$ в области больших токов

$$\tau_{\text{ЭК}} = \tau_{\tau N0} + \frac{\varphi_{\tau}}{I_{\text{К}}} C_{\Sigma} + \Delta\tau_{\tau N} \quad (6.112)$$

описывается увеличением $\Delta\tau_{\tau N}$ (6.110) для высоких уровней инжекции в различных диапазонах токов и напряжений коллектора. Зависимость, аналогичная (6.112) (рис. 6.15), может описывать инерционный процесс распространения сигнала от коллектора к эмиттеру с временем задержки $\tau_{\text{ЭК}}$.

6.6. КОМПОНЕНТЫ ТОКА БАЗЫ.

УМЕНЬШЕНИЕ КОЭФФИЦИЕНТА ПЕРЕДАЧИ ТОКА В МИКРОРЕЖИМЕ

Ток базы в электрической модели БТ можно представить в виде нестационарного и стационарного членов:

$$I_{\text{БЭ}} = \frac{dQ_{\rho}}{dt} + I_{\text{Брек}}. \quad (6.113)$$

Полный рекомбинационный ток базы можно разделить на два слагаемых, управляемых напряжением на эмиттерном и коллекторном переходах:

$$I_{\text{Брек}} = I_{\text{Б}}(\text{Э}) + I_{\text{Б}}(\text{К}). \quad (6.114)$$

Компоненты базовых рекомбинационных токов $I_B(\mathcal{E})$, $I_B(K)$ описываются экспоненциальными функциями с предэкспоненциальными коэффициентами, равными единице для идеальных (диффузионных) компонент, и коэффициентами неидеальности для рекомбинационно-генерационных токов в объеме и на поверхности эмиттерного и коллекторного p - n -переходов. Полные токи эмиттера и коллектора БТ в нестационарном и стационарном ($\frac{dI_I}{dt} = 0$, $\frac{dU_I}{dt} = 0$) режимах определяются выражениями, в которых доминирующий ток коллектора I_{KK} задан соотношениями (6.58) — (6.60):

$$I_{\mathcal{E}} = I_{KK} + I_B(\mathcal{E}) + \tau_{TN}(I, U) \frac{dI_{\mathcal{E}}}{dt} + C_{Je} \frac{dU_{\mathcal{E}}}{dt}, \quad (6.115)$$

$$I_K = I_{KK} - I_B(K) - \tau_{TI}(I, U) \frac{dI_K}{dt} + C_{Jc} \frac{dU_K}{dt}. \quad (6.116)$$

Уравнения (6.113) — (6.116) с соответствующим описанием зависимостей их членов составляют суть иерархической интегральной модели с управлением зарядами. Использование в эквивалентной схеме БТ последовательных сопротивлений областей, влияющих на изменение напряжений на переходах $U_{\mathcal{E}} = U_{\mathcal{E}}(\text{вн}) - IR$, учет микротопологии, распределенных эффектов, уточнение описания инерционных процессов БТ позволяют синтезировать более точную и эффективную модель БТ для машинных расчетов БИС.

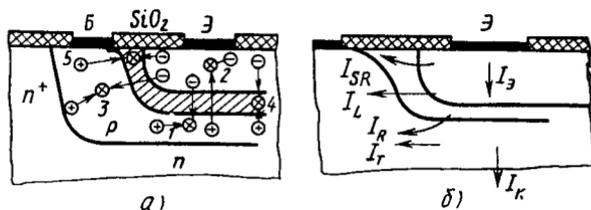


Рис. 6.16. Локализованные рекомбинационные процессы, влияющие на значение коэффициента передачи тока БТ и на характер его зависимости от тока коллектора (а); составляющие токов в БТ (б):

1 — в активной базовой области; 2 — в эмиттере; 3 — в пассивной базовой области; 4 — в объеме эмиттерного перехода; 5 — в приповерхностной области перехода

Рассмотрим рекомбинационные процессы, влияющие на коэффициент передачи тока B БТ в микрорежиме и на характер его зависимости в области малых коллекторных токов I_K (рис. 6.16, а). Процессы рекомбинации в активной 1 области базы, а также в эмиттере 2 определяют основную диффузионную компоненту базового тока I_T , в пассивной 3 области базы — периферийную компоненту базового тока I_L . Процессы генерации — рекомбинации в областях объемного заряда эмиттерного p - n -перехода 4 и выхода перехода на поверхность 5 обуславливают дополнительные компо-

ненты базового тока I_R, I_{SR} . Компоненты токов, протекающих в БТ, показаны на рис. 6.16, б.

Таким образом, полный базовый ток может быть представлен суммой четырех компонент:

$$I_B = I_T + I_B + I_{SR} + I_L, \quad (6.117)$$

а коэффициент передачи тока

$$B = I_K/I_B = I_K/(I_T + I_R + I_{SR} + I_L). \quad (6.118)$$

Для определения рекомбинационно-генерационной компоненты базового тока в объеме p - n -перехода воспользуемся моделью скорости рекомбинации Шокли—Рида—Холла (ШРХ). В данной модели генерация и рекомбинация осуществляются через один ловушечный уровень в середине запрещенной зоны полупроводника посредством следующих этапов: захвата электрона, эмиссии электрона, захвата дырки, эмиссии дырки. При этом скорость рекомбинации R [$\text{см}^{-3} \cdot \text{с}^{-1}$] определяется выражением

$$R = \frac{np - n_{ie}^2}{(n + n_{ie}) \tau_{p0} + (p + n_{ie}) \tau_{n0}}, \quad (6.119)$$

где τ_{p0}, τ_{n0} — времена жизни неосновных носителей заряда.

В равновесном состоянии ток генерации p - n -перехода компенсируется током рекомбинации. При обратном смещении, когда уменьшается концентрация носителей заряда ($pn \ll n_{ie}^2$), доминируют процессы экстракции дырок и электронов. Согласно выражению (6.119), при условии $p < n_{ie}, n < n_{ie}$ скорость генерации пар электрон — дырка

$$R = -[1/(\tau_{p0} + \tau_{n0})] n_{ie} = -n_{ie}/\tau_{эф}, \quad (6.120)$$

где $\tau_{эф}$ — эффективное время жизни носителей заряда в p - n -переходе; $\tau_{эф} \approx 2\tau$ при $\tau_{n0} \approx \tau_{p0} \approx \tau$.

Согласно выражению (6.120), ток генерации p - n -перехода

$$I_g = \int_V q|R| dV \approx \frac{qn_{ie}W_{pn}S\varepsilon}{\tau_{эф}} \approx \frac{qn_{ie}W_{pn}S\varepsilon}{2\tau}, \quad (6.121)$$

где W_{pn} — ширина перехода; $W_{pn} \sim (\varphi_0 + U_\varepsilon)^{1/2}$ для резких и $W_{pn} \sim (\varphi_0 + U_\varepsilon)^{1/3}$ для линейных переходов; φ_0 — контактная разность потенциалов.

С возрастанием обратного смещения компонента I_g увеличивается пропорционально ширине перехода: $I_g = I_{g0}W_{pn}/W_0$, где W_{pn}, W_0 — ширина перехода в неравновесном и равновесном состояниях.

При прямом смещении на эмиттерном p - n -переходе доминирующим является процесс рекомбинации (захват дырок и электронов на ловушечном уровне). В состоянии термодинамического равновесия $p_0n_0 = n_{ie}^2$.

Если к p - n -переходу приложено прямое смещение, а концентрации дырок и электронов определяются через потенциал V и квази-потенциалы Ферми φ_n , φ_p , то, согласно (6.15), произведение

$$pn = n_{ie}^2 \exp\left(\frac{\varphi_p - \varphi_n}{\varphi_T}\right) = n_{ie}^2 \exp\frac{U_{\Delta}}{\varphi_T}, \quad (6.122)$$

причем при $U_{\Delta} > 0$; $\varphi_p - \varphi_n > 0$, $pn > n_{ie}^2$.

Подставив (6.122) в выражение для скорости рекомбинации в p - n -переходе (6.120), получим

$$R = \frac{n_{ie}^2 \exp\left(\frac{U_{\Delta}}{\varphi_T}\right)}{(n + n_{ie})\tau_{p0} + (p + n_{ie})\tau_{n0}} \approx \frac{n_{ie}^2 \exp\frac{U_{\Delta}}{\varphi_T}}{\tau(n + p + 2n_{ie})}. \quad (6.123)$$

Максимальное значение скорости рекомбинации достигается при $V = (\varphi_n + \varphi_p)/2$; $(n + p) \rightarrow \min$. Тогда, согласно соотношениям (6.122), (6.123),

$$R_{\max} = \frac{1}{2} \frac{n_i}{\tau} \exp\frac{U_{\Delta}}{2\varphi_T}. \quad (6.124)$$

Оценку сверху тока рекомбинации в объеме p - n -перехода можно представить в виде

$$I_R = \int_V q |R| dV = \frac{qn_{ie}^2 W_{pn} S_{\Delta}}{2\tau} \exp\frac{U_{\Delta}}{2\varphi_T} = I_{R0} \exp\frac{U_{\Delta}}{2\varphi_T}, \quad (6.125)$$

где $I_{R0} = I_{g0} \frac{qn_{ie}^2 W_{pn} S_{\Delta}}{2\tau}$.

Сделанная оценка позволяет обобщить выражение для этого тока:

$$I_R = I_{R0} \exp\frac{U_{\Delta}}{m_r \varphi_T}, \quad 1 < m_r < 2. \quad (6.126)$$

Качественным объяснением выражения (6.126) служит тот факт, что примерно на металлургической границе p - n -перехода, где произведение $pn = \max$, $R_{\max} \sim \exp\frac{U_{\Delta}}{2\varphi_T}$, а на границах p - n -перехода $R_{\min} \sim \exp\frac{U_{\Delta}}{\varphi_T}$.

Следовательно, интегральная скорость рекомбинации $R \sim \exp\frac{U_{\Delta}}{m_r \varphi_T}$, $1 < m_r < 2$, что и позволяет рассчитать ток рекомбинации I_R согласно (6.126).

Нормальные компоненты плотности тока дырок и электронов связаны со скоростью генерации — рекомбинации на поверхности R_S соотношениями $J_p = -qR_S$, $J_n = qR_S$, где

$$R_S = \frac{p_S n_S - n_{ie}^2}{n + n_{ie}/s_p + (p + n_{ie})/s_n}; \quad (6.127)$$

p_s, n_s — концентрации дырок и электронов на поверхности; s_p, s_n — скорости поверхностной рекомбинации дырок и электронов (эквивалент времени жизни по поверхности полупроводника).

При допущениях, аналогичных сделанным при выводе формулы (6.123): $s_n \approx s_p \approx s$, выражение (6.127) упрощается:

$$R_S = \frac{(np + n_{ie})^2 s}{n_s + p_s + 2n_{ie}}. \quad (6.128)$$

В случае прямого смещения величина

$$R_S = \frac{n_{ie} \exp(U_3 / \varphi_T) s}{n_s + p_s + 2n_{ie}} \quad (6.129)$$

принимает максимальное значение при минимальной сумме $p_s + n_s$. Это условие достигается при $n_s = p_s = n_{ie} \exp(U_3 / 2\varphi_T)$. Тогда

$$R_{S \max} = \frac{1}{2} n_{ie} s \exp \frac{U_3}{2\varphi_T}. \quad (6.130)$$

Оценка сверху тока рекомбинации на поверхности p - n -перехода может быть сделана следующим образом:

$$\begin{aligned} I_{SR \max} &= \int_V q R_{S \max} dV = \frac{1}{2} \times \\ &\times n_{ie} q A_S \exp \frac{U_3}{2\varphi_T} = \\ &= I_{S0} \exp \frac{U_3}{2\varphi_T}, \quad I_{S0} = \frac{1}{2} n_{ie} q A_S, \end{aligned} \quad (6.131)$$

где A_S — эффективная площадь поверхностной рекомбинации. Для относительно небольших областей при выходе p - n -перехода на поверхность эффективная площадь поверхностной рекомбинации практически равна периметру p - n -перехода на поверхности эмиттера, $A_S \sim P$.

В общем виде ток поверхностной рекомбинации можно записать в виде

$$I_{SR} = I_{S0} \left(\exp \frac{U_3}{m_s \varphi_T} - 1 \right), \quad (6.132)$$

где параметр m_s в показателе экспоненты лежит в диапазоне $1 < m_s \leq 2$.

Ток поверхностной рекомбинации определяется скоростью поверхностной рекомбинации s дырок и электронов, а также плотностью заряда в окисле над поверхностью пассивной базы.

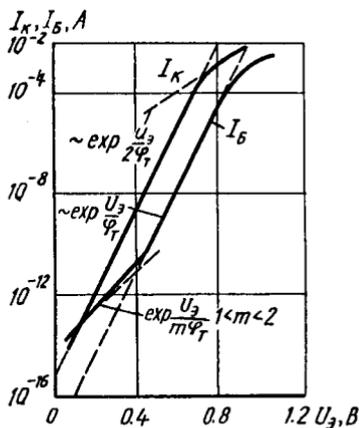


Рис. 6.17. Зависимости токов коллектора и базы от эмиттерного напряжения

Зависимость коэффициента передачи тока B от тока коллектора хорошо наблюдается из сравнения типичных экспериментальных ВАХ БТ $I_K(U_{\text{Э}})$, $I_B(U_{\text{Э}})$ в полулогарифмическом масштабе (рис. 6.17). Практически одинаковый наклон этих характеристик на протяжении одной или нескольких декад $U_{\text{Э}}$ подтверждает тот факт, что для БТ, определенного эффективностью эмиттера, коэффициент B в области средних токов не зависит от тока. На рис. 6.17 пунктирной линией показана диффузионная компонента базового тока $I_T(U_{\text{Э}})$, которая обеспечивала бы максимальное значение B_0 и его постоянство. Однако вследствие доминирующего влияния компонент рекомбинационных токов в объеме и на поверхности p - n -перехода увеличивается наклон ВАХ $I_B(U_{\text{Э}})$ и тем самым уменьшается коэффициент B при малых токах. В некоторых случаях удается выделить две ветви рекомбинационного тока с $m_r > 1$ и $m_s > 1$, $1 < m_r < m_s < 2$. Удобной оценкой, показывающей влияние совместного действия рекомбинационных процессов в объеме и по поверхности p - n -перехода, является эффективный m -фактор, описывающий ВАХ $I_B(U_{\text{Э}}) \approx I_{SR}(U_{\text{Э}})$:

$$I_{SR} \approx I_{SR0} \exp \frac{U_{\text{Э}}}{m\varphi_T}, \quad U_{\text{Э}} > 4m\varphi_T.$$

Если предположить, что в некотором диапазоне напряжений $U_{\text{Э}}$ величины I_{SR0} и m являются постоянными, то

$$m\varphi_T = \frac{dU_{\text{Э}}}{d \ln(I_B/I_{SR0})}. \quad (6.133)$$

Если $I_{SR0} \gg I_{T0} \gg I_{L0}$, то в области малых токов (полагая $I_K = I_{K0} \exp(U_{\text{Э}}/\varphi_T)$) зависимость $B(I_K)$ (см. рис. 6.4, а) дается выражением

$$\frac{1}{B} \approx \frac{I_{SR}}{I_K} = \frac{I_{SR0}}{(I_{T0})^{1/m}} (I_K)^{1/m-1} \quad (6.134)$$

или

$$B = K^{(m-1)/m}, \quad K = \frac{(I_{T0})^{1/m}}{I_{SR0}}.$$

Для инженерных расчетов в области малых токов верна оценка согласно (6.134):

$$B = B_1 (I_K/I_{K1})^{(\bar{m}-1)/\bar{m}}, \quad (6.135)$$

где B_1 — коэффициент при токе $I_K = I_{K1}$, например 50—100 мкА; \bar{m} — среднее значение m -фактора для диапазона токов $I_K < I_{K1}$, внутри которого при уменьшении тока происходит уменьшение коэффициента B .

Согласно выражению (6.135), при $m=2$ $B \approx B_1 (I_K/I_{K1})^{1/2}$, при $m=1.5$ $B \approx B_1 (I_K/I_{K1})^{1/3}$, при $m=1.2$ $B \approx B_1 (I_K/I_{K1})^{1/6}$. Для современных кремниевых БТ $1.1 \leq m \leq 1.6$ в области микротоков.

Периферийная компонента тока базы описывается соотношением $I_L = I_{L0} \exp(U_{\text{Э}} / m_L \varphi_T)$, где фактор неидеальности близок к единице $m_L \approx 1$ в большинстве случаев в микрорежиме практически не играет существенной роли.

С увеличением тока коллектора возникает значительное падение напряжения на сопротивлении активной базы от протекания тока I_B , приводящее к тому, что периферия эмиттерного перехода (ближе к контакту базы) будет находиться под большим напряжением, чем его центральная часть. Большая инжекция периферийной части перехода приводит, во-первых, к возрастанию роли составляющих тока базы I_{SR} и I_L по сравнению с I_T за счет более интенсивной рекомбинации в объеме пассивной базы и на поверхности p - n -перехода. При этом увеличение полного базового тока происходит с повышением внешнего напряжения $U_{\text{Э}}$ быстрее, чем увеличение тока коллектора. В этих условиях относительное увеличение составляющих I_{SR} , I_L (не протекающих по активной базе) может стать серьезной причиной уменьшения коэффициента B на средних и больших токах.

Базовый ток коллекторного перехода $I_B(U_K)$ состоит из идеальной (диффузионной) компоненты в областях базы и коллектора I_{BI} и рекомбинационного тока в объеме коллекторного p - n -перехода с доминированием ШРХ-механизма:

$$I_{BI} = I_{BI0} \left(\exp \frac{U_K}{\varphi_T} - 1 \right), \quad (6.136)$$

$$I_{BKR} = I_{BKR0} \left(\exp \frac{U_K}{m_K \varphi_T} - 1 \right), \quad 1 < m_K < 2. \quad (6.137)$$

Поверхностный ток коллекторного перехода I_{SI} состоит из компонент рекомбинационного тока в местах выхода коллекторного перехода на поверхность I_{SK} и рекомбинационного тока в приповерхностной области пассивной базы I_{LK} , связанного с интенсивной инжекцией электронов из пассивной области коллектора:

$$I_{SK} = I_{SK0} \left(\exp \frac{U_K}{m_{sK} \varphi_T} - 1 \right), \quad 1 < m_{sK} \leq 2, \quad (6.138)$$

$$I_{LK} = I_{LK0} \left(\exp \frac{U_K}{m_{LK} \varphi_T} - 1 \right), \quad m_{LK} \approx 1. \quad (6.139)$$

С помощью составляющих тока базы $I_B(U_K)$ (6.136) — (6.139) аналогично рассмотренному описанию зависимости коэффициента B_N можно описать зависимость коэффициента передачи тока B_I в нормальном инверсном режиме:

$$B_I = \frac{I_{\text{Э}}}{I_B} = \frac{I_{\text{Э}}}{I_{BI} + I_{BKR} + I_{SK} + I_{LK}}, \quad (6.140)$$

причем суть уменьшения B_I в микрорежиме будет отражена рекомбинационными составляющими токов (6.137), (6.138) с коэффициентами неидеальности m_K , m_{sK} , не равными единице.

6.7. МОДЕЛИРОВАНИЕ МНОГОМЕРНЫХ ФИЗИЧЕСКИХ ПРОЦЕССОВ В БЫСТРОДЕЙСТВУЮЩИХ БТ

Физические процессы в БТ можно условно разделить на две группы: одномерные и многомерные. Одномерные физические процессы связаны с переносом основных и неосновных носителей заряда в глубь БТ в направлении от эмиттера к коллектору вдоль координаты x . Двумерные и трехмерные процессы обусловлены специфическими особенностями топологии БТ. Омические контакты интегрального БТ расположены на поверхности кристалла БИС и занимают значительную

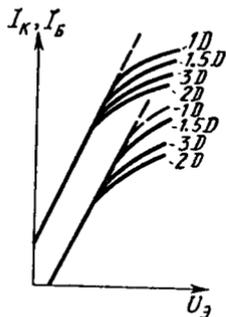


Рис. 6.18. Отличие ВАХ, рассчитанных в одномерном (1D), квазидвумерном (1.5D), двумерном (2D) и трехмерном (3D) приближениях

площадь даже при тенденции увеличения степени интеграции БИС. Следовательно, в реальном БТ помимо переноса подвижных носителей заряда в глубь БТ будет наблюдаться их растекание по областям базы и коллектора (в плоскости двух других координат y, z) под воздействием электрического поля. Электрические характеристики интегральных БТ в значительной степени определяются трехмерностью реального БТ, распределенностью его областей, конфигурацией и геометрическими размерами областей эмиттера, базы, коллектора и омических контактов в плоскости yz кристалла. С увеличением степени интеграции и быстродействия БИС на тонко-слоистых субмикронных БТ роль моделирования трехмерных физических процессов возрастает (рис. 6.18). Однако использование и

двумерных электрических моделей не потеряло актуальности.

Двумерные распределенные электрические модели БТ обеспечивают приемлемую точность расчетов, когда эффектами от протекания токов вдоль одной из координат (z) в ряде случаев можно пренебречь. Такие допущения справедливы в основном для БТ, у которого геометрические размеры по оси z значительно превышают размеры вдоль оси y , что может выполняться даже в сверхинтегрированных сверхскоростных БИС.

Общий подход к моделированию двумерных и трехмерных процессов в БТ. При допущении возможности снижения размерности и пренебрежения токами неосновных носителей заряда в плоскости yz от трехмерной задачи легко можно перейти к двумерной. Проинтегрировав ФСУ в трехмерном приближении в пределах от x_1 до x_2 , получим, что распределение потенциала в некотором диффузионном слое описывается уравнением вида

$$\nabla^2 V(g, z, t) = R_S(y, z) \left[J_{\text{рек}}(y, z, t) + C(V) \frac{\partial}{\partial t} V(y, z, t) \right], \quad (6.141)$$

где $R_S(y, z)$ — поверхностное сопротивление диффузионного слоя;

$J_{\text{рек}}(y, z, t) = q \int_{x_1}^{x_2} R dx$ — плотность тока рекомбинации; $C(V) = \frac{dQ_{1,2}(y, z, t)}{dV}$ — удельная емкость диффузионного слоя; $Q_1(y, z,$

$t) = q \int_{x_1}^{x_2} p(x, y, z, t) dx$, $Q_2 = q \int_{x_1}^{x_2} n(x, y, z, t) dx$ — плотности заряда

основных носителей. С учетом того, что сумма отдельных токов в правой части уравнения (6.141) равна суммарному току, инжектированному в некоторую диффузионную область p - n -переходами, уравнения для потенциала в базе и коллекторе можно привести к виду (рис. 6.19, а):

а) для активной области БТ (под дном эмиттера — область 1)

$$\nabla^2 U'_{\text{БЭ}} = R_{\text{СБЭ}} \left[(1 - \alpha'_N) J'_{\text{Э}} + (1 - \alpha'_I) J'_K + C'_{\text{Э}} \frac{dU'_{\text{БЭ}}}{dt} + C'_K \frac{dU'_{\text{БК}}}{dt} \right], \quad (6.142)$$

$$\nabla^2 U'_{\text{БК}} = R_{\text{СК}} \left[-J'_K + \alpha'_N J'_{\text{Э}} - (C'_K + C'_\Pi) \frac{dU'_{\text{БК}}}{dt} + C'_\Pi \frac{dU'_{\text{БЭ}}}{dt} \right]; \quad (6.143)$$

б) для пассивных областей вне площади эмиттера (области 2—4)

$$\nabla^2 U''_{\text{БЭ}} = R''_{\text{СБП}} \left[J''_K + C''_K \frac{dU''_{\text{БК}}}{dt} \right], \quad (6.144)$$

$$\nabla^2 U''_{\text{БК}} = R''_{\text{СК}} \left[-J''_K - (C''_K + C''_\Pi) \frac{dU''_{\text{БК}}}{dt} + C''_\Pi \frac{dU''_{\text{БЭ}}}{dt} \right]; \quad (6.145)$$

в) для периферийных областей базы (область 5)

$$\nabla^2 U''_{\text{БК}} = R''_{\text{СК}} C''_\Pi \frac{d}{dt} (U''_{\text{БЭ}} - U''_{\text{БК}}), \quad (6.146)$$

где $\nabla^2 = \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ — оператор дифференцирования, не зависящий от времени; $R_{\text{СБЭ}}$, $R_{\text{СБП}}$, $R_{\text{СК}}$ — поверхностные сопротивления слоев базы и коллектора; $J_{\text{Э}}$, J_K — плотности токов эмиттерного и коллекторного переходов; $C_{\text{Э}}$, C_K , C_Π — полные емкости эмиттерного, коллекторного и подложечного переходов.

В описанной двумерной задаче распределение потенциалов в базовой и коллекторной областях ($U_{\text{БЭ}}$, $U_{\text{БК}}$) может быть получено двумя методами: 1) непосредственным численным решением системы нелинейных дифференциальных уравнений в частных производных (6.141) — (6.146) с граничными и начальными условия-

но высокую точность расчета переходных процессов в схемах, если корректно проведена идентификация элементов внутренней и внешней областей БТ.

Стандартные электрические модели БТ с одним сосредоточенным сопротивлением базовой области, даже достаточно развитые, типа интегральной зарядоуправляемой модели Гуммеля — Пуна, BIRD, отражающие многие эффекты высокого уровня инжекции, являются довольно грубыми для расчетов на ЭВМ в первую очередь быстродействующих схем в области малых времен. Основная причина неудовлетворительной точности моделирования электрических процессов в БТ исследуемых фрагментов БИС — неучет не только доминирующих эффектов вытеснения в статике и динамике, специфики распределенности БТ, но также фазовой задержки в базовой и коллекторной областях БТ, эффектов накопления зарядов в ряде характерных областей БТ, модификации эффектов высокого уровня инжекции.

Модель статического эффекта вытеснения эмиттерного тока.

Суть статического эффекта вытеснения эмиттерного тока заключается в следующем. Базовый ток I_B , протекающий параллельно плоскости эмиттерного $p-n$ -перехода в направлении y , создает падение напряжения на высокоомном слое активной базы, что приводит к зависимости напряжения эмиттерного $p-n$ -перехода от координаты y : $U_{\Sigma}(y) = U_{\Sigma}(0) - I_B(y)R_B(I)$. Отметим, что, как было выяснено ранее, по активной базе БТ протекает не весь базовый ток, а только его часть за вычетом I_{SR} и I_L . Эта оговорка существенна, поскольку в дальнейших расчетах будет фигурировать параметр тока I_B , под которым будем подразумевать $I_B \approx I_T$. Далее напряжение, приложенное к краю эмиттерного перехода, равно разности внешнего приложенного напряжения от источника напряжения $U_{вн}$ и падения напряжения от протекания тока базы на практически неизменном сопротивлении пассивной базы $R_{БП}$ т. е. $U_{\Sigma}(y) = U_{\Sigma вн} - I_B R_{БП} - I_B(y)R_{БЛ}(I)$ (рис. 6.19, а). Поскольку плотность эмиттерного тока является экспоненциальной функцией приложенного напряжения, небольшие падения напряжения $I_B(y)R_{БЛ}(I)$, порядка долей и единиц фТ, могут привести к сильному влиянию распределенного напряжения $U_{\Sigma}(y)$ на плотность эмиттерного тока. Наибольшее напряжение U_{Σ} будет на краю эмиттерного перехода около базового контакта и наименьшее — в отдалении от него. При этом эмиттерный ток вытесняется к краю (со стороны базового контакта перехода). Если БТ имеет два базовых контакта, то минимальная плотность тока будет в центре БТ, а максимальная — по его краям (рис. 6.19, а). Для удобства аналитического описания статического эффекта вытеснения эмиттерного тока сделаем ряд непринципиальных допущений: 1) коэффициент передачи тока α_N (B или B_N) в активном нормальном режиме не зависит от координаты y и тока; 2) влияние эффектов высокого уровня инжекции в базе и коллектора пренебрежимо мало; 3) влияние температурных градиентов в БТ

вследствие рассеяния мощности пренебрежимо мало. Данные допущения могут быть сняты, однако это приведет к некоторому усложнению полученных инженерных формул. Кроме того, допущение о пренебрежимо малом влиянии эффекта модуляции проводимости базы соответствует реальной ситуации моделирования эффекта вытеснения в тонкослойных БТ с сильнолегированной базой: $S_A \approx 10^{18} \div 10^{19} \text{ см}^{-3}$. В дальнейшем под $U_{\text{Э}}(0)$ будем подразумевать напряжение $U_{\text{вн}} - I_{\text{Б}} R_{\text{бп}} - I_{\text{Э}} R_{\text{Э}}$ с учетом падений напряжений в омических областях пассивной базы и эмиттера.

Плотность эмиттерного тока

$$J_{\text{Э}}(y) = J_{\text{Э}0} \exp \frac{U_{\text{Э}}(y)}{\varphi_T}, \quad (6.147)$$

где $J_{\text{Э}0}$ — плотность тока насыщения эмиттера, не зависящая от координаты y и $J_{\text{Э}0}$.

Приращение падения напряжения от протекания базового тока $I_{\text{Б}y}$ можно выразить как

$$dU_{\text{Б}}(y) = -\frac{R_{\text{СБА}}}{l} I_{\text{Б}}(y) dx, \quad (6.148)$$

где $R_{\text{СБА}}$ — поверхностное сопротивление активной базы.

Напряжение на эмиттерном p - n -переходе, зависящее от координаты y , с учетом выражения (6.148) можно представить в виде

$$U_{\text{Э}}(y) = U_{\text{Э}0} - \int_0^y \frac{R_{\text{СБА}}}{l} I_{\text{Б}}(y) dx. \quad (6.149)$$

Изменение базового тока $dI_{\text{Б}}(y)$ пропорционально плотности эмиттерного тока в точке y :

$$dI_{\text{Б}}(y) = -(1 - \alpha_N) J_{\text{Э}}(y) l dy. \quad (6.150)$$

Используя выражения (6.147), (6.149), (6.150) и комбинируя члены при производных $\frac{d^2 I_{\text{Б}}(y)}{dy^2}$, $\frac{dI_{\text{Б}}(y)}{dy}$, получим обыкновенное дифференциальное уравнение второго порядка

$$\frac{d^2 I_{\text{Б}}(y)}{dy^2} + \frac{R_{\text{СБА}}}{\varphi_T l} I_{\text{Б}}(y) + I_{\text{Б}}(y) \frac{dI_{\text{Б}}(y)}{dy} = 0, \quad (6.151)$$

которое удовлетворяет двум граничным условиям:

$$I_{\text{Б}}(L) = 0, \quad I_{\text{Б}}(0) = I_{\text{Б}}. \quad (6.152)$$

Аналогично из выражений (6.147), (6.149), (6.150) получается дифференциальное уравнение для напряжения, распределенного вдоль эмиттерного перехода:

$$\frac{d^2 U_{\text{Э}}(y)}{dy^2} = R_{\text{СБА}} \frac{\hbar}{l} I_{\text{Э}0} \exp \frac{U_{\text{Э}}(y)}{\varphi_T}, \quad (6.153)$$

удовлетворяющее граничным условиям

$$U_{\text{Э}}(y=0) = U_{\text{Э}}, \quad \left. \frac{dU_{\text{Э}}(y)}{dy} \right|_{x=h} = 0. \quad (6.154)$$

Уравнение (6.151) с граничными условиями (6.152) имеет решение

$$I_{\text{Б}}(y) = \frac{2\varphi_T l}{hR_{\text{СБА}}} z \operatorname{tg} \left(z \left(1 - \frac{y}{h} \right) \right), \quad (6.155)$$

где параметр z определяется из дополнительного уравнения

$$z \operatorname{tg} z = \frac{I_{\text{Б}} h R_{\text{СБА}}}{2\varphi_T l} = \frac{a}{2}, \quad a = \frac{R_{\text{Б}} I_{\text{Б}}}{\varphi_T}. \quad (6.156)$$

Напряжение вдоль эмиттерного p - n -перехода $U_{\text{Э}}(y)$ находят из решения (6.153) с граничными условиями (6.154) или при использовании уравнения (6.153) в соотношении (6.149) и выполнении интегрирования:

$$U_{\text{Э}}(y) = U_{\text{Э}}(0) - 2\varphi_T \ln \frac{\cos(z(1-y/h))}{\cos z}. \quad (6.157)$$

Плотность эмиттерного тока рассчитывают при использовании выражения (6.157) в соотношении (6.147):

$$J_{\text{Э}}(y) = J_{\text{Э}0} \left(\exp \frac{U_{\text{Э}}(0)}{\varphi_T} \right) \frac{\cos^2 z}{\cos^2(z(1-y/h))}. \quad (6.158)$$

Уравнения (6.155)–(6.158) характеризуют распределение токов базы и эмиттера, напряжения эмиттерного p - n -перехода вдоль перехода от $y=0$ до $y=h$, а также эквивалент распределенного сопротивления базы, используемый в электрических моделях БТ.

Модели низкочастотных распределенных сопротивлений базы БТ с полосковым эмиттером. Определим эквивалентное базовое сопротивление по мощности БТ $R_{\text{Б1}}$ как сопротивление, рассеивающее мощность при протекании базового тока по активной базовой области $P_{\text{Б}}$:

$$R_{\text{Б1}} = P_{\text{Б}} / I_{\text{Б}}^2, \quad (6.159)$$

где

$$P_{\text{Б}} = \int_0^h \frac{R_{\text{СБА}} I_{\text{Б}}^2(y)}{l} dy. \quad (6.160)$$

Используя выражение (6.155) в (6.160), интегрируя его и учитывая (6.157), получим

$$R_{\text{Б1}} = \frac{hR_{\text{СБА}}(\operatorname{tg} z - z)}{lz \operatorname{tg}^2 z}. \quad (6.161)$$

Определим статическое эквивалентное базовое сопротивление как последовательное сопротивление $R_{Б2}$ в эквивалентной схеме для идеального диода с ВАХ $U_{Э}(0) = \varphi_T \ln(I_{Э}/I_0)$:

$$U'_{Э} = U_{Э}(0) - I_{Б2} R_{Б2}, \quad R_{Б2} = \frac{U_{Э}(0) - U'_{Э}}{I_{Б}}. \quad (6.162)$$

ВАХ базового тока не зависит от эффектов высокого уровня инжекции в БТ, поэтому

$$I_{Б} = I_{\tau 0} \exp \left[\frac{1}{\varphi_T} (U_{Э}(0) - I_{Б} R_{Б2}) \right]. \quad (6.163)$$

Подставляя (6.157) в (6.150), получим решение для тока базы при $I_{\tau 0} = -(1 - \alpha_N h I J_{Э0})$:

$$I_{Б} = -(1 - \alpha_N) h I J_{Э0} \frac{\sin z \cdot \cos z}{z} \exp \frac{U_{Э0}}{\varphi_T}. \quad (6.164)$$

Эквивалентное соотношение для $I_{Б} R_{Б2}$ можно получить, преобразуя напряжение (6.157) к виду

$$U_{Э} = U_{Э}(0) - U_{Б}, \quad U_{Б} = \varphi_T \ln \frac{z}{\sin z \cdot \cos z} = I_{Б} R_{Б2}. \quad (6.165)$$

Из выражений (6.164), (6.165) следует, что

$$R_{Б2} = \frac{U_{Б}}{I_{Б}} = \frac{h R_{SBA}}{I} \frac{1}{2z \operatorname{tg} z} \ln \frac{z}{\sin z \cdot \cos z} = \frac{R}{a} \ln \frac{z}{\sin z \cdot \cos z}. \quad (6.166)$$

Определим динамическое (малосигнальное) базовое сопротивление $R_{Б3}$ как низкочастотное сопротивление, обеспечивающее выполнение на выводах БТ условия

$$\frac{dU_{Э}(0)}{dI_{Б}} = R_{Б3} + \frac{dU_{Э}}{dI_{Б}}, \quad (6.167)$$

где $\frac{dU_{Э}}{dI_{Б}}$ — дифференциальное сопротивление перехода для случая пренебрежимо малого статического падения напряжения в базовой области.

Решением (6.167) после соответствующих преобразований уравнений (6.155), (6.157) является выражение

$$R_{Б3} = \left(\frac{R_{SBA} h}{z I} \right) \left(\frac{z \sin z + \frac{1}{z} \cos z - \operatorname{cosec} z}{\sin z + z \cos z + z \sin z \operatorname{tg} z} \right). \quad (6.168)$$

На рис. 6.20 представлены зависимости эквивалентных базовых сопротивлений от нормированного базового тока $\frac{I_{Б} h R_{SBA}}{2 \varphi_T I} =$

$$= z \operatorname{tg} z = \frac{a}{2}, \quad \text{где } R = \frac{R_{SBA} h}{I}, \quad a = \frac{I_{Б} R_{Б}}{\varphi_T}.$$

Для малых токов ($I_B \rightarrow 0, z \rightarrow 0$) все эквивалентные сопротивления равны $1/3$ полного сопротивления базы R :

$$\lim_{a \rightarrow 0} R_{B_i} = \frac{R}{3}, \quad R_{B_1}(0) = R_{B_2}(0) = R_{B_3}(0) = \frac{R_{SBA} h}{1.3} = \frac{R}{3}. \quad (6.169)$$

Эквивалентные сопротивления начинают отклоняться от $R/3$ при значениях $z \operatorname{tg} z \approx 0.5$, соответствующих $z \approx 0.65$ рад, или $a^* = a/2$. При больших токах и сопротивлениях R они становятся пренебрежимо малыми ($R_{B_i} \rightarrow 0$) уже при $a \geq 50$.

При больших значениях $a \gg 1$ ($z \rightarrow \pi/2, \sin z = 1$) выражение для эквивалентного сопротивления R_{B_1} упрощается:

$$R_{B_1} |_{a \gg 1} \approx 2\varphi_T / I_B \quad \text{при}$$

$$I_B \gg \varphi_T / R. \quad (6.170)$$

ВАХ нелинейного статического сопротивления R_{B_2} описывается характеристикой диода

$$I_{B_2} = \frac{2\varphi_T}{R} \exp \frac{U_B}{\varphi_T}, \quad \frac{dI_{B_2}}{dU_B} = \frac{\varphi_T}{2\varphi_T / R + I_{B_2}}. \quad (6.171)$$

Динамическое сопротивление R_{B_3} (6.167), изменяясь при больших токах согласно закону

$$R_{B_3} \sim \varphi_T / I_B, \quad (6.172)$$

быстрее сопротивлений R_{B_1}, R_{B_2} достигает наименьшего, слабо изменяющегося после $a \geq 30$ значения.

Формально для исследования процессов на малом сигнале в модели БТ необходимо использовать выражение R_{B_3} (6.168), в стационарном режиме — R_{B_2} (6.166), для расчета, например, шумов в БТ — R_{B_1} (6.161). Однако вследствие схожего характера изменения всех сопротивлений $R_{B_i} \sim \varphi_T / I_B$ в области больших токов ($a \gg 1$) для простоты описания нелинейной зависимости R_B целесообразно применять ВАХ R_{B_i} типа «квазидиодной» [см. выражение (6.171)]. Но использование в «квазидиодной» ВАХ обратного тока $I_{B_0} = 2\varphi_T / R$ приводит к значительной погрешности моделирования R_{B_i} (до 50%) в области $a \ll 1$, где $\lim R_{B_i} = R/3$, поскольку для ВАХ (6.171) $dU_B / dI_{B_2} = R/2$ при $a \ll 1$. Основным уравнениям модели распределенных сопротивлений базы (6.169), (6.172) можно поставить в соответствие универсальные электрические эквивалентные схемы БТ с нелинейным эквивалентным сопротивлением (рис. 6.21, а) согласно определениям R_{B_1} (6.161),

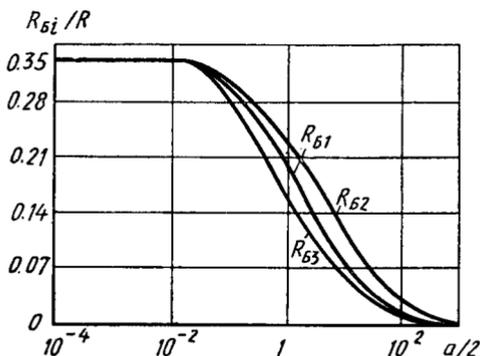


Рис. 6.20. Зависимости нормированных сопротивлений базы от параметра

$R_{Б2}$ (6.166), $R_{Б3}$ (6.168), несколько упрощенным описанием $R_{Б}$ диодом $D_{Б2}$ [см. формулу (6.171)], шунтированным резистором R (рис. 6.21, б) и диодом $D_{Б}$ (рис. 6.21, в) с ВАХ

$$I_{Б} = \frac{3\varphi_T}{R} \exp U_{Б}, \quad r_{Б} = \frac{\varphi_T}{3\varphi_T/R + I_{Б}}. \quad (6.173)$$

Поскольку погрешность аппроксимации

$$U_{Б} = \varphi_T \ln \frac{z}{\sin z \cdot \cos z} \approx \varphi_T \ln \left(1 + \frac{a}{3} \right) \quad (6.174)$$

составляет менее 10% при $a \leq 10$, менее 1% при $a \leq 1$, а также принимая в расчет, что условие $a \gg 10$ $dU_{Б}/dI_{Б} \approx \varphi_T/I_{Б}$ удовлетворяет рассмотренным асимптотическим выражениям (6.170)—(6.172), с помощью ВАХ (6.173) можно эффективно, наиболее просто, универсально описать изменения в области больших токов всех распределенных сопротивлений базы согласно принятым определениям.

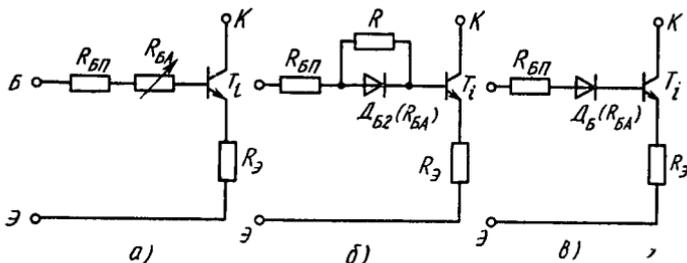


Рис. 6.21. Низкочастотные эквивалентные схемы распределенного сопротивления базы в электрической модели БТ

С целью уточнения низкочастотной модели распределенного сопротивления базы БТ для больших токов необходимо оценить отношение R/V_0 в случае влияния эффектов высокого уровня инжекции — модуляции проводимости и расширения базы. Влияние эффекта модуляции проводимости базы начинает сказываться начиная с некоторого значения плотности критического коллекторного тока J_N

$$J_N \approx \frac{1.4}{R} \left(\frac{h}{W_{Б}} \right)^2, \quad (6.175)$$

когда в активной области базы концентрации подвижных носителей заряда — электронов и дырок — приблизительно равны ($n \approx p$) и вкладом ионизированных акцепторных примесей в общий заряд базы можно пренебречь. Данный эффект приводит к значительному уменьшению поверхностного сопротивления активной базы $R_{СБА} \sim 2,8 \varphi_T/J_K$, что несколько компенсирует влияние эффекта статического вытеснения тока.

Будем считать, что влияние эффекта расширения базы будет сказываться начиная с критических плотностей тока (6.88), при этом эффективная ширина базы

$$W_{\text{Бэф}} = W_{\text{Б}} + W_{\text{К}}(1 - J_{\text{кр}}/J_{\text{К}}). \quad (6.176)$$

Подставив (6.175) в выражение для полного сопротивления базы $R = R_{\text{СБА}} h/l$, получим его оценку в условиях проявления только одного эффекта модуляции проводимости базы:

$$R \approx \frac{1.4\varphi_{\text{T}}}{W_{\text{Б}}^2 (l/h) J_{\text{К}}}, \quad (6.177)$$

причем значение этого сопротивления обратно пропорционально плотности режимного коллекторного тока.

В условиях совместного влияния двух эффектов: модуляции проводимости и расширения базы, подставляя (6.88), (6.176) в выражения для R , получим новое соотношение: при $\frac{W_{\text{К}}}{W_{\text{Б}}} \times$

$$\times \left(\frac{J_{\text{К}} - J_{\text{кр1}}}{J_{\text{К}}} \right) \gg 1$$

$$R \approx \frac{1.4\varphi_{\text{T}}}{W_{\text{К}}^2 (l/h) (J_{\text{К}} - J_{\text{кр}})^2}, \quad (6.178)$$

причем значение этого сопротивления обратно пропорционально плотности режимного коллекторного тока.

Воспользовавшись аппроксимативными выражениями для коэффициента передачи тока $B \sim B_0/J_{\text{К}}$ в случае проявления одного эффекта модуляции, $B \sim B_0/J_{\text{К}}^2$ — в случае совместного проявления двух эффектов высокого уровня инжекции, где B_0 — коэффициент передачи тока внутреннего БТ при малых токах, а также соотношениями (6.177), (6.178), легко показать, что практически во всем диапазоне токов поддерживается постоянство отношения

$$R/B_0 \approx \text{const}. \quad (6.179)$$

Данный факт позволяет считать решения базовых дифференциальных уравнений (6.151), (6.153) для всех $R_{\text{Б1}}$ инвариантными от уровня инжекции, так как

$$a = R \frac{l_{\text{Б}}}{\varphi_{\text{T}}} = \frac{R}{B_0} \frac{l_{\text{К}}}{\varphi_{\text{T}}} \approx \text{const},$$

и использовать эквивалентные схемы рис. 6.21 в широком режимном диапазоне с минимальной погрешностью.

Модель динамического эффекта вытеснения эмиттерного тока в БТ. На характеристики БТ, работающего на высоких частотах в режиме малого сигнала, или при подаче на его вход импульсных сигналов с крутыми фронтами может сильно влиять эффект вытеснения переменного тока или динамический эффект вытеснения тока в области малых времен переходного процесса. Поперечная составляющая базового тока проходит через RC -линию с распределенными параметрами (однородную или неоднородную), об-

разованную сопротивлением активной базовой области и эмиттерной (в основном диффузионной) емкостью. Статический эффект вытеснения модифицирует явление динамического вытеснения эмиттерного тока. Поскольку диффузионная емкость пропорциональна эмиттерному току, вследствие статического эффекта вытеснения значения диффузионных емкостей на краю эмиттера, вблизи базового контакта, будут выше, чем в других областях перехода, поэтому переменная составляющая эмиттерного тока

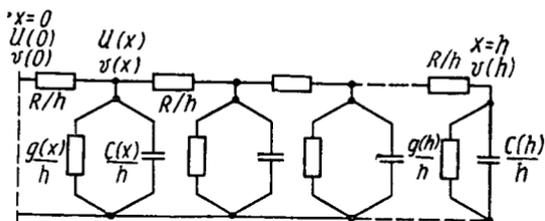


Рис. 6.22. Линия передачи, моделирующая активную базовую область БТ ($g(x)/h$, $c(x)/h$ — удельные проводимости и емкости)

будет концентрироваться на краю перехода, уменьшая эффективную площадь эмиттера, базовое сопротивление и влияя на другие параметры БТ.

Рассмотрим влияние динамического эффекта вытеснения на характеристики БТ с помощью однородной линии передачи, моделирующей распределение свойства актив-

тивной базовой области (рис. 6.22).

Однородная линия передачи описывается дифференциальным уравнением в операторном виде относительно переменного напряжения на переходе $v(x, p)$

$$\frac{d^2 v(x, p)}{dx^2} = \frac{1}{h^2} pRCv(x, p), \quad (6.180)$$

удовлетворяющим двум граничным условиям:

$$v(x, p)|_{x=0} = v_0, \quad \left. \frac{dv(x, p)}{dx} \right|_{x=h} = 0, \quad (6.181)$$

где v_0 — переменное напряжение, приложенное к p - n -переходу.

Решение дифференциального уравнения (6.180) имеет вид

$$v(x, p) = \frac{\text{ch}[(p\tau)^{1/2}(h-x)]}{\text{ch}[(p\tau)^{1/2}h]} v_0, \quad (6.182)$$

где $\tau = RC$ — постоянная времени однородной линии.

Модуль нормированного переменного напряжения в частотном диапазоне рассчитывают из выражения (6.182) путем замены оператора p на $j\omega$, где $\omega = 2\pi f$ — угловая частота, и записывают таким образом:

$$\left| \frac{v(x, \omega)}{v_0(x, \omega)} \right| = \frac{\left[\text{sh}^2 \left(\left(\frac{\omega RC}{2} \right)^{1/2} (h-x) \right) + \text{ch}^2 \left(\left(\frac{\omega RC}{2} \right)^{1/2} (h-x) \right) \right]^{1/2}}{\left[\text{sh}^2 \left(\left(\frac{\omega RC}{2} \right)^{1/2} h \right) + \text{ch}^2 \left(\left(\frac{\omega RC}{2} \right)^{1/2} h \right) \right]^{1/2}}. \quad (6.183)$$

Из выражений (6.182), (6.183) следует, что на высоких частотах (в режиме малых времен) или при увеличении сопротивления активной базовой области переменное напряжение значительно уменьшается вдоль эмиттерного перехода от базового контакта (рис. 6.23). Например, при $\omega RC \leq 0.3$ отношение переменного напряжения в центре эмиттерного перехода к его значению на краю практически неизменно: $N = \left| \frac{v(x, \omega)}{v_0(\omega)} \right| \approx 1$, однако при $\omega RC = 2$ $N = 0.6$, при $\omega RC = 10$ $N = 0.1$. В стандартных малосигнальных схе-

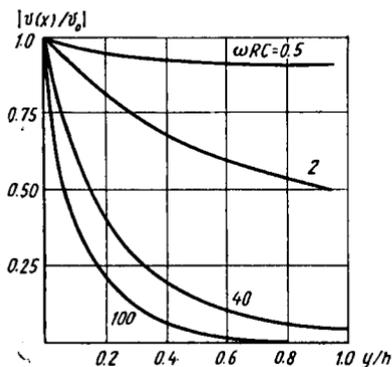


Рис. 6.23. Зависимости нормированного переменного напряжения эмиттера от координаты для различных значений параметра ωRC

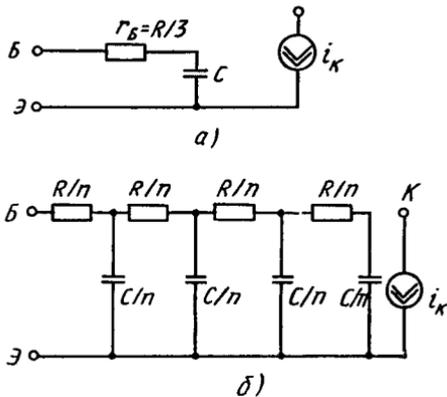


Рис. 6.24. Схемы односекционной малосигнальной модели БТ (а) и однородной малосигнальной распределенной модели БТ (б)

мах замещения, не учитывающих влияния динамического эффекта вытеснения, сопротивление базы представляется не зависящим от частоты и принимается равным $r_b = R_{Бз} = R/3$ (6.169) (рис. 6.24, а), что приводит к существенным погрешностям модели на высоких частотах (50—100% при $100 \leq f \leq 1000$ мГц). Идеализированная модель с распределенным сопротивлением базы (рис. 6.24, б), в которой R/n , C/n — эквивалентные сопротивления и емкости i -й секции, при $n \rightarrow \infty$ точно описывает поведение реальных БТ на высоких частотах, однако при $n > 5 \div 10$ слишком сложна для машинного расчета схем.

Для синтеза более простых, компактных, но достаточно точных малосигнальных моделей, учитывающих динамический эффект вытеснения в заданном частотном (временном) диапазоне, воспользуемся соотношением для входного сопротивления однородной линии передачи (см. рис. 6.22):

$$z_{вх}(p) = \frac{v_0}{i_{вх}(p)}, \quad i_{вх}(p) = -\frac{1}{R} \left. \frac{dv(x, p)}{dx} \right|_{x=0},$$

$$z_{вх}(p) = \frac{R \operatorname{cth}(p\tau)^{1/2}}{(p/\tau)^{1/2}}. \quad (6.184)$$

Используя разложение в ряд гиперболической функции относительно $p\tau$, $p=j\omega$, выражение (6.184) преобразуем к виду

$$z_{вх}(j\omega) = \frac{1 + \frac{j\omega\tau}{2} + \frac{(j\omega\tau)^2}{24} + \dots + \frac{(j\omega\tau)^n}{2n!} + \dots}{j\omega C \left[1 + \frac{j\omega\tau}{6} + \frac{(j\omega\tau)^2}{120} + \dots + \frac{(j\omega\tau)^{2n}}{(2n+1)!} + \dots \right]} \quad (6.185)$$

Сопоставление выражений (6.185) (с конечным числом членов разложения в числителе и знаменателе) и (6.184) позволяет оценить точность представления модуля

и фазы входного сопротивления n -секционной распределенной модели с конечным числом секций относительно идеальной модели с $n \rightarrow \infty$. Расхождение по модулю входного сопротивления двух распределенных моделей ($n \neq \infty$ и $n \rightarrow \infty$) при $n=12$ составляет менее 2%, при $n=8$ — менее 50%, при $n=4$ — менее 20% для частот $f \leq 3/\tau$. Точность представления фазы входного сопротивления более чувствительна к числу секций распределенной модели транзистора ($n \neq \infty$), чем модуль $z_{вх}$. Для той же точности представления фазы входного сопротивления, что и модуля, если это необходимо, в распределенную модель БТ требуется добавить дополнительно еще 1—2 секции. Выбором распределенной модели из трех секций достигается компромисс между точностью (20% по модулю, 30% по фазе) и сложностью данной модели (рис. 6.24, б) в достаточно широком диапазоне частот $0 < f < 1.5/\tau$.

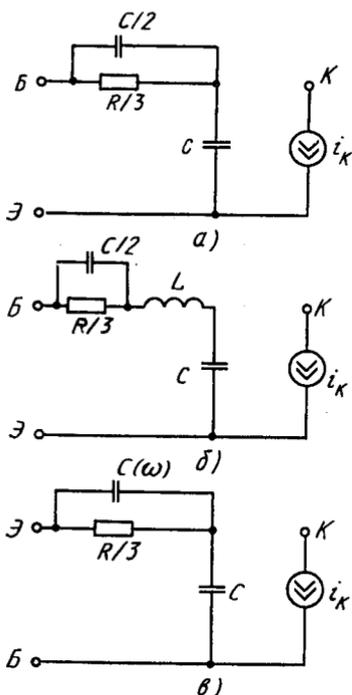


Рис. 6.25. Малосигнальные эквивалентные схемы БТ с различными формами замещения базового сопротивления в соответствующих частотных диапазонах

специфическими сосредоточенными элементами — эффективными реактивностями.

При $f < 0.15/\tau$, т. е. в режиме низких частот, из (6.185) получим простое выражение для входного сопротивления (рис. 6.24):

$$z_{вх} \approx \frac{1}{j\omega C} \left[1 + \frac{j\omega\tau}{3} \right] = \frac{R}{3} + \frac{1}{j\omega C} = r_B + \frac{1}{j\omega C} \quad (6.186)$$

При $f \leq 0.15\tau$ с учетом дополнительных членов разложения выражение входного сопротивления (6.185) принимает вид (рис. 6.25, а)

$$z_{вх} \approx \frac{1}{j\omega C} \left[1 + \frac{\frac{1}{3} j\omega\tau}{1 + \frac{1}{6} j\omega\tau} \right] = \frac{R}{3} \left\| \left(\frac{j\omega C}{2} \right)^{-1} + \frac{1}{j\omega C} \right\| = \quad (6.187)$$

$$= z_{Б0} + \frac{1}{j\omega C}.$$

Включение нескольких эффективных (частотонезависимых) реактивных компонентов в малосигнальную модель БТ позволяет учесть динамический эффект вытеснения в более широком диапазоне частот. Например, выражение для входного сопротивления модели с сосредоточенными параметрами (рис. 6.25, б)

$$z_{вх} = \frac{1}{j\omega C} \frac{1 + j\omega\tau/2 + (j\omega\tau)^2/24}{1 + j\omega\tau/6 + (j\omega\tau)^2/120} =$$

$$= \frac{1}{j\omega C} \left[1 + \frac{j\omega\tau/3}{1 + j\omega\tau/6} + \frac{(j\omega\tau)^2}{48} \right] = z_{Б0} + j\omega L + \frac{1}{j\omega C}, \quad (6.188)$$

где $C_B = C/2$, $L = R^2 C/48$, дает расхождение по модулю с выражением (6.185) ($n \rightarrow \infty$) менее 1% при $f = 0.3/\tau$, 6% при $f = 0.7/\tau$, 20% при $f = 1.5/\tau$.

С помощью базового сопротивления, представленного эффективной базовой емкостью, зависящей от частоты, включенной параллельно дифференциальному сопротивлению базы $r_B = R/3$, можно построить точную малосигнальную схему замещения БТ, учитывающую динамический эффект вытеснения для практически любого диапазона высоких частот (рис. 6.25, а—в). При $|j\omega\tau| > 2$, согласно (6.184), $z_{вх} \approx \frac{R}{3} \left(\frac{j\omega\tau}{9} \right)^{-1/2}$, что позволяет построить частотные зависимости базового сопротивления и базовой емкости:

$$z_B = \frac{R/3}{(1 + j\omega\tau/9)^{1/2}}, \quad C_B = \left[\frac{3}{j\omega\tau} (1 + j\omega\tau/9)^{1/2} - 1 \right] C. \quad (6.189)$$

Из анализа выражений (6.189) следует, что для рабочих частот $f < 3/\tau$ как базовое сопротивление, так и базовая емкость уменьшаются по амплитуде с ростом частоты: $z_B(\omega)$, $C_B(\omega) \sim \omega^{-1/2}$. Использование усредненной постоянной базовой емкости $\bar{C}_B = (0.25 \dots 0.1)C$ в диапазоне частот $0.3/\tau < f < 3/\tau$, как следует из выражения (6.180), дает погрешность относительно $C_B(\omega)$ не более 15%. В диапазоне частот $0.3/\tau < f < 1.5/\tau$, воспользовавшись разложением в ряд выражений (6.189), (6.190) относительно $j\omega\tau$:

$$z_B \approx \frac{R/3}{1 + j\omega\tau/18} = \frac{R}{3} \left\| \left(\frac{j\omega C}{6} \right)^{-1} \right\|, \quad (6.190)$$

$$C_B \approx \left[\frac{3}{j\omega\tau} \left(1 + \frac{j\omega\tau}{18} \right) - 1 \right] C = \frac{C}{6},$$

базовое сопротивление с точностью до 10—15% можно представить параллельным соединением дифференциального базового сопротивления и постоянной эффективной базовой емкости.

В более общем случае активную базовую область высокочастотного БТ нужно моделировать неоднородной линией передачи (с учетом эффекта вытеснения в статике).

Неоднородная линия передачи описывается дифференциальным уравнением в операторном виде в терминах определенных ранее параметров относительно переменного напряжения на переходе $v(x, p)$

$$\frac{d^2v(x, p)}{dx^2} = p(RCM)v(x, p) \frac{1}{h^2}, \quad M \approx \frac{3+a}{a}, \quad (6.191)$$

удовлетворяющим аналогичным граничным условиям для однородной линии.

Решение дифференциального уравнения (6.191) имеет вид ($a \leq 8$, $a = RI_B / \varphi_T = RI_K / \varphi_T B_0$)

$$v(x, p) \approx \frac{\operatorname{ch}[(pRC(3+a/3))^{1/2}(h-x)]}{\operatorname{ch}[(pRC)^{1/2}h]} v_0. \quad (6.192)$$

Модуль нормированного переменного напряжения

$$\begin{aligned} & \left| \frac{v(x, \omega)}{v_0(\omega)} \right| = \\ & = \frac{\left\{ \operatorname{sh}^2 \left[\left(\frac{\omega RC(3+a)}{6} \right)^{1/2} (h-x) \right] + \operatorname{ch}^2 \left[\left(\frac{\omega RC(3+a)}{6} \right)^{1/2} (h-x) \right] \right\}^{1/2}}{\left\{ \operatorname{sh}^2 \left[\left(\frac{\omega RC(3+a)}{6} \right)^{1/2} h \right] + \operatorname{ch}^2 \left[\left(\frac{\omega RC(3+a)}{6} \right)^{1/2} h \right] \right\}^{1/2}}. \end{aligned} \quad (6.193)$$

Из выражений (6.191), (6.193) следует, что при увеличении параметров a , ωRC переменное напряжение уменьшается вдоль эмиттерного перехода от базового контакта к центру, уменьшая там самым эффективную площадь эмиттера в динамическом режиме.

Входное сопротивление неоднородной линии БТ с учетом статического и динамического эффектов вытеснения записывается в виде

$$z_{\text{вх}}(\omega) = R \frac{\operatorname{cth} \left(j\omega RC \left(\frac{a+3}{3} \right) \right)^{1/2}}{\left(j\omega RC \left(\frac{a+3}{3} \right) \right)^{1/2}}. \quad (6.194)$$

При $|j\omega RC(a+3)/3| > 2$ выражение (6.194) приобретает вид

$$z'_{\text{вх}}(\omega) = \frac{R/(a+3)}{\left(j\omega RC \left(\frac{a+3}{3} \right) \right)^{1/2}} \quad (6.195)$$

или

$$z'_{\text{вх}} = \frac{R/(a+3)}{(j\omega/\omega_i)^{1/2}}, \quad \omega_i = \frac{3(a+3)}{RC} = \frac{3(a+3)\omega_r}{aB}.$$

Для малых значений параметра a в случае чисто динамического эффекта вытеснения

$$z'_{\text{вх}}|_{a \rightarrow 0} = \frac{R/3}{(j\omega/\omega_i)^{1/2}}, \quad \omega_i = \frac{9}{RC} = \frac{9\omega_r}{aB}. \quad (6.196)$$

Выражение (6.195) позволяет построить частотную зависимость базового сопротивления с учетом статического и динамического эффектов вытеснения в достаточно широком диапазоне частот:

$$z_B(j\omega) = \frac{R/(a+3)}{[1 + j\omega RC(a+3)/3]^{1/2}}. \quad (6.197)$$

В диапазоне $2 \ll |j\omega RC(a+3)/3| \ll 9$, воспользовавшись разложением в ряд знаменателя выражения (6.197) и принимая в расчет первые два члена разложения, базовое сопротивление можно представить как

$$z_B = \frac{R/(a+3)}{1 + j\omega \frac{R}{a+3} \frac{C}{6}} = \frac{R}{a+3} \left\| \left(\frac{j\omega C}{6} \right)^{-1} \right. \quad (6.198)$$

т. е. параллельным соединением низкочастотного дифференциального сопротивления $r_B = R/(a+3)$ и эффективной базовой емкости $C_B \approx C/6$. При этом дифференциальное сопротивление $R/(a+3)$ описывается ВАХ эффективного базового диода в модели БТ на большом сигнале (рис. 6.26) с

$$\text{характеристикой } I_B = \frac{3\varphi_T}{R} \left(\exp \frac{U_{ББ'}}{\varphi_T} - 1 \right).$$

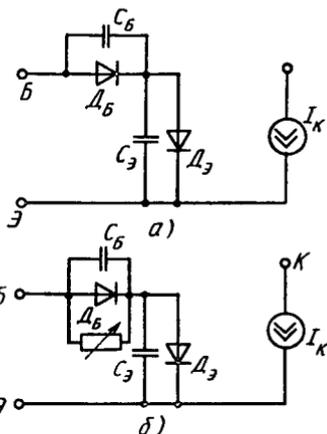


Рис. 6.26. Представление базовых сопротивлений в эквивалентных схемах БТ на большом сигнале

6.8. СИНТЕЗ ЭЛЕКТРИЧЕСКИХ МОДЕЛЕЙ БТ.

БИБЛИОТЕКА ИЕРАРХИЧЕСКИХ ЭЛЕКТРИЧЕСКИХ МОДЕЛЕЙ БТ

Необходимость учета микротопологии в электрической модели БТ. На рис. 6.27, а, б приведены топологические разрезы БТ для стандартных промышленных и перспективных субмикронных БИС с выделением элементов сосредоточенной схемы в статическом и динамическом режимах. Описанию диффузионных и барьерных емкостей на основе численного ФТ-моделирования будет уделено

значительное место в гл. 7. Здесь акцентируем внимание на важности и необходимости учета топологических особенностей БТ, в первую очередь омических сопротивлений областей, в общем случае нелинейных функций тока эмиттера, базы, коллектора. При уменьшении геометрических размеров БТ в горизонтальном и вер-

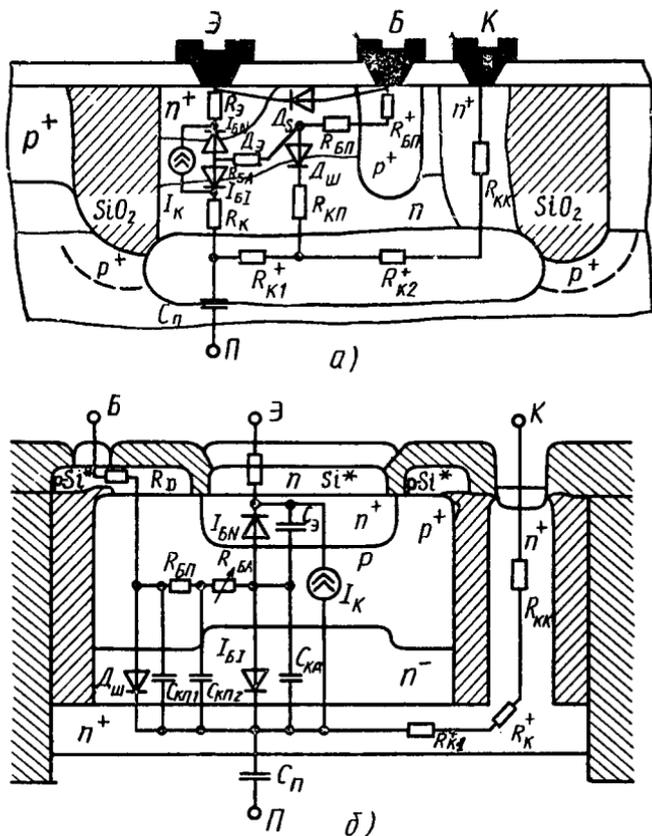


Рис. 6.27. Поперечное сечение и эквивалентная схема БТ, изготовленного по технологии «Изопланар 2» (а), и поликремниевой технологии сверхсамосовмещения (б)

тикальном направлении эти сопротивления начинают играть все более заметную роль, даже при достаточно сильном легировании этих областей. Усовершенствование технологических методов изготовления компонентов сверхскоростных БИС — субмикронных самосовмещенных БТ с поликремниевыми контактами — также приводит к росту сопротивления слоев эмиттера $R_{\text{Э}}$, поликремниевого базового слоя $\text{Si}^*(p^+) - R_{\text{Д}}$, последовательно включенного с сопротивлениями активной $R_{\text{БВ}}$ и пассивной базы $R_{\text{БП}}$. Причем в по-

ликремниевых субмикронных БТ часто $R_{\text{э}} \geq 10 \text{ Ом}$, $R_D \geq 50 \div 200 \text{ Ом}$. Диод $D_{\text{ш}}$, шунтированный соответствующей коллекторной емкостью пассивной части БТ, отражает способность в ряде электрических режимов насыщения шунтировать переход, перераспределять коллекторный и базовый токи и существенно влиять на электрические характеристики БТ. Наличие значительных базовых сопротивлений и шунтирующего диода $D_{\text{ш}}$ приводит к такому явлению: даже в отсутствие внешнего коллекторного тока имеется реальная возможность циркуляции в коллекторно-базовой цепи некоторого тока базы $I_{\text{б}}$ (по цепи $R_{\text{бп}} - D_{\text{ш}} - R_{\text{к}}$), когда диод $D_{\text{ш}}$ смещается в прямом направлении в большей степени, чем коллекторный диод $D_{\text{к}}$. В динамическом режиме при подаче на БТ импульсов с короткими фронтами с помощью диода $D_{\text{ш}}$ и емкости $C_{\text{кп}}$ моделируются эффекты накопления носителей заряда в пассивной части база — коллектор, искажение переходного процесса в коллекторно-базовой цепи БТ («улучшение» или затягивание фронтов на соответствующих входных перепадах напряжений).

Для получения высококачественных БТ, уменьшения коллекторных сопротивлений в технологии биполярных БИС наиболее часто применяют планарно-эпитаксиальные БТ, у которых достаточно высокоомное сопротивление эпитаксиального коллекторного слоя $R_{\text{к}}$ изменяется вследствие эффектов статического вытеснения тока и модуляции проводимости слоя. Поскольку при статическом вытеснении тока работает периферия БТ, это приводит к некоторому увеличению $R_{\text{к}}$, $R_{\text{к}} \geq R_{\text{к0}}$

$$R_{\text{к}} \approx R_{\text{к0}}(1 + a/3). \quad (6.199)$$

Однако эффекты высокого уровня инжекции в коллекторе компенсируют влияние статического эффекта вытеснения, что отражается в некотором уменьшении роста эффективного сопротивления $R_{\text{к}}$ с ростом тока коллектора больше критического $I_{\text{кр}}$. Несмотря на то что с помощью перспективных технологий БИС с изоляцией БТ канавками, заполненными двуокисью кремния или поликремния, удастся резко минимизировать емкость подложки $C_{\text{п}} < 10 \text{ фФ}$, ее влияние, особенно в области микротоков, весьма значительно. Распределенная емкость подложки моделируется сосредоточенным эквивалентом $C_{\text{п}}$, имеющим обычно постоянное значение.

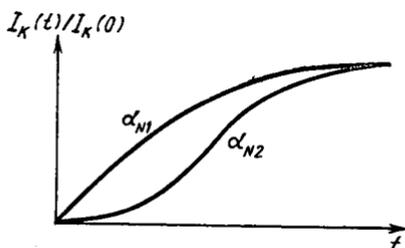
С помощью базового сопротивления БТ, зависящего от рабочего тока и входного фронта, моделируются эффекты статического и динамического вытеснения, а также эффекты высокого уровня инжекции в базе.

Учет фазовой задержки распространения сигнала в БТ. Дифференциальные уравнения первого порядка (6.113), (6.115), (6.116) электрической модели БТ, описывающие инерционный процесс изменения базового, коллекторного и эмиттерного токов, не учитывают дополнительной фазовой задержки распространения сигнала в БТ от эмиттера к коллектору:

$$\tau_{3N} \approx \tau_{\Sigma} + \tau_{\tau B} + \tau_i, \quad (6.200)$$

где τ_{Σ} определяется соотношениями (6.35), (6.38), (6.40), а под τ_i подразумевается вклад всех других областей и p - n -переходов БТ.

Для того чтобы строго учесть эффект фазовой задержки БТ, связанной с конечным временем переноса носителей заряда через БТ в переходной характеристике $I_K(t)$ (рис. 6.28), надо перейти от однополюсного представления коэффициента передачи тока



$$\alpha_{N1} = \frac{\alpha_N(0)}{1 + p\tau_{\tau N}} \quad (6.201)$$

Рис. 6.28. Описание переходного процесса тока коллектора БТ с помощью однополюсной и двухполюсной аппроксимации коэффициента передачи тока в схеме БТ с общей базой

к двухполюсному

$$\alpha_{N2}(p) = \frac{\alpha_N(0)}{(1 + p\tau_{3N})(1 + p\tau_{\tau N})}, \quad (6.202)$$

где p — оператор $p = d/dt$.

Учет фазовой задержки БТ в программе схемотехнического проектирования достаточно просто может быть осуществлен путем использования в односекционной сосредоточенной электрической модели БТ линии задержки с эффективными параметрами L и C (рис. 6.29). На рисунке показана часть полной эквивалентной схемы БТ, где формируется двухполюсное представление коэффициента передачи тока, когда токи эмиттера и коллектора модели сдвинуты с помощью задержки в LC -цепи, причем $\tau_{3N} = 2\sqrt{LC}$. Ток I_{Σ} в эмиттерном диоде модели БТ поступает на вход LC -цепи и, задержанный на время $\tau_{3N}(I_{\Sigma}^*)$, используется для описания генератора коллекторного тока в коллекторно-базовой цепи БТ. Для программ схемотехнического расчета БИС, не сильно лимитированных количеством узлов и пассивных двухполюсников, использование дополнительных моделей БТ с двухполюсными коэффициентами передачи тока с учетом эффектов высокого уровня инжекции может быть весьма эффективно. В случае ограниченных возможностей программ и машинного времени для снижения размерности могут быть применены другие алгоритмические методы учета собственной фазовой задержки БТ, особенно для расчета сверхскоростных переключений в области малых времен. В частности, проблема повышения точности модели БТ в области малых времен может быть решена без существенного увеличения затрат машинного времени, если в рассматриваемой передаточной модели БТ с управлением зарядом коллекторный ток описывать обыкновенным дифференциальным уравнением второго порядка:

$$\tau_{3N}\tau_{TN}\frac{d^2I_K}{dt^2} + (\tau_{3N} + \tau_{TN})\frac{dI_K}{dt} + I_K(t) = I_S \exp \frac{U_{\text{Э}}}{\varphi_T}. \quad (6.203)$$

Уравнению (6.203) соответствует переходная характеристика коллекторного тока (см. рис. 6.28) с фазовой задержкой. В этом уравнении τ_{3N} , τ_{TN} — постоянные времени, аппроксимирующие двухполюсную функцию (6.202). Уравнение (6.203) в программе схемотехнического расчета интегрируется численно с уравнениями, описывающими узловые потенциалы рассчитываемой БИС. Причем на каждом $n+1$ временном шаге интегрирования в соответствующем блоке программы «библиотечная модель БТ» вычисляется «задержанный» коллекторный ток I_K^{n+1} , зависящий от тока на двух других предыдущих шагах I_K^{n-1} , I_K^n временных шагов Δt^n , Δt^{n+1} и текущих значений переменных τ_{3N}^{n+1} , τ_{TN}^{n+1} , $U_{\text{Э}}^{n+1}$. При реализации модели в программе в данном случае требуются дополнительные затраты памяти для хранения I_K^n , I_K^{n-1} на двух предыдущих шагах и небольшие дополнительные затраты машинного времени для текущих расчетов с использованием блока программы «библиотечная модель БТ».

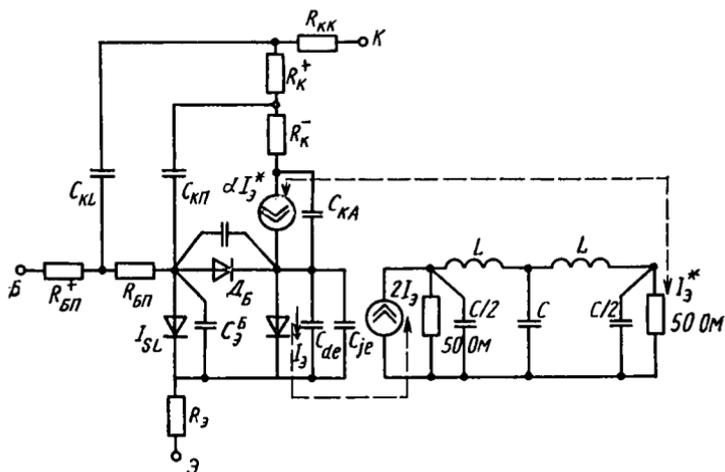


Рис. 6.29. Эквивалентная сосредоточенная схема БТ на большом сигнале

Однако необходимо отметить, что для схем, работающих от генератора напряжения в режиме большого сигнала, учет эффекта фазовой задержки практически искажен сильным влиянием высокоомной базовой цепи БТ. Использование эффективных распределенных электрических моделей БТ, в которых, например, активная базовая область представлена несколькими секциями, позволяет получать минимальную погрешность описания переходного процесса в БТ и с помощью стандартной формы представления то-

ков обыкновенными дифференциальными уравнениями первого порядка (6.113), (6.115), (6.116) с однополюсными коэффициентами передачи тока (6.201).

Эффективные распределенные электрические модели БТ. Пред-

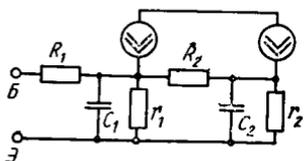


Рис. 6.30. Схема двухсекционной модели и распределенной эмиттерно-базовой области БТ

ставляет значительный интерес синтез эффективных распределенных моделей, состоящих из двух секций (в общем случае неоднородных), адекватно в некотором интервале фронтов (частот) входных сигналов описывающих переходные процессы в ненасыщенной схеме.

Для однородной распределенной малосигнальной модели БТ (см. рис. 6.24) входное сопротивление и полный коллекторный ток в операторной форме описы-

ваются выражениями

$$z_{\text{вх}}(p) = \frac{R}{(p\tau)^{1/2}} \frac{\text{ch}(p\tau)^{1/2}}{\text{sh}(p\tau)^{1/2}}, \quad I_K(p) = \frac{U_0}{r(p\tau)^{1/2}} \frac{\text{sh}(p\tau)^{1/2}}{\text{ch}(p\tau)^{1/2}}. \quad (6.204)$$

Воспользовавшись разложением в ряд гиперболических функций, получим в определенном временном интервале следующие выражения для $z_{\text{вх}}(p)$, $I_K(p)$:

$$z_{\text{вх}}(p) = \frac{1}{pC_{\text{Э}}} \frac{1 + \frac{p\tau}{2} + \frac{(p\tau)^2}{24} + \dots + \frac{(p\tau)^n}{n!} + \dots}{1 + \frac{p\tau}{6} + \frac{(p\tau)^2}{120} + \dots + \frac{(p\tau)^n}{(n+1)!} + \dots}, \quad (6.205)$$

$$k(p) = \frac{v_0}{r} \frac{1 + \frac{p\tau}{6} + \frac{(p\tau)^2}{120} + \dots + \frac{(p\tau)^n}{(n+1)!}}{1 + \frac{p\tau}{2} + \frac{(p\tau)^2}{24} + \frac{(p\tau)^3}{720} + \dots + \frac{(p\tau)^n}{n!}}, \quad (6.206)$$

причем погрешность выражений для $z_{\text{вх}}$ и I_K с ограниченным числом членов разложения в числителе и знаменателе будет определяться величиной τ и фронтом входного сигнала в области рассматриваемых малых времен.

Поставим задачу так: подобрать параметры элементов в двухсекционной эффективной модели БТ (рис. 6.30) в некотором временном интервале, чтобы она достаточно адекватно отражала в нем динамические характеристики распределенной модели БТ с бесконечно большим количеством секций.

Для двухсекционной модели БТ (рис. 6.30) напомним соотношение напряжений двух узлов U_1 , U_2 :

$$U_1(p) = v_0 \frac{1 + pR_2C_2}{(1 + pR_1C_1)(1 + pR_2C_2) + pR_1C_1}, \quad (6.207)$$

$$U_2(p) = v_0 \frac{1}{(1 + pR_1C_1)(1 + pR_2C_2) + pR_1C_2}. \quad (6.208)$$

Использование выражений (6.207), (6.208) для определения базового и эмиттерного токов в соответствующих ветвях двухсекционной модели (рис. 6.30) при коэффициенте передачи тока $\alpha_N \approx 1$ дает возможность определить для данной модели входное сопротивление $z_{вх}^{(2)}(p)$ и полный коллекторный ток $I_K^{(2)}(p)$:

$$z_{вх}^{(2)}(p) = \frac{1 + p(\tau_1 + \tau_2 + \gamma_2 \tau_2) + p^2 \tau_1 \tau_2}{p C_{\Sigma} (1 + p \tau_1 \tau_2) / (\tau_1 + \gamma_2 \tau_2)}, \quad (6.209)$$

$$I_K^{(2)}(p) = \frac{v_0}{r_{\Sigma}} \frac{1 + p \gamma_1 \tau_2}{1 + p(\tau_1 + \tau_2 + \gamma_2 \tau_2) + p^2 \tau_1 \tau_2}, \quad (6.210)$$

где $\tau_1 = C_1 R_1$, $\tau_2 = C_2 R_2$, $\gamma_1 = r_2 / (r_1 + r_2)$, $C_{\Sigma} = C_1 + C_2$, $\gamma_2 = R_1 / R_2$, $r_{\Sigma} = r_1 r_2 / (r_1 + r_2) = \gamma_1 r_1$.

Ограничиваясь двумя членами разложения в числителе (знаменателе) и знаменателе (числителе) выражений (6.205), (6.206), приравнивая соответствующие члены в формулах (6.209), (6.210) и (6.205), (6.206), получаем систему уравнений:

$$\begin{aligned} \tau_1 \tau_2 / (\tau_1 + \gamma_2 \tau_2) &= \tau / 6, & \tau_1 + \tau_2 + \gamma_2 \tau_2 &= \tau / 2, \\ \tau_1 \tau_2 &= \tau^2 / 24, & \gamma_1 \tau_2 &= \tau / 6, & \tau_1 + \tau_2 + \gamma_2 \tau_2 &= \tau / 2. \end{aligned} \quad (6.211)$$

Решением системы (6.211) являются равенства $\gamma_1 = 2/3$, $\gamma_2 = 1/3$, $\tau_1 = 2/3 \tau_2$ и следующие соотношения для эффективных параметров двухсекционной модели БТ:

$$r_2 = 2r_1, \quad R_1 = R_2/3, \quad C_1 = 2C_2 \quad (6.212)$$

Таким образом, в синтезированной эффективной двухсекционной модели БТ (рис. 6.30) полный коллекторный ток, согласно (6.210), (6.212), описывается в оперативном виде:

$$I_K(p) = \frac{v_0}{r_{\Sigma}} \frac{1 + p \tau_3}{(1 + p \tau_1)(1 + p \tau_2)}, \quad (6.213)$$

где $\tau_1 = \tau/0.2$, $\tau_2 = \tau/0.8$, $\tau_3 = \tau/0.5$, а временная зависимость коллекторного тока — выражением

$$\begin{aligned} I_K(t) &= \frac{v_0}{r_{\Sigma}} \left(1 - \frac{\tau_1 - \tau_3}{\tau_1 - \tau_2} e^{-\frac{t}{\tau_1}} + \frac{\tau_2 - \tau_3}{\tau_1 - \tau_2} e^{-\frac{t}{\tau_2}} \right) = \\ &= \frac{v_0}{r_{\Sigma}} \left(1 - 0.8e^{-\frac{0.2t}{\tau}} - 0.2e^{-\frac{0.8t}{\tau}} \right), \end{aligned} \quad (6.214)$$

которое может быть использовано в простейших аналитических расчетах переходных процессов в БТ.

Для трехсекционной распределенной модели БТ (рис. 6.31), аналогично рассмотренной методике составления уравнений, получим другие решения: $R_2 = 3R_1$, $R_3 = 1.8R_1$, $r_2 = 2r_1$, $r_3 = 2.3r_1$; $C_2 = 0.5C_1$, $C_3 = 0.22C_1$.

Эффективные модели БТ для большого сигнала. Описание емкостей p - n -переходов. Эффективные электрические модели внутреннего БТ для большого сигнала изображены на рис. 6.32. Для адекватного описания полной модели БТ к внутренним точкам B' , K' необходимо добавить двухполосники — резисторы, емкости диода, отражающие электрофизические процессы в пассивных областях БТ, а также влияние микротопологии на выходные статические и динамические характеристики БТ.

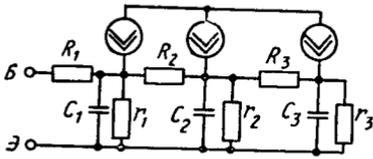


Рис. 6.31. Схема трехсекционной модели распределенной эмиттерно-базовой области БТ

фузионных емкостей эмиттера C_{de} через зависимость от рабочего тока параметра $\tau_{TN}(I_K)$ в полных емкостях $C_{Э}$, состоящих также из барьерных компонент $C_{Э1} = C_{dei} + C_{jei}$, $C_{dei} = \tau_{TN}(I) \frac{1}{\varphi_T} I_{Ki}$, позволяет значительно улучшить адекватность моделирования переходных процессов в области малых времен и высоких плотностей тока.

Включение эффектов высокого уровня инжекции в описание токов генераторов тока I_{Ki} , а также диффузионных емкостей эмиттера C_{de} через зависимость от рабочего тока параметра $\tau_{TN}(I_K)$ в полных емкостях $C_{Э}$, состоящих также из барьерных компонент $C_{Э1} = C_{dei} + C_{jei}$, $C_{dei} = \tau_{TN}(I) \frac{1}{\varphi_T} I_{Ki}$, позволяет значительно улучшить адекватность моделирования переходных процессов в области малых времен и высоких плотностей тока.

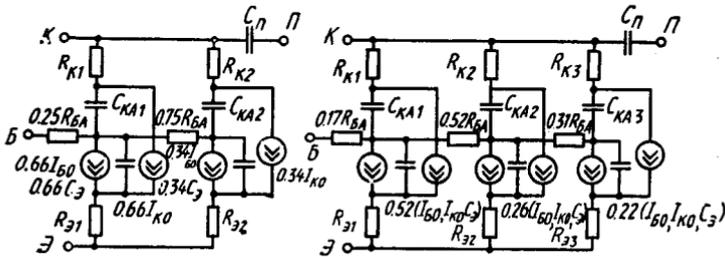


Рис. 6.32. Схемы двух- и трехсекционных распределенных моделей БТ

Зависимость барьерных емкостей переходов от приложенных напряжений может быть описана тремя типами выражений:

$$C_{je1} = C_0 / (1 - U/\varphi_0)^n = \frac{C_0}{v^n}, \quad C_{je1}(C_0, \varphi_0, n), \quad v = U/\varphi_0, \quad (6.215)$$

$$C_{je2} = \frac{C_0}{(v^2 + b)^{n/2}} \left(1 + \frac{n}{1-n} \frac{b}{v^2 + b} \right), \quad C_{je2}(C_0, \varphi_0, n, b), \quad (6.216)$$

$$C_{je3} = \frac{C_0}{(v^2 + b)^{1/2}} \left\{ \frac{v + (v^2 + b)^{1/2}}{2} \right\}^{1-n}, \quad C_{je3}(C_0, \varphi_0, n, b). \quad (6.217)$$

Обычно в большинстве программ схемотехнического проектирования зависимость барьерной емкости от приложенного напря-

жения описывается выражением (6.215) (рис. 6.33). Однако использование выражения (6.215) в библиотеке моделей БТ для расчета быстродействующих схем, работающих при средних и больших токах (0.1—10 мА), сопряжено с серьезным усложнением расчетов при $U \sim \varphi_0$. Когда напряжение на переходе приближается к φ_0 , значения емкости устремляются к бесконечности, поэтому обычно значения φ_0 увеличивают, стремясь сделать их превышающими наибольшее ожидаемое напряжение на переходе. Формула (6.215) просто выводится в предположении, что φ_0 — контактная разность потенциалов перехода (равная 0.6—0.8 В для кремниевых БТ), n — показатель плавности p - n -перехода: для резкого перехода $n=1/2$, для плавного перехода $n=1/3$. Однако даже при удачном выборе C_0 и n и завышенном значении $\varphi_0 \approx 1 \div 1.5$ В значительно ухудшается точность вычисления барьерной емкости при прямых смещениях и теряется адекватность моделирования.

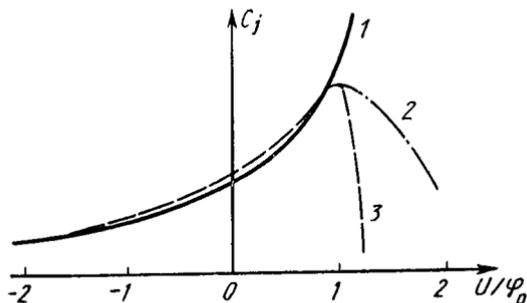


Рис. 6.33. Аппроксимации вольт-фарадной зависимости барьерной емкости перехода БТ: 1 — по формуле (6.215); 2 — по формуле (6.216); 3 — по формуле (6.217)

Модифицированные выражения для емкости (6.216), (6.217) позволяют улучшить описание реального БТ для приложенных смещений, равных и превышающих φ_0 (рис. 6.33). В большинстве случаев параметры моделей емкостей (C_0 , φ_0 , n , b) — это параметры аппроксимации, определяемые на основе натурального или машинного эксперимента. Например, значения подстроечного параметра b лежат в диапазоне 10^{-3} — $5 \cdot 10^{-1}$. Для обратных и небольших прямых напряжений формулы для барьерных емкостей (6.215)—(6.217) дают слабо различающиеся результаты (рис. 6.33), но при прямом напряжении $U = \varphi_0$ значения емкостей, вычисленных по формулам (6.216), (6.217), достигают экстремума:

$$C_{j \max} \approx \frac{C_0}{1-n} b^{n/2} \approx \frac{C_0}{b^{1/2}} \left(\frac{b^{1/2}}{2} \right)^{1/n}.$$

Однако при напряжениях $U > \varphi_0$ барьерная емкость, согласно данным формулам, убывает по различным законам, более быстро в (6.217) (рис. 6.33).

Настраиваемые упрощенные двухсекционные модели БТ. Параметры разбиения БТ на эффективные секции в рассмотренных примерах выбирались для самого общего случая исходя из того, чтобы синтезируемая эквивалентная схема с минимальной погрешностью описывала переходную характеристику БТ в области ма-

лых времен. Однако необходимо заметить, что в принципе очень трудно, а практически невозможно подобрать параметры эквивалентной схемы таким образом, чтобы они были оптимальными для всех режимов работы БТ, например в диапазоне $0 \leq p\tau \leq \infty$. Использование многосекционных ($n > 10$) моделей БТ, несмотря на достижение приемлемой погрешности расчетов, не может служить критерием целесообразности вследствие огромных затрат машинного времени при расчетах переходных процессов. Синтез компактных малосекционных ($n = 2 \div 3$) электрических моделей значительно упрощается, если удастся выделить специфику переключения в ограниченном диапазоне малых времен или рабочих токов.

Например, существенно упрощается описание процесса выключения быстродействующих БТ, где базовое сопротивление становится сравнительно большим. В двухсекционной модели с разбиением активной области БТ на две настраиваемые области суммарный коллекторный ток равен сумме токов двух секций:

$$\begin{aligned} I_K &= I_{K1} + I_{K2}, \\ I_{K1} &= z_0 h l I_s \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right), \\ I_{K2} &= (1 - z_0) h l I_s (\exp U_{\text{Э}} / \varphi_T - 1), \end{aligned} \quad (6.218)$$

где z_0 — коэффициент, равный отношению длин секции; h, l — ширина и длина эмиттера.

Активные базовые сопротивления двух секций представляют с помощью двух настраиваемых параметров a_1 и a_2 :

$$R_{\text{БА}1} = a_1 R, \quad R_{\text{БА}2} = a_2 R, \quad (6.219)$$

где

$$R = \begin{cases} R_{\text{СБА}}/hl, & k=1, \\ R_{\text{СБА}}/4hl, & k=2, \\ R_{\text{СБА}}/16hl, & k=4. \end{cases}$$

Параметры z_0, a_1, a_2 выбираются из условия, чтобы двухсекционная эквивалентная схема по возможности точнее описывала переходную характеристику в режиме выключения БТ. Отметим, что в процессе выключения распределенное сопротивление базы можно описать одним эквивалентным сосредоточенным сопротивлением базы $R_{\text{БА}}^*$, которое, однако, должно зависеть от тока и времени. В общем виде сопротивление $R_{\text{БА}}^*$ описывается выражением

$$R_{\text{БА}}^* \approx R (1 - I_K(t) / I_K(0)), \quad (6.220)$$

справедливым для достаточно большого переключающего базового тока $|I_B| \gg \varphi_T / R$, когда неосновные носители заряда первыми удаляются из эмиттерного p - n -перехода, т. е. в режиме резкого

перемещения зарядового фронта в глубь активной базы от эмитера. На рис. 6.34 представлен график переходного процесса тока коллектора, рассчитанного для различных моделей БТ с учетом неравенства $|I_B| \gg \varphi_T/R$ во время приложения импульса запирающего (отрицательного) напряжения, когда базовое сопротивление R_{BA}^* (6.220) сильно изменяется. Кривая 1 приведена для модели БТ с использованием R_{BA}^* (6.220), кривая 3 — для эквивалентной схемы БТ, содержащей одно постоянное среднее сопротивление \bar{R}_{BA} . В последнем случае наблюдается линейное уменьшение тока. Если принять $R_{BA} = R/4$, то кривые 1 и 3 совпадут в точке $I_K = I_K/2$, однако для остальных времен модель с одним сосредоточенным (средним) сопротивлением \bar{R}_{BA} дает неудовлетворительные результаты. Значительно лучшие результаты получаются для настраиваемой двухсекционной модели БТ. При $|I_B| \gg \varphi_T/R$ разрядка диффузионной емкости C_{de2} , отдаленной от базового контакта секции, начинается только после полной разрядки емкости C_{de1} , ближней к базовому контакту секции. Следовательно, результирующая кривая 2 переходного процесса будет состоять из двух прямых (рис. 6.34). После определения точки излома z_0 и точек z_1, z_2 , соответствующих пересечению кривых 2, 3, настраиваемые параметры a_1, a_2 в формуле (6.129) могут быть рассчитаны таким образом:

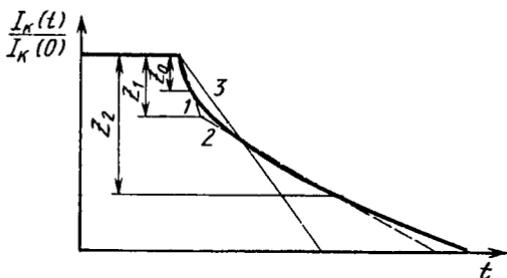


Рис. 6.34. График переходного процесса тока коллектора и его аппроксимации

Вывод: Значительно лучшие результаты получаются для настраиваемой двухсекционной модели БТ. При $|I_B| \gg \varphi_T/R$ разрядка диффузионной емкости C_{de2} , отдаленной от базового контакта секции, начинается только после полной разрядки емкости C_{de1} , ближней к базовому контакту секции. Следовательно, результирующая кривая 2 переходного процесса будет состоять из двух прямых (рис. 6.34). После определения точки излома z_0 и точек z_1, z_2 , соответствующих пересечению кривых 2, 3, настраиваемые параметры a_1, a_2 в формуле (6.129) могут быть рассчитаны таким образом:

$$a_1 = \frac{z_1}{2}, \quad a_2 = \frac{z_2}{2} \frac{z_2 - z_1}{z_2 - z_0}. \quad (6.221)$$

Выбирая z_0, z_1, z_2 в соответствии с рис. 6.34, получим типовые значения $a_1 = 0.15, a_2 = 0.5, z_0 = 0.4$, которые используют для нормирования диффузионных емкостей, базовых сопротивлений и коллекторных токов двухсекционных эквивалентных схем БТ [см. выражения (6.218) — (6.220)]. В большинстве случаев настраиваемые параметры a_1, a_2, z_0 не зависят от других параметров эквивалентной БТ. Сравнение результатов расчета процесса выключения БТ, полученных на основе «настраиваемой» двухсекционной модели БТ и n -секционной модели ($n = 10 \div 20$), позволяет сделать вывод о достаточно высокой адекватности «настраиваемых» эффективных двухсекционных моделей при значительном выигрыше в вычислительных затратах, увеличивающемся с ростом размерности рассчитываемых на ЭВМ БИС.

Библиотека иерархических адаптивных электрических моделей БТ. Анализ погрешностей аппроксимаций двух- и трехсекционными эффективными моделями одновременно входного сопротивления и тока коллектора идеальной распределенной модели показывает возможность (с погрешностью 10—15%) использования двухсекционных моделей БТ во временном интервале при $1 \leq \rho\tau \leq 16$ и трехсекционных моделей при $1 \leq \rho\tau \leq 24$.

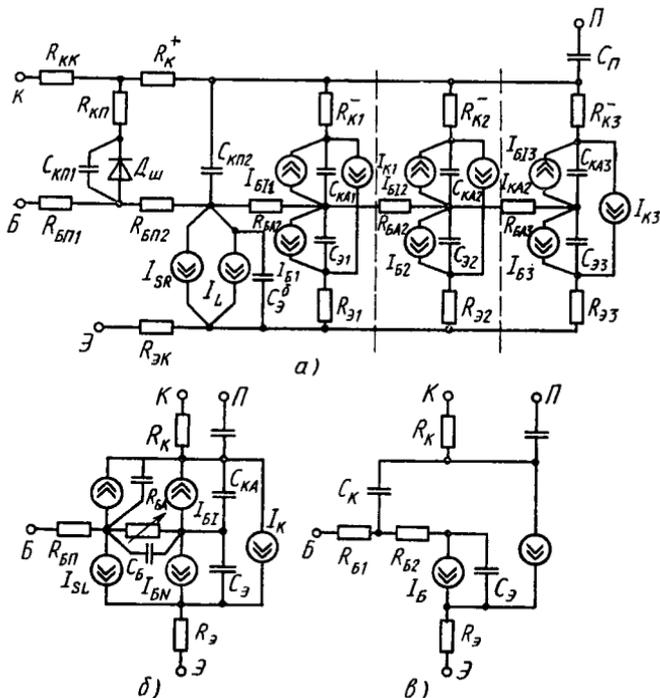


Рис. 6.35. Набор иерархических адаптивных электрических моделей БТ

В схеме эффективной распределенной модели БТ для большого сигнала (рис. 6.35, а) пунктиром выделена область активной базы БТ, моделируемой двумя или тремя неоднородными секциями с зависящими от режимного тока базовыми сопротивлениями секций.

Генераторы базовых составляющих токов для простоты промоделированы одним эффективным генератором базового тока I_{B1} . Нелинейные емкости эмиттера $C_{Э1}$, состоящие из суммы барьерной и диффузионной емкостей, учитывают эффективное разбиение БТ на секции. Барьерные емкости каждой из секций пропорциональны соответствующим им площадям, а диффузионные емкости — токам коллектора: $C_{dei} \approx \tau_{TN} I_{Ki} / \Phi T$. Эффекты высоких плотностей тока одномерных транзисторов каждой из секций отраже-

ны в рассмотренных ранее нелинейных зависимостях I_{ki} , C_{dei} , R_{Bi} , R_{ki} .

Организация библиотеки адаптивных моделей на две, три, пять, семь неоднородных секций в процессе счета позволяет оперировать моделями требуемой точности, причем контролирование очередности задействования каждой из моделей, несмотря на некоторые затраты машинного времени, дает в конечном итоге экономии за счет обращения к моделям невысокой сложности (две секции) на большинстве участков переходного процесса.

В зависимости от задач электрического моделирования БИС из иерархического ряда моделей БТ требуется не одна модель БТ высокой точности (рис. 6.35, а), а несколько адаптивных, ориентированных на применение в том или ином электрическом и температурном режимах, диапазон входных фронтов и др. Для обеспечения компромисса точность — сложность моделирования фрагментов БИС формализованный синтез модели ($k-1$)-го уровня из более сложной и точной модели k -го уровня может быть осуществлен с помощью стандартных программ оптимизации. При этом для адаптивных моделей низких уровней сложности параметры двухполюсников в общем случае, возможно, будут не всегда «физичными». Для сохранения «физичности» параметров упрощенных моделей иерархического ряда (рис. 6.35, а), например, в насыщенном (рис. 6.35, б) и активном нормальном (рис. 6.35, в) режимах могут быть использованы исключение некоторых элементов модели БТ (если влияние соответствующих цепей в статике или динамике пренебрежимо мало), группировка элементов (емкостей, сопротивлений, генераторов тока областей и переходов БТ) в общий эффективный элемент, а также значительное упрощение аналитических зависимостей модели I_{kk} , C_{de} , z_B . Данный подход, несмотря на «инженерный» характер, легко формализуется на ЭВМ, позволяет синтезировать модели БТ из стандартных двухполюсников и дает удовлетворительную точность расчета (5—20%) статических и динамических параметров элементов и фрагментов БИС.

В параметры двухполюсников модели БТ включены температурные зависимости токов насыщения, сопротивлений базы и диффузных емкостей согласно предыдущему ФТ-расчету в температурном диапазоне. Иерархический принцип построения универсальной k -й модели БТ включает априорную адаптацию модели к режимным (рабочие токи и напряжения), температурным и сигнальным (на входе) изменениям в рассчитываемой схеме. Процедура адаптации иерархической модели БТ к совокупности режимных и температурных изменений основана на согласовании режимных и температурных факторов U_i , I_i , $t_{\phi i}$, T_i с текущими значениями переменных U , I , t_{ϕ} , T , а также на выработке требований точности и вычислительных затрат каждой электрической модели БТ из иерархического набора.

ОБЪЕДИНЕННОЕ ФИЗИКО-ТЕХНОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ. ИДЕНТИФИКАЦИЯ И КОРРЕЛЯЦИЯ ПАРАМЕТРОВ ЭЛЕМЕНТОВ БЫСТРОДЕЙСТВУЮЩИХ БИС

7.1. МЕТОДОЛОГИЯ МАШИННОГО РАСЧЕТА И ИДЕНТИФИКАЦИИ ХАРАКТЕРИСТИК И ПАРАМЕТРОВ ИНТЕГРАЛЬНЫХ БТ С ПОМОЩЬЮ ПРОГРАММ ТЕХНОЛОГИЧЕСКОГО И ФИЗИКО-ТОПОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ

Традиционные экспериментальные методы идентификации электрических параметров, несмотря на их важность, актуальность и широкое использование в процессе разработок ИМС и БИС, в настоящее время нельзя считать единственно достоверными и эффективными вследствие значительных погрешностей измерений, высокой трудоемкости, сложности и длительности измерений основных параметров.

Одним из главных результатов совместного технологического (Т) и ФТ-моделирования является расчет электрических характеристик БТ и идентификация электрических параметров его электрической модели. После данной процедуры расчет БИС может быть продолжен на элементном и фрагментном уровнях. При этом задача идентификации электрической модели БТ должна быть поставлена так, чтобы данная модель достаточно адекватно описывала всю совокупность статических и динамических характеристик БТ. В частном случае она должна хорошо описывать совокупность электрических характеристик в ограниченных режимах использования БТ в БИС, например в активном нормальном режиме.

Общим фундаментальным подходом к идентификации параметров электрической модели БТ является метод параметрической оптимизации — настройки, например по среднеквадратическому отклонению, модельных характеристик двух уровней — ФТ и Э — с целью определения искомого параметра. Этот метод, несмотря на универсальность и хорошую формализуемость, имеет существенные недостатки: 1) необходимость предварительной оценки значений настраиваемых параметров; 2) значительные трудности определения функциональной зависимости параметров от электрического режима; 3) неоднозначность результатов настройки, т. е. существование нескольких локальных экстремумов (особенно при большом количестве настраиваемых параметров и «гладкости целевой функции»), а также ряд других недостатков чисто вычисли-

тельного плана. Третий недостаток является самым распространенным и трудно устранимым, что достигается такими специфическими приемами, как выбор начальных приближений (детерминированно или случайно), настройка совокупности зависимостей и др. Наряду с универсальными методами идентификации параметров электрической модели с помощью математического аппарата и параметрической оптимизации, широко используемой в данной работе, будут описаны специальные методы идентификации электрической модели и ее параметров, опирающиеся на электрофизическую природу каждого из параметров.

Методы машинной идентификации только на основании ФТ-расчетов характеристик элементов БИС с приблизительным заданием профиля примесей бесперспективны.

Кратко опишем этапы и последовательность применения и взаимодействия программ при машинном расчете, а также идентификации характеристик и параметров БТ.

1. Необходимым условием процесса моделирования является наличие и согласование с разработчиками БИС полного технологического маршрута изготовления БТ (библиотечного набора БТ) и топологического чертежа. При согласовании маршрута и чертежа должны быть выяснены сведения о погрешностях задания технологических параметров, а также литографических погрешностях топологической прорисовки областей БТ при различных термооперациях. Для расчета параметров пассивных компонентов: резисторов, емкостей, формируемых в определенных слоях БТ, — также требуется информация о маршруте, топологии и погрешностях их параметров.

2. С помощью одномерной технологической программы ТЕХИС 1 рассчитывают профиль распределения примеси в наиболее критичном сечении БТ — под дном эмиттера. Для указанных при согласовании с технологами допустимых отклонений технологических параметров (дозы легирования, температуры, времени термоопераций) и реальных экспериментальных данных по профилю БТ (глубины залегания переходов, поверхностного сопротивления слоев и др.) имеется принципиально важная возможность скорректировать данные моделирования и технологического варианта изготовления БТ, варьируя технологические (и даже некоторые подстроечные параметры) модели для получения удовлетворительного согласования между расчетом и экспериментом.

3. Профиль базовой примеси, рассчитанной по одномерной программе ТЕХИС 1, до начала введения эмиттерной примеси представляют в двумерную технологическую программу ТЕХИС 2, далее рассчитывают в двумерном приближении распределение эмиттерной и базовой примесей вблизи эмиттера.

4. Одномерный профиль (при $y=0$) эмиттерной и базовой примесей, рассчитанный по программе ТЕХИС 2, и профиль примеси скрытого слоя переносят по точкам в одномерную ФТ-программу ТРАН 1.

Производят пробный расчет одной-двух точек ВАХ для низких уровней инжекции с ориентировочной площадью эмиттера (взятой, например, из топологического чертежа), далее, если необходимо, проверяют согласование с экспериментальной ВАХ (например, $I_K(U_a)$, $U_K=0$).

5. Двумерный профиль эмиттерной и базовой примесей из программы ТЕХИС 2 и одномерный профиль из программы ТЕХИС 1 подставляют в программу ТРАН 2 для моделирования активного нормального режима БТ. Производят расчет ВАХ и параметров БТ по программе ТРАН 2 и сравнение с результатами расчета по программе ТРАН 1. Отметим, что сетка разбиения по оси x (в глубину БТ) для программ ТРАН 1 и ТРАН 2 совпадает.

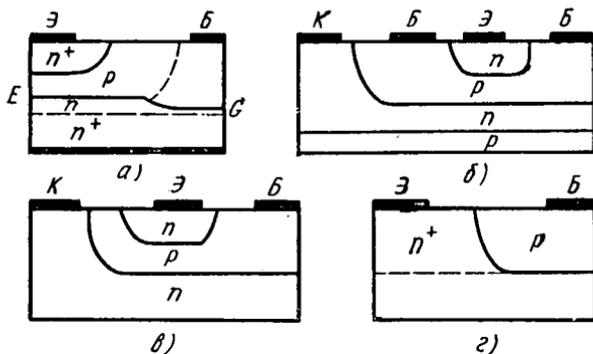


Рис. 7.1. Типовые геометрии интегральных БТ, используемые в ФТ-расчетах

6. По совпадению коллекторных токов, рассчитанных по программам ТРАН 1 и ТРАН 2 при низких уровнях инжекции, определяют эффективную площадь эмиттера, требуемую для надежных расчетов по программе ТРАН 1.

7. ВАХ и параметры для активного инверсного режима работы БТ с непланарным расположением контактов рассчитывают в рамках программ ТРАН 1 и ТРАН 2, причем в программе ТРАН 2 предусмотрена возможность интегрирования заряда и скорости рекомбинации по всему БТ и по специально выбранной области, а также прямого вычисления токов в заданных областях.

На рис. 7.1, а—г приведены типовые геометрии интегральных БТ, используемые в ФТ-расчетах.

7.2. ОБЪЕДИНЕНИЕ ПРОГРАММ ТЕХНОЛОГИЧЕСКОГО И ФИЗИКО-ТОПОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ ИНТЕГРАЛЬНЫХ БТ В ЕДИНЫЙ КОМПЛЕКС ФИЗИКО-ТЕХНОЛОГИЧЕСКОГО РАСЧЕТА

Для практического использования программ двумерного технологического моделирования (ТЕХИС 2) и физико-топологического моделирования (ТРАН 2) БТ необходимо объединить их в еди-

ный комплекс, позволяющий производить расчеты характеристик БТ с минимальными затратами ручного труда. При этом основную сложность представляют организация обмена между отдельными программами, формирование профиля примеси в БТ из различных фрагментов, рассчитанных по Т-программе, и автоматический выбор сетки пространственного разбиения (СПР) для ФТ-программы ТРАН 2.

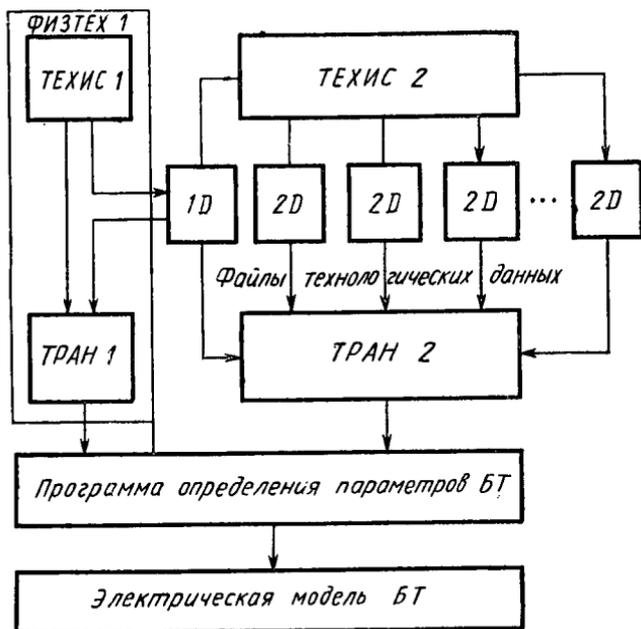


Рис. 7.2. Структурная схема единого комплекса физико-технологического расчета

Объединение программ одномерного технологического (ТЕХИС 1) и физико-топологического (ТРАН 1, ТРАН 1,5) моделирования с автоматизированной генерацией сетки разбиения рассчитываемых приборных структур реализовано в одномерном программном комплексе физико-технологического моделирования — ФИЗТЕХ 1. Объединение программ двумерного моделирования ТЕХИС 2, ТРАН 2 и программного комплекса ФИЗТЕХ 2 представлено структурной схемой на рис. 7.2.

Одномерные профили примеси, рассчитываемые технологической программой ТЕХИС 1 комплекса ФИЗТЕХ 1, передаются для ФТ-расчета либо записываются в специальный файл данных и впоследствии могут быть использованы в программах ТЕХИС 2 (в качестве начального распределения) и ТРАН 2 (в качестве одного из нескольких фрагментов). Двумерные профили примеси

в различных областях БТ, рассчитываемые программой ТЕХИС 2, также записываются в файлы данных.

Программа ТРАН 2 производит считывание фрагментов из указанных пользователем файлов и их объединение на единой СПР с большим количеством узлов, генерирует оптимальную СПР и производит двумерный ФТ-расчет БТ. Результаты ФТ-расчета передаются программе определения параметров электрической модели БТ по среднеквадратическому отклонению электрических и ФТ-характеристик.

Опишем последовательность расчета при моделировании БТ (рис. 7.2) в комплексе ФИЗТЕХ 2. Программой ТЕХИС 2 рассчитывают следующие фрагменты: Ф1 — профиль акцепторной и донорной примесей вблизи эмиттера; Ф2 — профиль скрытого n^+ -слоя; Ф3 — профиль акцепторной примеси вблизи боковой части перехода база — эмиттер; Ф4 — акцепторный профиль подлегирования контакта базы (p^+); Ф5 — донорный профиль подлегирования контакта коллектора (n^+). Все эти двумерные профили записываются в соответствующие файлы данных. Далее производится формирование профиля БТ из всех частных профилей, записанных в файлы данных.

В комплексе ФИЗТЕХ 2 использован специальный входной язык, который позволяет пользователю считывать эти профили и размещать их в нужных местах БТ. Пользователь задает номера файлов, в которых находятся те или иные фрагменты полного профиля БТ, а также координаты центров фрагментов по оси x . Файлы данных располагают в определенном порядке: сначала файлы с профилем эмиттера и файл с профилем примеси скрытого слоя, затем файлы, в которых хранятся профили подлегирующих контакты примесей. Пользователь указывает также координаты контактов.

Один из наиболее сложных вопросов, который необходимо решить при объединении двумерных программ технологического и физико-топологического моделирования, — автоматизированный выбор СПР для программного комплекса физико-технологического моделирования.

Из опыта практических ФТ-расчетов известно, что шаг сетки необходимо уменьшать в области объемного заряда p - n -переходов, в то время как в квазинейтральных областях (КО) шаг может быть значительно увеличен. Можно предположить, что для уменьшения ошибки аппроксимации в областях с более нелинейным ходом потенциала нужно делать более мелкий шаг СПР. Мерой нелинейности потенциала является его вторая производная по координате. Таким образом, для нахождения второй производной потенциала следует решать задачу ФТ-моделирования. В программе ТЕХИС 2 предложен и реализован метод выбора шагов на основе предварительно вычисленного распределения потенциала $V(x, y)$.

После считывания всех фрагментов профиля моделируемого БТ фрагменты интерполируют на вспомогательной равномерной

сетке с большим количеством узлов, покрывающей всю область моделирования. На этой сетке решают уравнение Пуассона в приближении квазиравновесия:

$$\Delta V(x, y) = -\left(\frac{q}{\epsilon\epsilon_0}\right) \left(n_{ie} \exp(-V(x, y)) - n_{ie} \exp(V(x, y)) + C(x, y) \right). \quad (7.1)$$

Решение ведут строчным итерационным методом Зейделя, для которого хорошая сходимость достигается уже за 5—7 итераций. В каждом узле вспомогательной сетки вычисляют значения вторых производных $\frac{\partial^2 V(x, y)}{\partial x^2}$ и $\frac{\partial^2 V(x, y)}{\partial y^2}$. На каждой горизонтальной и вертикальной линиях сетки находят средние значения производных (k, l — номера узлов вспомогательной сетки соответственно по осям x, y):

$$A(x_k) = \sum_l \left(\frac{\partial^2 V}{\partial x^2} \right)_{kl},$$

$$B(y_l) = \sum_k \left(\frac{\partial^2 V}{\partial y^2} \right)_{kl}.$$

Наконец, находят набор оптимальных шагов искомой СПР соответственно по осям x и y ($i=2, m_x; j=2, m_y$), обратно пропорциональных средним вторым производным потенциала по x, y .

Целесообразно ввести ограничение на максимальный разброс шагов сетки $h_{\max}/h_{\min}=100$, для того чтобы в квазинейтральных областях шаг не был слишком велик. Необходимо также добавить несколько узлов в сетку по оси y : по узлу вблизи границ $y=0, y=y_{\max}$ для задания граничных условий и по несколько узлов в областях, где начинаются или кончаются металлические контакты. На рис. 7.3 приведен пример автоматизированной генерации СПР для типового БТ БИС.

После вычисления СПР, а также распределения примесей в узлах СПР информация о БТ передается для расчета программе физико-топологического моделирования, одномерный профиль активной области транзистора — программе ФИЗТЕХ 1. Это связано с тем, что для определения всех параметров электрической модели БТ требуются результаты как двумерного, так и одномерного расчета.

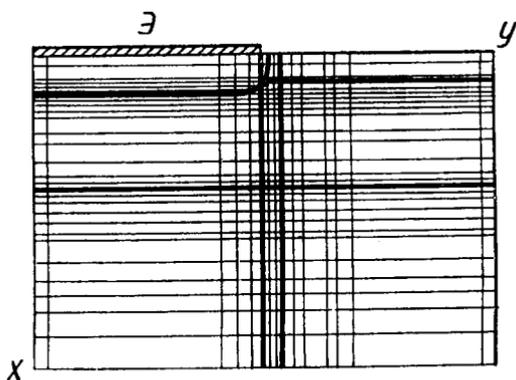


Рис. 7.3. Пример автоматизированной генерации сетки пространственного разбиения для типового БТ вблизи эмиттера (масштаб $5 \cdot 10^4 : 1$)

7.3. ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ ЭЛЕКТРИЧЕСКОЙ МОДЕЛИ БТ В АКТИВНОМ НОРМАЛЬНОМ РЕЖИМЕ С ПОМОЩЬЮ ПРОГРАММ ФТ-МОДЕЛИРОВАНИЯ

В рамках ФТ-моделирования БТ в одномерном, двумерном и квазидвумерном приближениях на ЭВМ рассчитывают статические (ВАХ, сопротивления) и динамические (емкости, времена задержки) характеристики и параметры, на основе которых в принципе может быть синтезирована электрическая модель БТ произвольной конфигурации и сложности, не привязанная к общепринятым стандартам. Однако ввиду того, что в инженерной практике разработки БИС наибольшее распространение получили электрические модели из набора электрических двухполюсников — генераторов тока, емкостей, сопротивлений с соответствующим описанием параметров, — в качестве базисных электрических моделей могут быть использованы модели типа рассмотренных в гл. 6. Одной из наиболее привлекательных сторон ФТ-моделирования является принципиальная возможность определения всех параметров электрической модели БТ на основе полученных расчетных ФТ-зависимостей. Формально это может быть осуществлено с помощью аппарата настройки. Однако проблема многоэкстремальности настраиваемых параметров значительно усложняет их корректную идентификацию. Рассматриваемые ниже методы идентификации параметров электрической модели БТ, опирающиеся на физический смысл параметра, позволяют более качественно и эффективно проводить процесс идентификации статических и динамических параметров БТ в рамках численного одномерного, двумерного и квазидвумерного моделирования.

Описание методик определения характеристик и параметров БТ будем соотносить с эквивалентной схемой иерархической электрической модели БТ для активного нормального режима (рис. 7.4), в которой каждый двухполюсник отражает строго детерминированные электрофизические процессы в активной и других областях БТ, включая объем и поверхность p - n -переходов, а также пассивные и периферийные области. Рассмотрим типовые ВАХ для активного нормального режима БТ БИС (рис. 7.5).

Коллекторный ток рассчитывают суммированием плотностей токов J_n , J_p , определяемых непосредственно по разностным аналогам двух уравнений (5.32), (5.34), вдоль линии EG (см. рис. 7.1, a), находящейся в области обратномещенного коллекторного перехода.

Для ВАХ $I_K(U_{\text{Э}})$, $I_B(U_{\text{Э}})$ параметры соответствующих элементов электрической модели — токи насыщения и коэффициенты неидеальности в показателях — определяют непосредственно из экспоненциальных участков для низких уровней инжекции. Параметр I_N — начальный ток высокого уровня инжекции — определяют из ВАХ $I_K(U_{\text{Э}})$, рассчитанной по программе ТРАН 1, когда последовательное сопротивление базы $R_{\text{БА}} = 0$ и отклонение ВАХ $I_K(U_{\text{Э}})$ от идеальной с наклоном φ_T оцениваются эффектом высо-

кого уровня инжекции в базе и коллекторе (на пересечении двух асимптот с наклонами φ_T и $2\varphi_T$ идентифицируется I_N). Коэффициент M_K рассчитывают по формуле $M_K = -\frac{\Delta U_{\text{Э}}}{\Delta U_K}$ из определения тока генератора в базовой цепи I_B — по формуле $M_K = -\frac{\Delta I_B}{\Delta U_K} \cdot \frac{\varphi_T}{I_K}$. Для этого в программе ТРАН 1 рассчитывают характеристику $I_B(U_K)$ и проводят необходимое численное дифференцирование. Выражение ВАХ генератора тока, учитывающего эффект Эрли во входной цепи, записывают в виде $I_{\text{ЭБ}} = M_K U_K / \varphi_T I_{K0} \left(\exp \frac{U_{\text{Э}}}{\varphi_T} - 1 \right)$ (см. рис. 7.4).

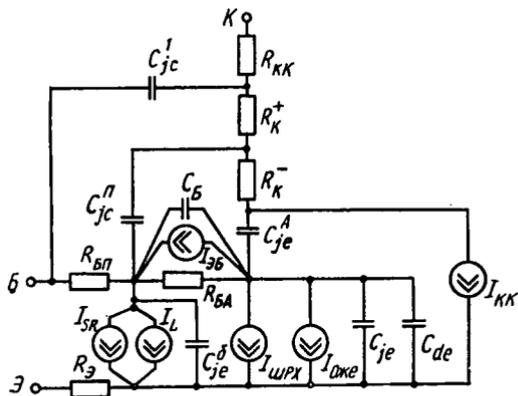


Рис. 7.4. Схема электрической модели БТ для активного нормального режима

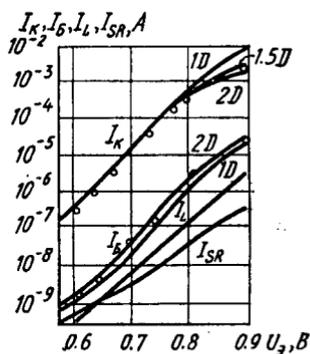


Рис. 7.5. ВАХ составляющих тока БТ, рассчитанные с помощью ФТ-моделей (о — эксперимент)

Методика определения основных электрических параметров БТ базируется на сравнении и анализе ФТ-расчетов БТ по программам ТРАН 1, ТРАН 1.5, ТРАН 2 при выделении в электрической модели БТ (см. рис. 7.4) специфических двухполюсников, моделирующих активную область БТ (программы ТРАН 1, ТРАН 1.5) и пассивные, поверхностные области (программа ТРАН 2).

Зависимости базового тока $I_B(U_{\text{Э}})$: I_B^{2D} , I_B^{1D} рассчитывают интегрированием члена рекомбинации по всему объему БТ. Для тонкослойных БТ с прозрачным эмиттером ($x_{je} \leq 0.1 \div 0.25$ мкм), у которых время переноса в эмиттере много меньше времени жизни неосновных носителей заряда, базовый ток рассчитывают с учетом тока неосновных носителей, не прорекомбинировавших в эмиттере (при конечной скорости рекомбинации на контакте).

Граничные условия на поверхности (4.41)—(4.45) несут исходную информацию о поверхности полупроводника ИМС и БИС;

корректное задание параметров s_n , s_p , ρ_s , площади поверхности позволяет определить ток поверхностной рекомбинации I_{SR} , интегрируя $R_s(S_{n,p})$ по всей поверхности. Типовые значения в ФТ-расчетах ВАХ для БТ: $s_n \approx 500$ см/с, $s_p \approx 3000$ см/с, $\rho_s/q = 2 \cdot 10^{10}$ см⁻². Приведенные значения по скоростям рекомбинации достаточно хорошо соотносятся с экспериментальными данными. В рамках численного ФТ-моделирования существует принципиальная возможность корректного определения s_n , s_p . Для этого по экспериментальной методике, основанной на выявлении поверхностной составляющей тока базы из сравнения токов базы БТ, расположенных на одном кристалле в непосредственной близости друг от друга и существенно различающихся отношениями площади эмиттера к его периметру, строят зависимость $I_{SR}(U_{\text{Э}})$. Из сравнения и настройки экспериментальной и модельной ФТ-зависимостей $I_{SR}(U_{\text{Э}})$ (с учетом оценок тока периферийной инжекции I_L) определяют параметры s_n и s_p реальных БИС.

ВАХ $I_L(U_{\text{Э}})$ определяется таким образом:

$$I_L(U_{\text{Э}}) = I_{\text{Б}}^{2D} - I_{\text{Б}}^{1,5D} - I_{\text{SR}}^{2D}, \quad (7.2)$$

где зависимости базового тока в двумерном и квазидвумерном приближениях $I_{\text{Б}}^{2D}$, $I_{\text{Б}}^{1D}$ рассчитываются по ФТ-программам ТРАН 2, ТРАН 1.5.

ВАХ генераторов тока составляющих токов поверхностной рекомбинации I_{SR} и I_L (периферийной боковой инжекции) описываются выражениями

$$I_{SR} = I_{S0} \exp\left(\frac{U_{\text{Э}}}{m_s \varphi_T} - 1\right), \quad (7.3)$$

$$I_L = I_{L0} \exp\left(\frac{U_{\text{Э}}}{m_L \varphi_T} - 1\right), \quad (7.4)$$

где I_{S0} , I_{L0} — начальные токи; m_s , m_L — коэффициенты неидеальности соответствующих компонент, а ВАХ генераторов тока составляющих токов Шокли — Рида — Холла- и Оже-рекомбинации — выражениями

$$I_{\text{ШРХ}} = I_{\text{Ш0}} \left(\exp \frac{U_{\text{Э}}}{m_{\text{Ш}} \varphi_T} - 1 \right), \quad (7.5)$$

$$I_{\text{Оже}} = I_{\text{О0}} \left(\exp \frac{U_{\text{Э}}}{m_{\text{О}} \varphi_T} - 1 \right). \quad (7.6)$$

На рис. 7.6, а, б приведены типовые расчетные зависимости концентраций подвижных носителей заряда и скорости рекомбинации от координаты для различных уровней инжекции, полученные с помощью программы ТРАН 1, причем при малых уровнях инжекции в объеме p - n -перехода доминирует механизм ШРХ-рекомбинации, при больших — Оже-рекомбинации в квазинейтраль-

ном эмиттере. Для отражения двух специфических механизмов рекомбинации в электрической модели БТ для активной области введены двухполюсники ШРХ- и Оже-рекомбинации, описываемые выражениями (7.5), (7.6). В простейшем случае эти два двухполюсника могут быть свернуты в один, наиболее часто используемый в стандартных моделях БТ.

Анализ рекомбинационных членов в квазинейтральных областях эмиттера, базы и p - n -перехода позволяет заключить, что для БТ рекомбинационный ток в квазинейтральном эмиттере сравним и даже превышает рекомбинационные токи в квазинейтральной базе (для толщин $W_B \leq 0.3$ мкм).

Отмечен рост процентного содержания поверхностной и периферийной составляющих тока базы у данных БТ. Анализ отношений $I_{SR}/I_B^{1.5D}$ показывает, что с увеличением отношения периметра к площади эмиттерного перехода p_3/S_3 БТ (а также с уменьшением толщины W_B) токи I_{SR} , $I_B^{1.5D}$ соизмеримы при $I_K \approx (1 \div 5) \cdot 10^{-7}$ А. Если влияние токов поверхностной рекомбинации существенно для БТ в диапазоне $I_K = 5 \cdot 10^{-7} \div 10^{-8}$ А, то деградацию коэффициента передачи тока B существенно в области средних и повышенных уровней инжекции: $I_K > 3 \cdot 10^2$ А/см². Причем возможный спад B в условиях отсутствия эффектов высокого уровня инжекции может происходить именно вследствие значительного влияния токов I_L , инициируемых эффектами сильного легирования и статического вытеснения. Анализ ВАХ, рассчитанных с помощью одномерных и двумерных ФТ-моделей, позволяет получить полезную зависимость коэффициента передачи тока базы в области малых и

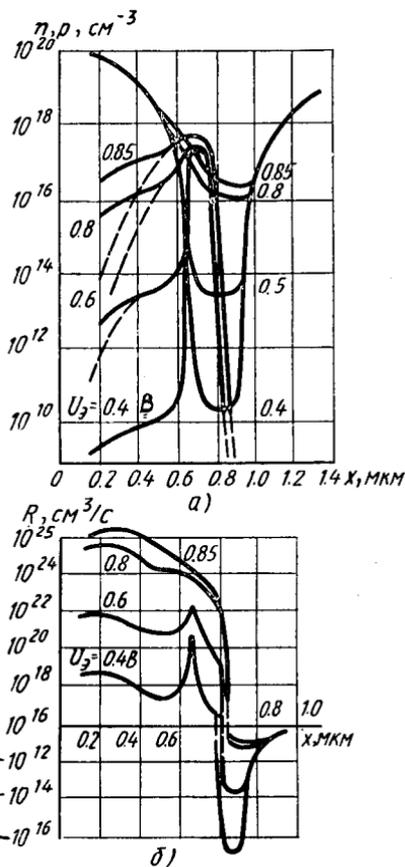


Рис. 7.6. Зависимости концентраций электронов и дырок от координаты и напряжения на эмиттерном переходе для БТ с фосфорным эмиттером (а) (пунктир соответствует расчету концентраций дырок в эмиттере БТ без учета эффектов сильного легирования); зависимости скорости рекомбинации с учетом ШРХ- и Оже-механизмов от координаты и напряжения на эмиттерном переходе БТ (б)

средних токов I_K от величины $S_{\text{Э}}/p_{\text{Э}}$ при фиксированной плотности тока эмиттера ($I_{SR} = p_{\text{Э}} J_{SR}$, $I_B^{1D} = S_{\text{Э}} J_B^{1D}$, $I_L = p_{\text{Э}} x_{je} J_L$):

$$B \approx \frac{K(S_{\text{Э}}/p_{\text{Э}})}{S_{\text{Э}}/p_{\text{Э}} + L + M} - 1, \quad (7.7)$$

где $K = J_{\text{Э}}/J_B^{1D}$, $L = J_{SR}/J_B^{1D}$, $M = (J_L/J_B^{1D}) x_{je}$.

Уменьшение коэффициента B с уменьшением отношения $S_{\text{Э}}/p_{\text{Э}}$ в (7.7) при различных фиксированных токах эмиттера характеризуется в конечном итоге возрастающее влияние токов I_{SR} , I_L (см. рис. 7.5) при увеличении степени интеграции согласно требованию обеспечения быстродействия за счет создания тонкослойных БТ.

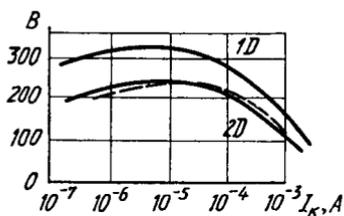


Рис. 7.7. Расчетные зависимости коэффициента передачи тока коллектора БТ от тока коллектора, рассчитанные с помощью одномерной (1D) и двумерной (2D) ФТ-моделей (пунктиром обозначена экспериментальная кривая)

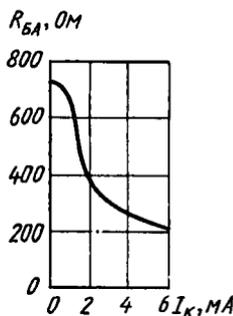


Рис. 7.8. Расчетная ФТ-зависимость активного сопротивления базы БТ от тока коллектора

Таким образом, параметры БТ БИС существенно отличаются от аналогичных параметров ИМС малой и средней степеней интеграции из-за возрастания влияния поверхностных и приповерхностных областей полупроводника. Это связано не только с увеличением вклада поверхностной составляющей тока в ток инжекции. Выявлено значительное возрастание потока неосновных носителей заряда в приповерхностные высоколегированные области (за счет сужения ширины запрещенной зоны полупроводника при его сильном легировании) (рис. 7.7), которое составляет 40—130% против аналогичных структур с относительно слабым легированием в эмиттере и базе.

Для тестовых образцов БТ с «пристеночной» окисной изоляцией, изготовленных по соответствующим технологическим маршрутам, получено хорошее согласование экспериментальных и расчетных зависимостей $B(I_K)$ в широком диапазоне токов (рис. 7.7).

На рис. 7.7 показано также сравнение ФТ-расчетов по одномерной (1D) и двумерной (2D) программам.

Поскольку причинами, вызывающими отклонение зависимости $\lg I_K(U_{\text{Э}})$ от идеальной с наклоном φ_T , являются не только последовательные сопротивления, но и эффекты высокого уровня инжекции в базе и коллекторе, оценку сопротивлений из данной зависимости проводят со значительной погрешностью. С другой стороны, ввиду отсутствия эффекта высокого уровня инжекции в эмиттере (модуляции проводимости в эмиттере) вследствие высокого уровня легирования ($C_{\text{Э}0} > 10^{20} \text{ см}^{-3}$) последовательные сопротивления, вызывающие отклонения зависимости $I_{\text{Б}}^{2D}(U_{\text{Э}})$ от идеальной, могут быть оценены из зависимости $I_{\text{Б}}^{2D}(U_{\text{Э}})$ согласно выражению

$$\varphi_T \ln \frac{I_{\text{Б}}^{2D}}{I_{\text{Б}0}} = U_{\text{Э}} - I_{\text{Б}}(R_{\text{БА}} + R_{\text{БП}}) - I_{\text{Э}}R_{\text{Э}}. \quad (7.8)$$

В рамках ФТ-моделирования представляется более простая возможность идентификации сопротивлений базы, зависящих от режимного тока. Сопротивление активной базы, обуславливающее отличие ВАХ $I_{\text{Б}}(U_{\text{Э}})$, рассчитанных по программам ТРАН 1 и ТРАН 1.5, вычисляют непосредственно по формулам (5.73), (5.73а).

На рис. 7.8 приведена типовая зависимость $R_{\text{БА}}(I_K)$. Различия в ВАХ $I_{\text{Б}}(U_{\text{Э}})$, рассчитанных по программам ТРАН 2 и ТРАН 1.5, дает пассивное базовое сопротивление $R_{\text{БП}}$, вычисляемое по формуле

$$R_{\text{БП}}(I_K) = \frac{|U_{\text{Э}}^{2D} - U_{\text{Э}}^{1.5D}|}{I_{\text{Б}}^{2D}}, \quad (7.9)$$

где $U_{\text{Э}}^{2D}$, $U_{\text{Э}}^{1.5D}$ — напряжения эмиттер—база в двумерном и квазидвумерном расчетах, соответствующие одному и тому же коллекторному току; $I_{\text{Б}}^{2D}$ — базовый ток, рассчитанный по программе ТРАН 2 при напряжении $U_{\text{Э}}^{2D}$.

Отметим, что сопротивление $R_{\text{БП}}(I_K)$, вычисленное данным способом, является эффективной величиной (зависящей от тока), учитывающей также отличие моделирования эффектов высокого уровня инжекции в программах ТРАН 2 и ТРАН 1.5, что отражено в электрической модели БТ (см. рис. 7.4). В большинстве случаев (кроме самосовмещенных БТ БИС с высоким уровнем легирования в активной базе: $C_{\text{Б}} \max > (1 \div 10) \cdot 10^{18} \text{ см}^{-3}$, $R_{\text{БП}} < R_{\text{БА}}$ и в модели (рис. 7.4) вместо зависимости $R_{\text{БП}}(I_K)$ можно использовать $R_{\text{БП}} = \text{const}$ или среднее значение $\bar{R}_{\text{БП}}$ в рабочем диапазоне токов.

Сопротивление эмиттера определяют с помощью программы ТРАН 1. Для этого к профилю примеси исследуемого БТ добав-

ляют дополнительный профиль высокого легирования в базе и коллекторном переходе, что на несколько порядков увеличивает плотности токов начала эффектов высокого уровня инжекции. Помимо этого, моделирование проводят без учета влияния подвижных носителей заряда на коэффициенты подвижностей носителей в эмиттере и базе и при уменьшенных значениях подвижности дырок в $10-10^3$ раз для области квазинейтрального эмиттера $x \leq x_{je}$. При соблюдении данных условий отклонение зависимости $I_{\Sigma}(U_{\Sigma})$ от идеальной с наклоном в области больших плотностей токов ($J > 3 \cdot 10^3 \div 5 \cdot 10^4$ А/см²) можно приписать падению напряжения эмиттерного тока на сопротивлении тела эмиттера R_{Σ} электрической модели БТ (см. рис. 7.4). Отметим, что в данном способе идентификации не учитывается сопротивление контакта эмиттера $R_{\Sigma k}$.

Выражение (7.8) может быть переписано в виде

$$R_{\Sigma} = \left(R_{\Sigma} + \frac{R_{\text{БА}}}{B_{N0}} \right) + \frac{R_{\Sigma} + R_{\text{БП}}}{B_N} = \frac{\varphi_T}{I_K} \ln \frac{I_{B0}}{I_B^{2D}}, \quad (7.10)$$

где I_{B0} — идеальный ток базы (с коэффициентом φ_T); B_{N0} — коэффициент передачи тока активного БТ, рассчитанный по программе ТРАН 1; B_N — коэффициент передачи тока, рассчитанный по программе ТРАН 2.

В данном методе идентификации сопротивления БТ строят зависимость $R_{\Sigma}(1/B)$ для различных режимных токов, в предположении $R_{\text{БА}}/B_{N0} \approx \text{const}$ по экстраполяции $1/B_N \rightarrow \infty$ определяют R_{Σ} , а по наклону функции $R_{\Sigma}(1/B_N)$ — суммарное сопротивление $R_{\Sigma} + R_{\text{БП}}$, или в конечном итоге $R_{\text{БП}}$. Соотношение

$$\frac{\varphi_T}{I_K} \ln \frac{I_{B0}}{I_B} = R_{\Sigma} + \frac{R_{\text{БА}}}{B} + (R_{\Sigma} + R_{\text{БП}})/B \quad (7.11)$$

будет идентично (7.10) только в том случае, когда интегральный коэффициент $B = B_{N0}$, что в принципе выполнимо лишь для БТ, у которых процессами рекомбинации на поверхности и в пассивных областях БТ можно пренебречь. Поэтому оценки сопротивлений R_{Σ} , $R_{\text{БП}}$ надо согласовывать с отношением B/B_{N0} .

В большинстве современных электрических моделей БТ используется концепция управления зарядом при моделировании электрических процессов в динамике. При этом в электрической модели БТ (см. рис. 7.4) специфическим инерционным процессам соответствуют двухполюсники барьерных и диффузионных емкостей для тех или иных областей БТ.

Принципиальная возможность определения инерционных параметров — задержек и емкостей БТ — с помощью ФТ-моделирования основана на расчете «квазистатических» зарядов, накопленных в различных областях БТ. Для этого в рамках ФТ-моделирования проводят численное дифференцирование, вычисление производных от накопленного заряда по току или напряжению.

Полную емкость эмиттерного перехода БТ рассчитывают численным дифференцированием по напряжению U_3 полного потока заряда во всем БТ:

$$C_{Te} = \frac{dQ}{dU_3}, \quad Q = \int_V n dV, \quad (7.12)$$

где $n(x, y)$ — концентрация электронов.

Аналогично вычисляют полную емкость коллекторного перехода $C_{Tc} = dQ/dU_K$.

Барьерная емкость эмиттерного перехода рассчитывается по формулам

$$C_{Je} = \frac{d \int_{V_{p-n}} (n_S - n) dV}{dU}, \quad (7.13)$$

$$\begin{aligned} p_S n_S &= n_{ie}^2 \exp(U_3/\varphi_T), \\ p_S - n_S &= C_A - C_D, \end{aligned} \quad (7.14)$$

где p_S, n_S — концентрации дырок и электронов, соответствующие условию зарядовой нейтральности в объеме полупроводника.

По аналогичным формулам определяют коллекторные барьерные емкости БТ $C_{Jc}^a, C_{Jc}^L, C_{Jc}^n$ (см. рис. 7.4). Отметим, что положение границ интегрирования в формуле (7.13) несущественно при выполнении условия, когда объем V_{p-n} заведомо включает в себя область p - n -перехода. По программам ТРАН 1 и ТРАН 1.5 определяют емкость торцевой части дна p - n -перехода (полную и барьерную емкости C_{Te}^{1D}, C_{Je}), по программе ТРАН 2 — полную и барьерную емкости C_{Te}^{2D}, C_{Je}^{2D} перехода. Емкость боковой части эмиттерного p - n -перехода определяют как разность емкостей, рассчитанных по программам ТРАН 2 и ТРАН 1.

Важнейший параметр БТ — время задержки эмиттер — коллектор — оценивают по формуле

$$\tau_{ЭК} = \frac{d \int_V n dV}{dI_3}. \quad (7.15)$$

Меняя объем интегрирования V , можно вычислить времена задержки БТ в соответствующих областях для двумерной и одномерной ФТ-моделей.

На рис. 7.9, 7.10 приведены типовые результаты моделирования зависимостей полной $C_{Te}(U_3)$ и барьерной $C_{Je}(U_3)$ емкостей эмиттерного перехода, а также времена задержки $\tau_{ЭК}(I_K)$ при $U_K = 0$ В БТ, рассчитанные по программам ТРАН 1 (1D), ТРАН 1.5 (1.5D), ТРАН 2 (2D).

Диффузионная емкость с учетом эффектов высокого уровня инжекции в базе при пренебрежимо малом влиянии эффекта расширения базы записывается в виде

$$C_{de} \approx \tau_{TN} (1 + k_1) \frac{I_K}{\varphi_T}, \quad k_1 = \begin{cases} 0, & I_K \leq I_N, \\ k_1 \left(\frac{I_K - I_N}{I_N} \right), & I_K > I_N, \end{cases} \quad (7.16)$$

где k_1 — коэффициент, учитывающий эффективно в первом приближении квазилинейное увеличение τ_{TN} из-за совокупности эффектов уровня инжекции, I_N — параметр начала эффектов высокого уровня инжекции в базе.

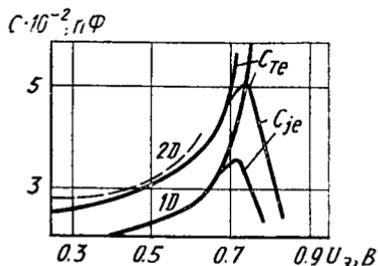


Рис. 7.9. Зависимости полной C_{te} и барьерной C_{je} емкостей эмиттерного перехода БТ от напряжения на эмиттерном переходе, рассчитанные в двумерном (2D) и одномерном (1D) приближениях (пунктиром обозначена экспериментальная кривая)

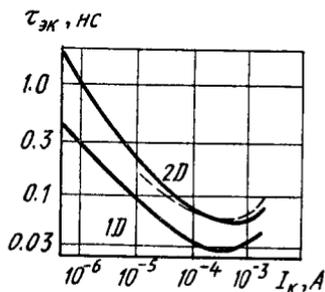


Рис. 7.10. Зависимости времени задержки эмиттер — коллектор БТ от эмиттерного тока, рассчитанные с помощью ФТ-модели в одномерном (1D) и двумерном (2D) приближениях (пунктиром отмечена экспериментальная кривая)

В общем случае, согласно определению, интегральная диффузионная емкость учитывает накопление заряда не только в квазинейтральной области базы, но также в коллекторе, эмиттере и эмиттерном p - n -переходе.

Барьерные емкости (рис. 7.9) хорошо аппроксимируются выражением (6.217). Емкость боковой части эмиттерного p - n -перехода $C_{te}^0 = C_{je}^0 + C_{de}^0$, где C_{je}^0 определяется формулой, аналогичной (6.217), а C_{de}^0 — формулой (7.16), отличающейся модифицированным процессом накопления в эмиттере из-за ЭСЛ.

Влияние барьерной емкости C_{je}^0 боковой части p - n -перехода весьма существенно, что следует из сравнения результатов расчета $C_{je}(1D)$, $C_{je}(2D)$ для небольших $U_{э}$.

Проведенные расчеты с учетом и без учета ЭСЛ в ФТ-модели в одномерном и двумерном приближениях показали значительное влияние ЭСЛ на характеристики барьерных емкостей.

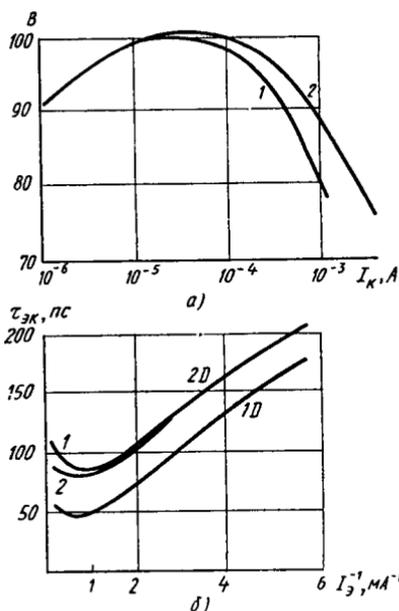


Рис. 7.11. Зависимость коэффициента передачи тока от коллекторного тока (а); зависимости времени задержки эмиттер — коллектор от эмиттерного тока (б) для одномерного (1D) и двумерного (2D) ФТ-расчета, планарные и непланарные контакты эмиттера, базы, коллектора соответственно 1, 2

Аналитическое выражение для расчета начального напряжения φ_0 (6.217)

$$\varphi_0 = \frac{2}{3} \varphi_T \ln \left(\frac{a^2 \epsilon \epsilon_0 \varphi_T}{8 n_{ie}^2} \right)$$

дает незначительное расхождение с результатами, полученными из ФТ-расчета. Здесь a — градиент распределения примесей.

Включение в ФТ-модель ЭСЛ (n_{ie}) приводит к уменьшению φ_0 на 30—40 мВ в программах ТРАН 1 и ТРАН 1.5 и на 40—55 мВ в программе ТРАН 2, при этом барьерная емкость увеличивается на 35—40% в программе ТРАН 1 и на 35—45% в программе ТРАН 2. Таким образом, можно сделать вывод о важности учета эффектов ЭСЛ и влияния боковой сильнолегированной части p - n -перехода.

Время задержки эмиттер — коллектор для двумерного расчета несколько отличается от результатов одномерного ФТ-расчета, что обусловлено в основном увеличением

постоянных времени зарядки боковой емкости p - n -перехода $\frac{\varphi_T}{I_Э} C_{je}^0$,

τ_{TN} для низких плотностей тока. При высоком уровне инжекции различия в результатах одномерного и двумерного ФТ-моделирования обусловлены различием в постоянных времени накопления в эмиттерной и базовой областях и влиянием процессов накопления в боковой части p - n -перехода. Расчет по двумерной ФТ-модели в квазистатическом режиме позволяет учесть влияние эффекта вытеснения на периферийную инжекцию и совокупность эффектов высокого уровня инжекции в базе и коллекторном переходе, тем самым уточнить одномерную ФТ-модель и синтезировать точную электрическую модель БТ.

На рис. 7.11, а б приведены для сравнения типовые зависимости $B(I_K)$, $\tau_{ЭК}(I_Э)$ ФТ-расчета БТ для модельных структур с коллектором внизу и сверху, т. е. с планарным расположением контактов. Анализ результатов расчета параметров B , $\tau_{ЭК}$ поз-

воляет сделать вывод о заметном влиянии сопротивлений коллектора R_{K^+} , R_{KK} и емкостей коллекторного p - n -перехода в режиме квазинасыщения. В области малых и средних токов ($< 5 \cdot 10^{-5}$ А) отличие расчетных характеристик для двух рассмотренных вариантов топологии БТ составляет менее 1—5%.

Сопротивление скрытого слоя коллектора R_{K^+} рассчитывают по данным программы ТЕХИС 1 — поверхностному сопротивлению и топологии скрытого слоя. Для определения подлежащего слоя к контакту коллектора R_{KK} в программу ТРАН 2 вносят также профиль примеси, соответствующий боковому коллекторному переходу и подлежащую n^+ -слою (рис. 7.12). Затем напряжения $U_{ЭБ}$, $U_{КБ}$ выбирают так, чтобы p - n -переход был смещен в обратном направлении и $U_{КЭ} \approx \varphi_T$. Рассчитав теперь ток I_K через поверхность EG (см. рис. 7.1), находим $R_{KK} = U_{КЭ} / I_K$.

Для определения R_{K^-} по программе ТРАН 1 рассчитывают характеристику $I_K(U_K)$ при $0 < U_K \leq 0.5$ В, и при наличии линейного участка этой характеристики, определяемого как $\Delta U_K = I_K R_{K^-}$, находят сопротивление $R_{K^-} = \Delta U_K / \Delta I_K$.

В ФТ-программе ТРАН 1 предусмотрен расчет пробивного напряжения коллекторного p - n -перехода. Для этого при фиксированном прямом напряжении на эмиттере $U_{Э}$ увеличивают с малым шагом обратное напряжение на коллекторном переходе U_K и на каждом шаге рассчитывают величину A :

$$A = \int_{x_1}^{x_2} \alpha_n \exp \left[- \int_{x_1}^{x_2} (\alpha_n - \alpha_p) dx \right] dx.$$

Здесь $\alpha_n = 3.8 \cdot 10^6 \exp(1.75 \cdot 10^6 / |E|)$, $\alpha_p = 2.27 \exp(3.2 \cdot 10^6 / |E|)$ — коэффициенты ионизации, зависящие от напряженности поля E ; x_1 , x_2 — границы коллекторной области пространственного заряда.

При выполнении условия $A \geq 1/B + 1$, где $B = I_K / I_B$, определяют момент наступления лавинного пробоя и критическое пробивное напряжение $U_{пр}$. Отметим, что данный метод, во-первых, не учитывает туннельного пробоя и справедлив при $U_{пр} > 5 \div 6$ В, во-вторых, зависимости $B(U_{Э})$, $U_{пр}(U_{Э})$ дают равенство $U_{пр} = U_{пр \max}$ при $B = B_{\max}$.

7.4. ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ ЭЛЕКТРИЧЕСКОЙ МОДЕЛИ БТ В ИНВЕРСНОМ И НАСЫЩЕННОМ РЕЖИМАХ С ПОМОЩЬЮ ПРОГРАММ ФТ-МОДЕЛИРОВАНИЯ

При двумерном ФТ-моделировании БТ с микротопологией в инверсном и насыщенном режимах, за исключением расчета активного БТ, первостепенную важность приобретает задание реальных геометрий исследуемых БТ БИС с планарным расположением контактов. Рассмотрим методики с использованием программ ТРАН 2 и ТРАН 1 при расчете БТ с планарным расположением контактов для идентификации в активном инверсном режиме дополнительных электрических параметров: емкостей коллектора —

активной S_{KA} , пассивной $S_{KП}$ и боковой S_{KL} областей (с выделением барьерной и диффузионной составляющих); ВАХ генераторов тока базы I_B , $I_{BKР}$ активной области БТ; пассивной и поверхностной составляющих токов I_{LK} , $I_{СК}$.

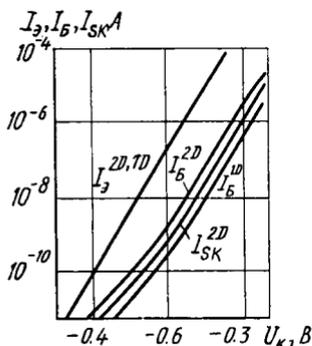


Рис. 7.12. ВАХ в активном инверсном режиме, рассчитанные по ФТ-моделям в одномерном (1D) и двумерном (2D) приближениях

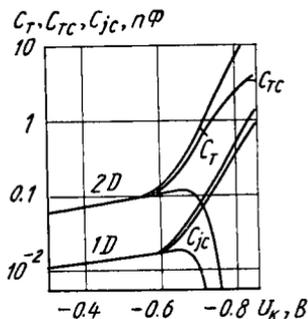


Рис. 7.13. Зависимости полной C_T , полной коллекторной C_{TC} и барьерной C_{jc} емкостей коллекторного перехода, рассчитанные по ФТ-моделям в двумерном (2D) и одномерном (1D) приближениях

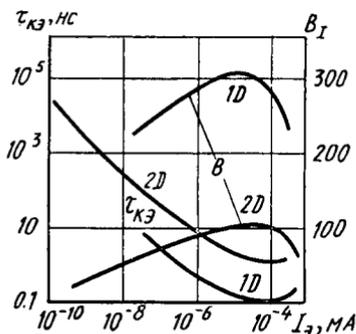


Рис. 7.14. Зависимости времени задержки коллектор — эмиттер и коэффициента передачи тока от тока эмиттера в активном инверсном режиме ($U_E = 0.1$ В), рассчитанные с помощью ФТ-моделей в двумерном (2D) и одномерном (1D) приближениях

Результаты ФТ-моделирования статических и динамических параметров в нормальном инверсном режиме ($U_E = 0.1$ В) приведены на рис. 7.12—7.14. Основа методик идентификации основных «инверсных» параметров БТ — сравнительный анализ ФТ-расчетов БТ по программам ТРАН 2, ТРАН 1 и выбор соответствующего описания двухполюсников в полной электрической модели БТ (рис. 7.15).

Анализ ВАХ доминирующего тока I_K БТ в активном нормальном $I_K(U_E)$ и активном инверсном $I_E(U_K)$ режимах подтверждает наличие единого тока насыщения I_s (при совпадении расчетов

по программам ТРАН 2, ТРАН 1 для низких и средних уровней инжекции). Коэффициенты обратной связи по напряжениям M_{Σ} , M_K в соответствующих цепях БТ рассчитывают с помощью численного дифференцирования по программе ТРАН 1. Начальный ток высокого уровня инжекции в инверсном режиме I_I определяют аналогично I_N по ВАХ $I_{\Sigma}(U_K)$, рассчитанной по программе ТРАН 1 (без учета R_{BA}), как точки пересечения двух асимптот с наклонами φ_T (для низкого уровня инжекции) и $2\varphi_T$ (для высокого уровня инжекции).

Расчет базового тока по программе ТРАН 1 позволяет выделить диффузионную рекомбинационную составляющую I_{BI} в областях базы и коллектора с преобладанием Оже-механизма, а также рекомбинационную составляющую в объеме коллекторного перехода I_{BKR} (с преобладанием ШРХ-механизма), причем

$$I_B^{1D} = I_{BI0} \left(\exp \frac{U_K}{\varphi_T} - 1 \right) + I_{BKR0} \left(\exp \frac{U_K}{m_K \varphi_T} - 1 \right), \quad 1 < m_K < 2. \quad (7.17)$$

Ток поверхностной рекомбинации в активном инверсном режиме I_{SI} рассчитывают

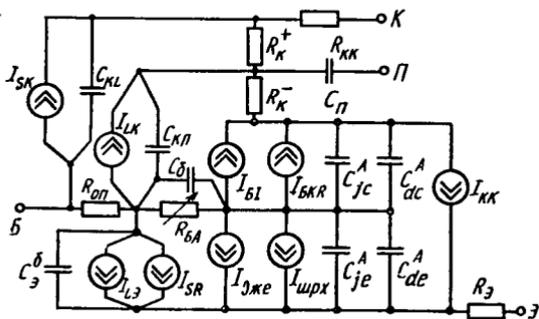


Рис. 7.15. Полная электрическая (компонентная) модель БТ БИС

интегрированием члена рекомбинации R_S , стоящего в граничном условии для ФТ-модели, по всей поверхности раздела окисел — полупроводник. При низких уровнях инжекции $|U_K| < < 0.55$ В доминирующий вклад в составляющую тока I_{SI} (6.138) вносит составляющая рекомбинационного тока I_{SK} (6.138) в местах выхода коллекторного тока на поверхность. При средних и высоких уровнях инжекции доминирующий вклад вносит составляющая рекомбинационного тока I_{LK} в приповерхностной области пассивной базы, связанная с интенсивной инжекцией электронов из пассивной области коллектора.

Таким образом, анализ составляющих полного тока базы в инверсном режиме $I_B^{2D} = I_B^{1D} + I_{SI} = I_{BI} + I_{BKR} + I_{SK} + I_{LK}$ позволяет, используя соотношения (7.17), (6.138), (6.139), выделить токи насыщения и коэффициенты неидеальности непосредственно на графиках рис. 7.12. В ФТ-расчетах по программе ТРАН 2 скорости поверхностной рекомбинации для дырок и электронов составляли $s_n = s_p = 10^3$ см/с.

На рис. 7.13 приведены зависимости полной C_T , полной коллекторной C_{Tc} и барьерной коллекторной C_{jc} емкостей, рассчитанных по программам ТРАН 1 (1D) и ТРАН 2 (2D). Значения барьерных емкостей активной части перехода C_{jc}^a и всего перехода C_{jc} БТ при изменении напряжения $0 < |U_K| < 0.65$ В изменяются всего в 2.5—3 раза, достигая максимума; при дальнейшем увеличении напряжения U_K происходит резкий спад емкостей C_{jc}^{1D} , C_{jc}^{2D} . По разности

$$C_{jc}^{2D} - C_{jc}^{1D} = C_{jc}^L + C_{jc}^n \quad (7.18)$$

идентифицируют барьерные емкости пассивной коллекторной области C_{jc}^n и боковой части p - n -перехода C_{jc}^L . Составляющую C_{jc}^L вычисляют как разность полных барьерных емкостей, рассчитанных с помощью программы ТРАН 2 для планарного расположения контактов и расположения коллектора внизу. Совместный ФТ-расчет коллекторных емкостей в одно- и двумерном приближении, а также с планарным и непланарным расположением контактов позволяет получить исчерпывающую информацию о составляющих коллекторных барьерных емкостей C_{jc}^A , C_{jc}^L , C_{jc}^n в широком диапазоне напряжений на коллекторе и для каждой из них определить параметры аппроксимации барьерной емкости C_0 , φ_0 , n , b .

При напряжении на коллекторе $|U_K| \geq 0.62$ В доминирующий вклад в полную емкость коллекторного перехода как в активной области БТ, так и во всем переходе вносят соответствующие диффузионные составляющие емкостей, связанные с накоплением подвижных носителей заряда в областях коллектора, базы и собственно p - n -перехода. На рис. 7.13 C_{Tc}^{1D} , C_{Tc}^{2D} — полные емкости, рассчитанные по программам ТРАН 1 и ТРАН 2, связанные с накоплением подвижных носителей заряда только в собственно p - n -переходе (так называемые «нейтральные» емкости); емкости C_T^{2D} , C_T^{1D} учитывают накопление подвижных носителей заряда не только в p - n -переходе, но и в соответствующих областях базы и коллектора.

Из сравнения расчетов емкостей C_T^{2D} , C_T^{1D} , а также концентрационных зависимостей можно сделать вывод о доминирующем накоплении подвижных носителей заряда в пассивных областях коллектора, что позволяет характеризовать активную и пассивную области коллектора резко отличающимися эффективными диффузионными емкостями $C_{dc}^n \gg C_{dc}^A$, $C_{dc}^n = C_{dc}^{2D} - C_{dc}^{1D}$. Как показали ФТ-расчеты, для случая пренебрежимо малого расширения базы в область коллектора эффективные диффузионные емкости можно с достаточной точностью описать соотношением

$$C_{dc} = \tau_{Tl} (1 + k_2) \frac{I_{\text{э}}}{\varphi_T} \begin{cases} 0, & I_{\text{э}} < I_l, \\ k_2 \frac{I_{\text{э}} - I_l}{I_l}, & I_{\text{э}} \geq I_l. \end{cases} \quad (7.19)$$

Параметр тока высокого уровня инжекции в инверсном режиме $I_3 \approx I_I$. Диффузионные емкости активной и пассивной областей коллектора БТ характеризуются резко отличающимися постоянными накопления τ_{TI}^a, τ_{TI}^n , причем $\tau_{TI}^n \gg \tau_{TI}^a$. ФТ-расчеты зависимости времени задержки коллектор — эмиттер от тока эмиттера в инверсном режиме (см. рис. 7.14) подтверждают этот факт, давая следующие оценки τ_{TI} для высоких значений тока эмиттера: $\tau_{TI}^a = 0.11$ нс, $\tau_{TI}^n = 1.1$ нс, $I_I^A \approx I_I^n \approx I_I$.

Таким образом, увеличение инерционности процессов накопления в пассивной области базы по сравнению с достаточно малой инерционностью в области активного БТ приводит к резкому возрастанию времени задержки коллектор — эмиттер БТ и выполнению неравенств $\tau_{кэ} \approx \tau_{TI}^n$ и $C_{Tc} \approx C_{Tc}^n$. В то же время сильное влияние процессов рекомбинации в приповерхностных областях пассивной базы ($I_B^{2D} \approx I_{Cк}$) обуславливает существенное отличие полного коэффициента передачи тока B_I , рассчитываемого по программе ТРАН 2, и коэффициента БТ B_{I0} , рассчитываемого по программе ТРАН 1 (см. рис. 7.14); $B_I = (3 \div 6) B_{I0}$.

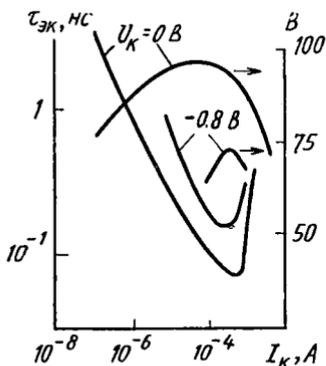


Рис. 7.16. Расчетные (в двумерном приближении) ФТ-зависимости времени задержки эмиттер — коллектор и коэффициента передачи тока от тока коллектора в активном нормальном режиме ($U_k = 0$ В) и в режиме нормального насыщения ($U_k = -0.8$ В)

Для всех режимов функционирования БТ БИС важна информация о емкости переходов скрытый слой — подложка. Данная информация была получена из ФТ-расчетов по программе ТРАН 1 паразитного «подложечного» БТ типа $p-n-p$, когда профиль подложки использовался в качестве эмиттера p -типа, профиль скрытого слоя — в качестве базы n -типа, а профиль базы — в качестве коллектора p -типа. Для площади перехода скрытый слой — подложка $S_n = 350$ мкм² (с учетом расширения мезаобласти в глубь БТ емкость $C_n = C_{пн}(0) \approx 0.1$ пФ). Аналогично может быть получена информация о токе инжекции подложки и ее диффузионной составляющей емкости.

На рис. 7.16 приведены расчетные ФТ-зависимости коэффициента передачи тока и времени задержки эмиттер — коллектор для БТ в режиме нормального насыщения ($U_k = -0.8$ В). Для сравнения приведены аналогичные характеристики в активном нормальном режиме ($U_k = 0$ В). Из-за накопления носителей заряда в областях базы и коллектора в режиме насыщения возрастает ток базы (и его составляющая I_{SR}), что приводит к резкому падению интегрального коэффициента передачи тока B и возрастанию времени задержки $\tau_{эк}$.

7.5. РАСЧЕТ НА ЭВМ СТАТИЧЕСКИХ И ДИНАМИЧЕСКИХ ПАРАМЕТРОВ БТ БИС В РАЗЛИЧНЫХ ТЕМПЕРАТУРНЫХ РЕЖИМАХ

Вопросы оценки температурной стабильности статических и динамических параметров интегральных структур и БИС весьма актуальны при разработке современной и перспективной элементно-технологической базы с использованием ФТ-моделей БТ в изотермических и неизотермических условиях для прогнозирования в широком диапазоне температур параметров БТ и ЛЭ БИС. Экспериментальные исследования в температурном диапазоне динамических параметров БТ — времен задержки, полных и барьерных емкостей, сопротивлений базы — трудоёмки, характеризуются большой экспериментальной сложностью, а также значительной погрешностью измерений для мелких БТ с микротопологией.

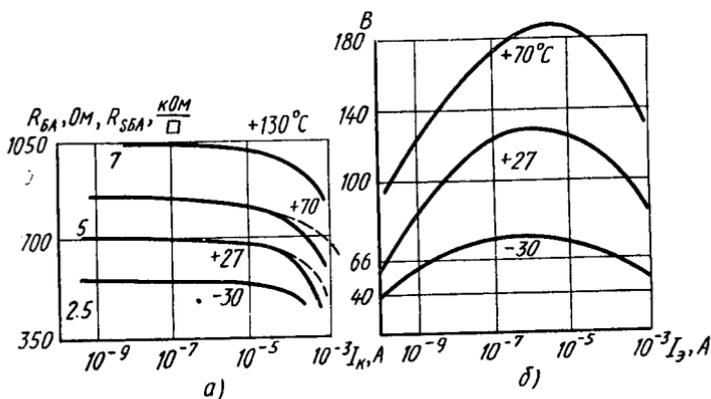


Рис. 7.17. Зависимости активного сопротивления базы и поверхностного сопротивления активной базы от тока коллектора при различных внешних температурах (а); зависимость коэффициента передачи тока от эмиттерного тока при различных внешних температурах (б):

— расчет; - - - эксперимент

Правомочность использования в широком температурном диапазоне изотермической ФТ-модели БТ, рассмотренной в гл. 3, основана на реальном факте постоянства небольшой потребляемой мощности (≤ 1 мВт) для сверхбыстродействующих ЛЭ БИС на переключателях тока и малых напряжениях коллектор — база БТ ($|U_K| < 0.5$ В). Потенциальная возможность влияния изотермических условий в коллекторе БТ данных ЛЭ на динамические параметры, вероятно, будет сказываться при плотностях тока $\geq (8 \div 10) \cdot 10^4$ А/см², что не типично для современных БИС.

Температурные и концентрационные зависимости электрофизических параметров легированного кремния, заложенные в ФТ-программах ТРАН 1, ТРАН 2, были рассмотрены в гл. 4.

В качестве иллюстрации приведем результаты ФТ-расчета статических и динамических параметров типового интегрального БТ в различных температурных режимах (рис. 7.17, а, б; 7.18, а, б).

Из ФТ-расчета ВАХ коллекторного тока I_K , а также составляющих базового тока: идеальной $I_{Oжe}$, $I_{ШРХ}$, токов поверхностной рекомбинации I_{SR} и периферийной инжекции I_L в широком температурном диапазоне — $30 \div 130^\circ\text{C}$ — получены зависимости составляющих токов насыщения от температуры: $I_{K0}(T)$, $I_{O0}(T)$, $I_{ш0}(T)$, $I_{S0}(T)$. Расчетные зависимости полных и удельных компонент то-

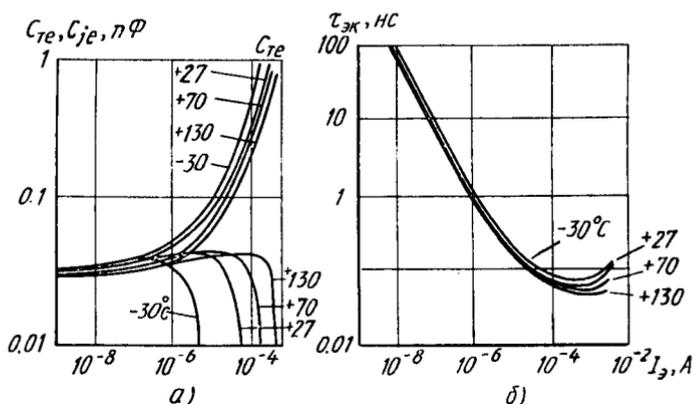


Рис. 7.18. Зависимости полной и барьерной емкостей от тока эмиттера при различных внешних температурах (а); зависимость времени задержки эмиттер — коллектор от тока эмиттера при различных внешних температурах (б)

ков насыщения от параметра $1/kT$ (эВ⁻¹) в предположении равенства внешней температуры и температуры на переходе T дают различный наклон, причем

$$\begin{aligned} I_{K0}(T) &\sim \exp\left(-\frac{1.2}{kT_j}\right), & I_{O0} &\sim \exp\left(-\frac{1.12}{kT_j}\right), \\ I_{ш0}(T) &\sim \exp\left(-\frac{0.6}{kT_j}\right), & I_{S0} &\sim \exp\left(-\frac{0.3}{kT_j}\right). \end{aligned} \quad (7.20)$$

Отметим, что фактор неидеальности m ШРХ-составляющей базового тока изменяется от 1.5 (127°C) до 1.82 (30°C), относительное изменение идеального и эффективного коэффициентов передачи тока:

$$\Delta B_{\max}/\Delta T = 1 \text{ град}^{-1}, \quad \Delta B_{\max}/\Delta T = 0.77 \text{ град}^{-1}.$$

Использование в электрической модели БТ, рассчитанных с помощью одномерных и двумерных ФТ-моделей, температурных зависимостей токов насыщения коллектора и составляющих базового тока в активном нормальном и инверсном режимах (7.20),

позволяет провести более эффективный анализ статических характеристик ЛЭ БИС, включая их помехоустойчивость в широком температурном диапазоне. Исследования на ЭВМ барьерных емкостей эмиттерного и коллекторного переходов обнаружили их слабую зависимость от температуры для напряжения смещения на переходе 1—0.65 В, причем относительное изменение емкостей в диапазоне температур $-30 \div +130^\circ\text{C}$ составило не более 5—15%. Экспериментальные исследования на тестовых БТ подтвердили данные расчетов на ЭВМ. В то же время из-за сильной зависимости контактной разности потенциалов переходов от температуры

$$\varphi_0 \approx \frac{2kT}{3q} \ln \left(\frac{a^2 kT}{8qn_{ie}^2} \right) \text{ при } U \approx \varphi_0 \text{ наблюдается значительное (50—}$$

200%) расхождение барьерных емкостей для различных температурных режимов. Однако в режиме $U \approx \varphi_0$ в полной емкости переходов доминируют диффузионные составляющие, связанные с накоплением носителей заряда в квазинейтральных областях эмиттера, базы и p - n -перехода, причем $C_{de} \gg C_{je}$. Для рассматриваемых БТ температурная зависимость полных емкостей $C_{Te} \approx C_{de} \approx \tau_{TN} I_K (q/kT)$ (рис. 7.18) обусловлена не только температурным потенциалом, но также температурной зависимостью τ_{TN} , $\tau_{TN} \approx \tau_\alpha + \tau_{\tau_B} + \tau_{p-n} + \tau_K$, где составляющие τ_α , τ_{τ_B} , τ_K учитывают зависящие от температуры подвижности носителей. В достаточно тонкослойных БТ зависимости C_{Te} , τ_{τ_K} от температуры нельзя объяснить только зависимостью базовой составляющей $\tau_{\tau_B} \sim W_B^2 / (4D_n)$, как это общепринято для БТ с толстой базой. Для рассматриваемых БТ в области малых токов 10^{-9} — 10^{-4} мА изменение τ_{τ_K} в температурном диапазоне $-30 \leq T \leq 130^\circ\text{C}$ составляет менее 10—17%, что объясняется слабой зависимостью от температуры составляющих $\varphi_T C_{je} / I_\alpha$; при достаточно высоком уровне инжекции в том же температурном диапазоне изменение составляет уже 20—70%.

В указанном температурном диапазоне сильно проявляется зависимость $R_{ГА}(T)$, причем при $I_\alpha < 10^{-5}$ А $\Delta R_{ГА} / \Delta T = 3$ Ом/град. Данное изменение достаточно корректно объясняется зависимостью от температуры в первую очередь поверхностного сопротивления активной базы $R_{СБА} = \left[\int (\mu_{pn} + \mu_{pp}) dx \right]^{-1}$, где подвижности μ_n и μ_p описываются в ФТ-модели соответствующими зависимостями от температуры.

7.6. РАСЧЕТ НА ЭВМ ПЕРЕХОДНЫХ ПРОЦЕССОВ В ТРАНЗИСТОРНОЙ СХЕМЕ С ОБЩЕЙ БАЗОЙ С ПОМОЩЬЮ ПРОГРАММЫ ДВУМЕРНОГО НЕСТАЦИОНАРНОГО МОДЕЛИРОВАНИЯ

Актуальность и необходимость нестационарного ФТ-моделирования тонкослойных субмикронных БТ БИС диктуется чрезвычайной сложностью внутренних инерционных механизмов их функционирования при переключении ЛЭ. Одной из наиболее сложных за-

дач является нахождение реакции коллекторного и базового токов в одно- и двумерном приближениях в транзисторной схеме с общей базой (рис. 7.19, а, б) на импульс эмиттерного напряжения.

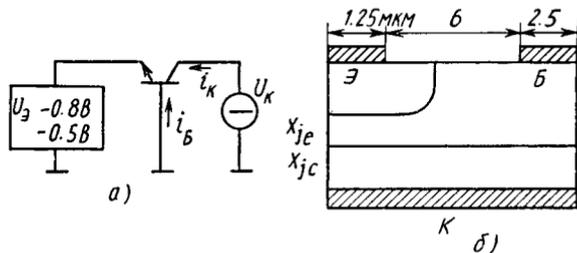


Рис. 7.19. Транзисторная схема с общей базой при нулевой коллекторной нагрузке (а); геометрия БТ, используемых в нестационарных ФТ-расчетах (б)

На рис. 7.20, а, б приведены расчетные зависимости токов коллектора $i_K(t)$ и базы $i_B(t)$ от времени в полулогарифмическом и линейном масштабах при изменении эмиттерного напряжения от

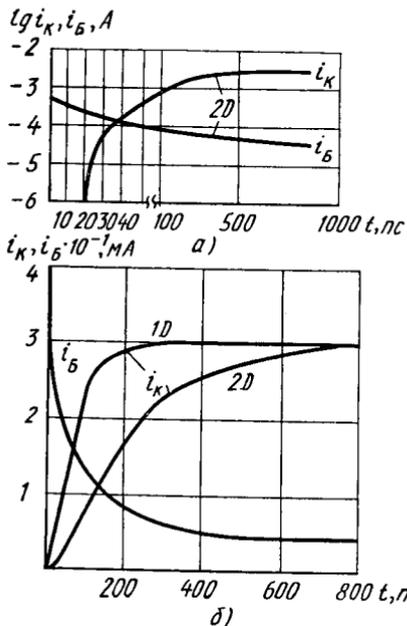


Рис. 7.20. Расчетные ФТ-зависимости токов коллектора и базы от времени в полулогарифмическом (а) и линейном (б) масштабах в одномерном (1D) и двумерном (2D) приближениях

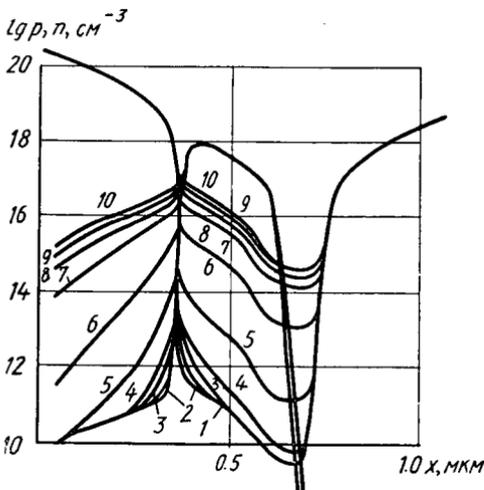


Рис. 7.21. Расчетные зависимости концентраций электронов и дырок от координаты x в центре эмиттера ($y=0$) для различных времен

—0.5 до —0.8 В с фронтом 1 пс, при отсутствии внешнего нагрузочного сопротивления ($R_n=0$) и напряжении на коллекторе $U_k=0$ В. При $t=1\div 2$ пс происходит скачок коллекторного тока $i(t)=4\cdot 10^{-3}$ А, затем резкий спад $\sim -10^{-3}$ А ($t\approx 3$ пс). До $t\approx 15$ пс полный ток коллектора определяется током смещения. В диапазоне 15—1000 пс ток смещения коллектора изменяется от -10^{-4} до 10^{-7} А. Конструктивно-технологические параметры БТ:

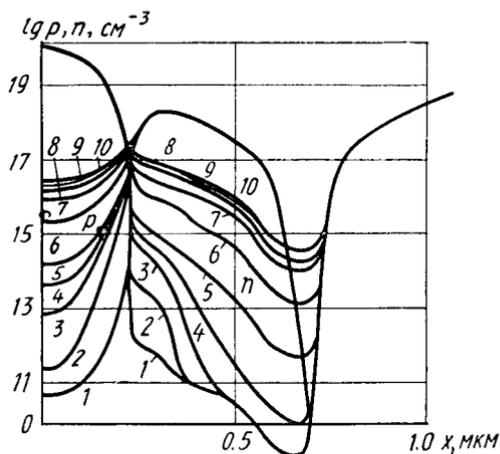


Рис. 7.22. Расчетные зависимости концентраций электронов и дырок от координаты x на краю эмиттера ($y=1.36$ мкм) для различных времен

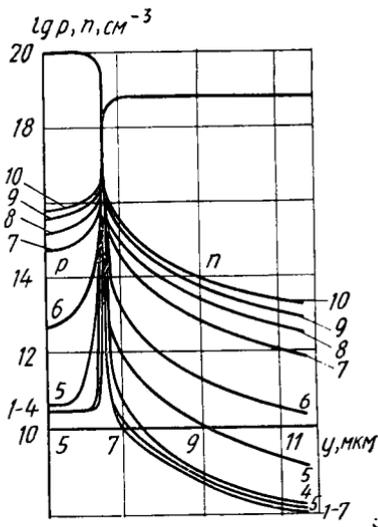


Рис. 7.23. Зависимости концентраций электронов и дырок от координаты y ($x=0.13$ мкм) для различных времен

$x_{je}=0.36$, $x_{jc}=0.68$, $y_{je}=0.22$, $h=2$, $l=6$, $d=6$ мкм, $C_{э\max}=2 \times 10^{20}$ см $^{-3}$, $C_{б\max}=8 \cdot 10^{17}$ см $^{-3}$; скрытый слой коллектора в первом варианте расчета врезается в базовую примесь.

С помощью данного расчета можно детально исследовать инерционные процессы переключения БТ, определяемые цепью базы ($R_n=0$), что актуально для анализа переключателей тока и оценки их потенциального быстродействия. Приведенные на рис. 7.21—7.23 зависимости концентраций дырок и электронов от координат x , y для различных времен переходного процесса в центре и на краю эмиттерного p - n -перехода иллюстрируют инерционные процессы накопления неосновных носителей заряда в базе и эмиттере, а также установление статических напряжений в БТ. Зависимости концентраций носителей заряда от координаты y (рис. 7.23) иллюстрируют эффекты двумерного нестационарного распределения носителей в БТ, влияющие на инерционность изменения токов в области малых времен.

На рис. 7.21—7.23 цифры у кривых означают: 1—0.5 пс; 2—1 пс; 3—2 пс; 4—3.9 пс; 5—10 пс; 6—30 пс; 7—90 пс; 8—150 пс; 9—242 пс; 10—392 пс.

При модельном допущении, что значения полных эмиттерных емкостей БТ, рассчитанных по программам стационарного и не-

стационарного моделирования, совпадают, из времен задержки БТ по уровням $0.1 i_K(t)$, $0.5 i_K(t)$ можно определить эффективное сопротивление пассивной базы $R_{БП}$ и активное базовое сопротивление $Z_{БЛ}(t)$:

$$t_s(0.1 i_K(t)) \approx R_{БП} C_{je} \times \times (2D) + z_{БЛ} C_{je} (1D), \quad (7.21)$$

$$t_s(0.5 i_K(t)) \approx \tau_{тN} + + R_{БП} C_{te} (2D) + z_{БЛ} C_{te} \times \times (1D). \quad (7.22)$$

Для идентификации необходимо использовать расчетные значения полных и барьерных емкостей $C_{te}(2D)$, $C_{te}(1D)$, $C_{je}(2D)$, $C_{je}(1D)$, а также времени переноса $\tau_{тN}$, вычисленных по стационарным программам ТРАН 2, и ТРАН 1 в двумерном (2D) и одномерном (1D)

приближениях. Из анализа зависимостей $t_s(0.1 i_K(t))$, $t_s(0.5 i_K(t))$ (см. рис. 7.20), выражений (7.21), (7.22) следует, что $R_{БП} = 240 \text{ Ом}$, $z_{БЛ} \ll R_{БП}$. Оценка $R_{БП}$ БТ из данных стационарного моделирования по программам ТРАН 1, ТРАН 2 составляла 220 Ом. В данном режиме вследствие динамического эффекта вытеснения в БТ «работает» только периферия эмиттерного перехода, в полной задержке доминирует составляющая, связанная с $R_{БП}$.

Из сравнительного анализа зависимостей $i_K(t)$, $i_B(t)$ для различных фронтов входного сигнала $1 \leq t_f \leq 300$ пс по оценке $Z_{БЛ}(t, I)$ рассчитывают базовую емкость $C_{Б} \approx (0.1 \div 0.4) [C_{te}(1D) + + C_{кА}(1D)]$, где $C_{кА}(1D)$ — активная емкость коллекторного перехода. Быстродействие БТ, работающего в режиме высоких плотностей тока, снижается при меньших токах в условиях большого

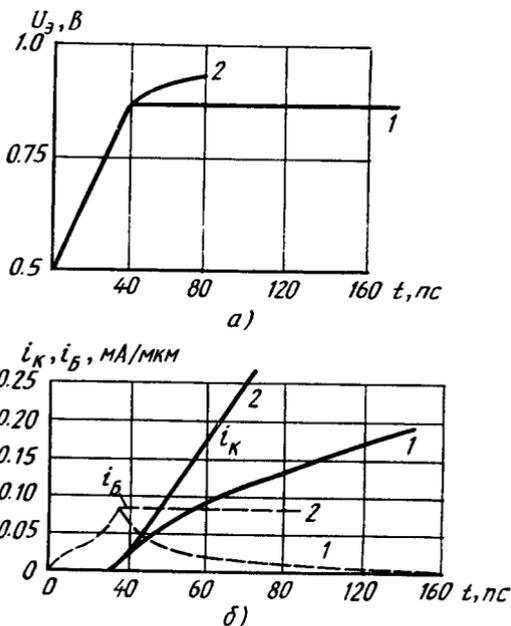


Рис. 7.24. Зависимость напряжения на эмиттерном переходе от времени (а); зависимости коллекторного и базового токов от времени (б), рассчитанные с помощью токов от времени двумерной ФТ-модели при задании на входе БТ генератора напряжения (кривая 1) и генератора тока (кривая 2)

взаимодействия эффектов расширения базы и динамического вытеснения с ростом частоты или уменьшенном фронте входного сигнала.

Большие возможности для идентификации динамических параметров БТ в рамках нестационарного ФТ-моделирования в двумерном (и трехмерном) приближении дают использование управляющих входных сигналов от генераторов тока и напряжения, исследование реакций на эти сигналы токов базы и коллектора в режиме малых времен (рис. 7.24, а, б). Для тонкослойного субмикронного БТ с глубинами залегания p - n -переходов $x_{je}=0.1$ мкм, $x_{jc}=0.2$ мкм расчетный базовый ток доминирует в начале переходного процесса (в схеме, работающей от генератора напряжения), обеспечивая зарядку током смещения барьерных емкостей эмиттерного и коллекторного переходов. Барьерные емкости C_{je} , C_{jc} и сопротивление базы можно оценить из расчетных данных переходного процесса в БТ при управлении от генератора напряжения. При $t=25$ пс напряжение на эмиттере достигает достаточно значения, чтобы доминирующим стал процесс зарядки диффузионной емкости эмиттера.

Из наклона зависимости коллекторного тока от времени в переходном процессе для схемы, работающей от генератора тока (кривая 2 на рис. 7.24, б), определяют постоянную времени τ_{TN} .

Таким образом, численные расчеты переходных процессов в БТ в двумерном приближении позволяют не только определить важные динамические параметры τ_{TN} , $R_{БП}$, $z_{БЛ}$, $C_{Б}$, $R_{БЛ}$, C_{je} , C_{jc} электрической модели БТ, но и предсказать согласуемое с экспериментом ухудшение быстродействия БТ вследствие эффектов высокого уровня инжекции, инициируемых динамическим эффектом вытеснения тока. Это дает в руки разработчиков интегральных приборов и схем эффективный инструмент численного двумерного нестационарного моделирования для оптимизации профиля и топологии БТ с целью устранения нежелательных потерь быстродействия логических элементов в режиме высоких плотностей токов, характерном для работы С³БИС.

7.7. ИССЛЕДОВАНИЕ НА ОСНОВЕ ФИЗИКО-ТЕХНОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ ВЛИЯНИЯ ИЗМЕНЕНИЯ ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ НА ЭЛЕКТРИЧЕСКИЕ ХАРАКТЕРИСТИКИ БТ БИС

Коэффициенты влияния. Широкие возможности совместного использования программ Т- и ФТ-моделирования наиболее ярко проявляются при решении актуальных вопросов влияния тех или иных технологических факторов на интересующие разработчиков электрические параметры схем. Подобные вопросы могут возникнуть, например, на заключительном этапе отработки нового технологического маршрута, когда за счет вариации технологических параметров необходимо довести (улучшить) основные электрические параметры транзисторной структуры, влияющие на быстро-

действие ЛЭ БИС. Варьируемыми параметрами в Т-модели будут дозы и времена разгонки эмиттерной и базовой примесей для формирования оптимальных профилей, обеспечивающих минимизацию быстрой деградации и фактора действия.

Отметим также, что одной из существенных причин уменьшения выхода годных БИС является разброс основных электрических параметров БТ (B , $\tau_{ЭК}$, R_B и др.) и ЛЭ (время задержки, запас помехоустойчивости и др.) вследствие недостаточно жесткого контроля технологических параметров.

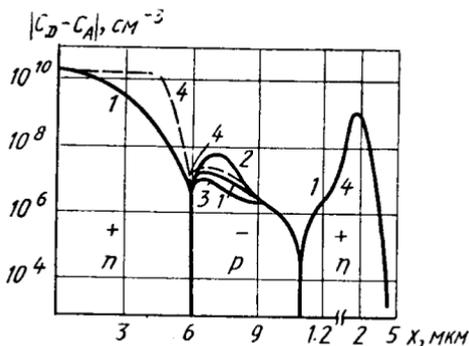


Рис. 7.25. Расчетные профили распределения примеси БТ четырех технологических вариантов

технологии (Т) (например, дозы и времена разгонки) позволяют для основной, в некотором смысле оптимальной, технологии обосновать рекомендации по модификации профиля БТ.

Для оценки влияния вариации технологических параметров на разброс электрических параметров рассчитывались коэффициенты влияния технологических параметров на электрические параметры БТ по формуле

$$K_{вдj}^i = \frac{\Delta \mathcal{E}_i / \mathcal{E}_i}{\Delta T_j / T_j} \approx \frac{\mathcal{E}_0 - \mathcal{E}_i}{(\mathcal{E}_0 \mathcal{E}_i)^{1/2}} \bigg/ \frac{T_0 - T_j}{(T_0 T_j)^{1/2}},$$

где \mathcal{E}_0 , T_0 — значения электрического и технологического параметров номинального БТ; \mathcal{E}_i и T_j — значения i -го электрического и j -го технологического параметров.

Отметим достаточно сложный, в общем случае нелинейный, характер влияния Т-параметров на Э-параметры. Коэффициенты влияния, рассчитанные для одного и того же технологического параметра в соседних, но не очень близких режимных точках технологического маршрута, могут сильно отличаться друг от друга. Данные по $K_{вд}$ позволяют определить, разброс какого из технологических параметров сильнее всего скажется на разбросе электрических параметров, а следовательно, на выходе годных БИС, изготовленных по данной технологии.

Исследование на ЭВМ влияния профиля распределения примесей в эмиттере, базе и коллекторе на электрические характеристики БТ БИС. Данная рубрика посвящена исследованию на ЭВМ влияния профиля распределения примесей в эмиттере, базе и коллекторе на инерционные характеристики БТ в рамках ФТ-моделирования с целью выдачи рекомендаций по выбору технологических маршрутов изготовления БИС и топологий БТ. С помощью программы ТЕХИС 1 были рассчитаны концентрационные зависимости профиля БТ (рис. 7.25), технологический маршрут изготовления которого (модифицированный вариант БТ 1) включал следующие основные операции: имплантацию сурьмы в подложку, разгонку сурьмы с окислением, эпитаксию, окисление, имплантацию бора, разгонку бора, имплантацию фосфора (мышьяка), разгонку фосфора (мышьяка).

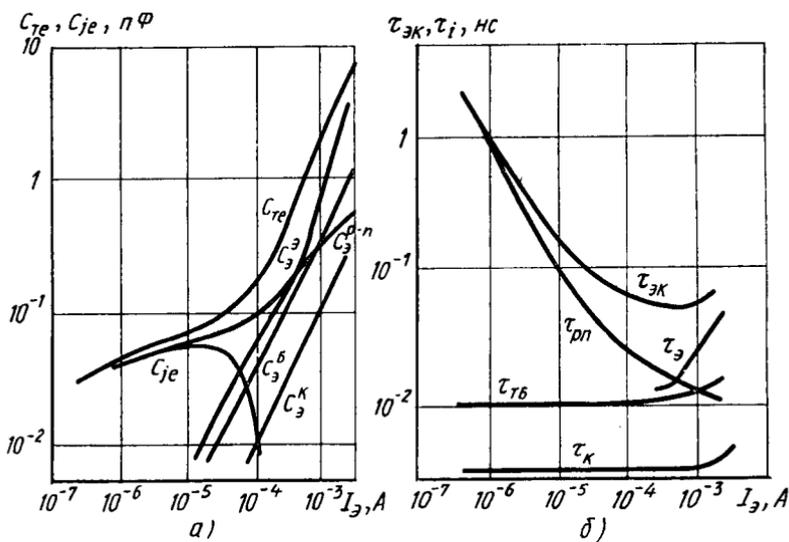


Рис. 7.26. Расчетные зависимости полной емкости эмиттерного перехода и ее составляющих от тока эмиттера (а); расчетные токовые зависимости (б) времени задержки эмиттер — коллектор и его составляющих (для Т-варианта 1)

Относительно базового Т-варианта 1 для вариантов 2 и 3 доза бора в базе была выбрана так, чтобы максимальная эффективная концентрация в базе была в два раза больше или меньше максимального уровня концентрации в варианте 1. В качестве эмиттерной примеси Т-варианта 4 был выбран мышьяк. Дополнительные варианты 5, 6 отличались от базового варианта 1 только различным распределением примеси в коллекторной области; в варианте 5 использовался эпитаксиальный слой толщиной $W_K = 0.3$ мкм с концентрацией $C_K = 5 \cdot 10^{15}$ см $^{-3}$ до скрытого слоя

сурьмы, в варианте 6 — слой толщиной $W_K=0.15$ мкм и концентрацией $1 \cdot 10^{15}$ см $^{-3}$.

Для Г-варианта 1 на рис. 7.26, а приведены расчетные токовые зависимости полной и барьерной емкостей эмиттерного перехода БТ, а также составляющих полной емкости, связанных с накоплением заряда в эмиттере C_3^3 , p - n -переходе C_{3n}^p , базе C_3^b , коллекторе C_3^k при соответствующем разбиении БТ на ряд областей, а на рис. 7.26, б — расчетные токовые зависимости суммар-

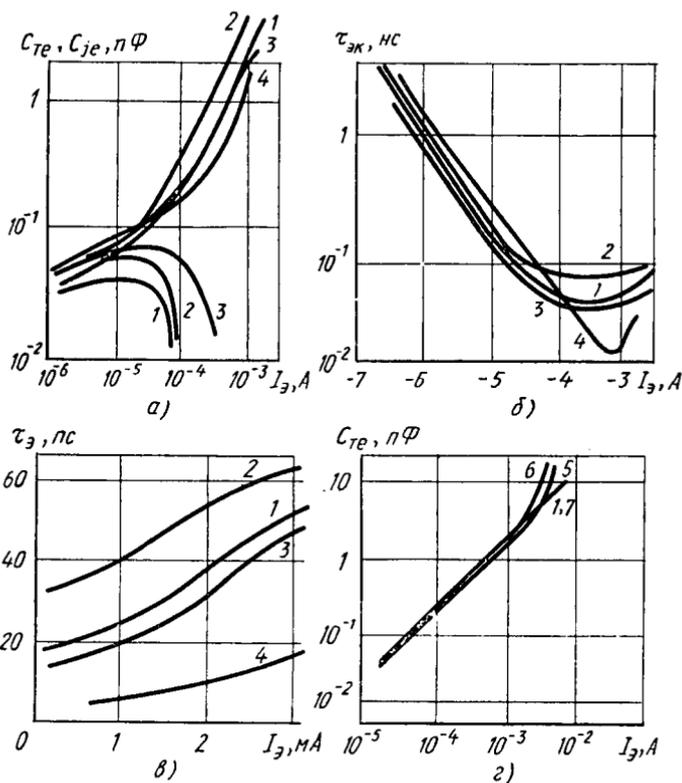


Рис. 7.27. Зависимости полной и барьерной емкостей (а), времен задержки эмиттер — коллектор (б), времен накопления в эмиттере от тока эмиттера (в) для четырех технологических вариантов; зависимости полной емкости эмиттера от тока для БТ с различными профилями в коллекторе $S_3 = 2.5 \times 6$ мкм (г)

ного времени задержки эмиттер — коллектор $\tau_{ЭК}$ и его составляющих — постоянных времени накопления заряда в эмиттере τ_3 , эмиттерном p - n -переходе τ_{p-n} , базе τ_{TB} , коллекторном p - n -пере-

ходе и коллекторе τ_k . Наблюдается доминирующий вклад постоянной времени τ_3 в $\tau_{эк}$; для диапазона токов эмиттера $10^{-4} \leq I_3 \leq 5 \cdot 10^{-3}$ А выполняется соотношение параметров $\tau_3 = (1.5 \div 5) \tau_{тБ}$. В том же диапазоне токов имеется доминирующий вклад составляющей C_3^3 , причем $C_3^3 = (1.4 \div 5) = C_3^5$. Важно подчеркнуть, что аналогичные расчеты по ФТ-программе, проведенные для БТ, изготовленных по стандартным технологиям с шириной базы 0.6—0.8 мкм, дают доминирующий вклад $\tau_{тБ}$ в $\tau_{эк}$ и C_3^5 в C_3^3 , причем $\tau_3 \ll \tau_{тБ}$ и $C_3^3 \ll C_3^5$. Дополнительная составляющая времени задержки τ_{p-n} ($I \geq 10^{-3}$ А), связанная с «нейтральной» емкостью C_{p-n}^3 , сравнима с $\tau_{тБ}$: $\tau_{pn} \approx \tau_{тБ} \cdot \tau_{тБ}$.

При увеличении тока $I_3 \geq 10^{-4}$ А БТ с «мышьяковистым» эмиттером характеризуются минимальной емкостью, примерно в 1.5—3.5 раза меньшей $C_{те}$ БТ Т-вариантов 1—3 (рис. 7.27, а). Для данных БТ с высоким градиентом распределения примесей ($a = 10^{-24}$ см $^{-4}$) по сравнению с другими БТ (Т-варианты 1—3) минимальны постоянные времени как τ_{pn} , так и τ_3 (рис. 7.27, б, в). Вследствие более сильной инжекции дырок в эмиттер для БТ Т-варианта 2 с более высоким легированием базы справедливо условие $\tau_{31} > \tau_{32} > \tau_{31} > \tau_{34}$. (Для данных расчетов интегрирование заряда электронов велось от контакта эмиттера до x_{je} .)

На рис. 7.27, г приведены для сравнения четыре токовые зависимости полных емкостей эмиттера для БТ с одинаковыми профилями в эмиттере и базе (аналогично фосфорному профилю 1), но с различными распределениями примеси в коллекторе: с врезанием скрытого слоя в базу (1), эпитаксиальными слоями толщиной $W_k = 0.3$ мкм, с концентрацией $C_k = 5 \cdot 10^{15}$ см $^{-3}$ (5), $W_k = 0.15$ мкм, $C_k = 1 \cdot 10^{15}$ см $^{-3}$ (6), $W_k = 0.01$ мкм, $C_k = 5 \cdot 10^{15}$ см $^{-3}$ (7).

Важным результатом анализа зависимостей рис. 7.26, 7.27 является также смещение границы эмиттерного перехода в сторону эмиттера при высоких уровнях инжекции (начальная плотность тока эффекта расширения базы в эмиттер $J_k = 5 \cdot 10^3$ А/см 2). Это смещение обусловлено процессом нейтрализации объемного заряда (сильнее на эмиттерной стороне, чем на базовой) из-за асимметрии распределения примеси относительно глубины залегания перехода. Нейтрализация области эмиттера на эмиттерной стороне $p-n$ -перехода при высоких уровнях инжекции приводит к интенсификации процесса накопления подвижных носителей заряда в эмиттере и эмиттерной части перехода и соответственно к уменьшению быстродействия БТ. Пунктиром на рис. 7.28 обозначены зависимости напряженности полей в эмиттере при $U_3 = 0.4; 0.9$ В для БТ с одинаковыми профилями примеси в базе и коллекторе и геометрией, но с мышьяковистым эмиттером, характеризующимся более высоким градиентом распределения примеси в переходе ($a_{As} = 10^{24}$ см $^{-4}$; $a_p = 10^{23}$ см $^{-4}$). Можно сделать вывод, что использование мышьяка в качестве эмиттерной примеси позволяет создавать мелкие БТ с более высоким быстродействием, чем у

«фосфорных» БТ, за счет уменьшения постоянных времени накопления в эмиттере и эмиттерном переходе.

Взаимодействие программ физико-технологического комплекса может быть проиллюстрировано на примере двумерного расчета и

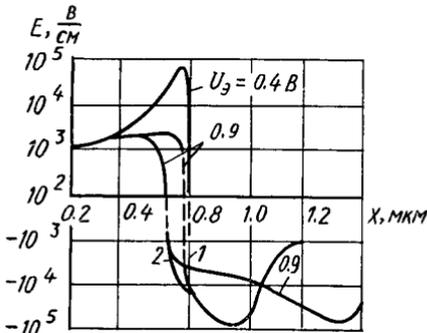


Рис. 7.28. Расчетные зависимости напряженности электрического поля от координаты при различных эмиттерных напряжениях для БТ с мышьяковистым (кривая 1) и фосфорным (кривая 2) эмиттерами

оптимизации поликремниевое тонкослойного сверхсамосовместимого БТ (см. рис. 1.3) с профилями, показанными на рис. 7.29. Эффект изменения дозы легирования примеси бора на 30% по сравнению с базовой технологией отражен в изменении зависимостей коэффициента передачи тока и поверхностного сопротивления активной базы (рис. 7.30, а, б); с уменьшением уровня легирования базы значения B и $R_{сБА}$ возрастают. На рис. 7.30, в показаны соответствующие зависимости полных емкостей эмиттера и коллектора с различным профилем в коллекторе. При низких режимных токах БТ с

профилем ν имеют, например, большую полную коллекторную емкость, обусловленную ее барьерной составляющей. Однако с ростом коллекторного тока увеличение полных емкостей эмиттера и коллектора БТ с профилем ν значительно меньше вследствие того, что расширение базы начинается при больших плотностях тока.

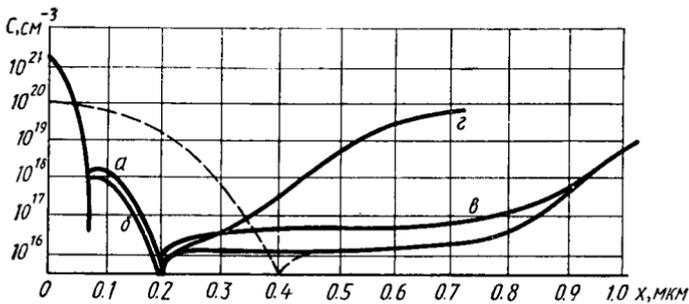


Рис. 7.29. Расчетные профили распределения примесей тонкослойного БТ при ФТ-моделировании

Взаимодействие программ технологического и нестационарного физико-топологического моделирования проиллюстрировано на примере расчета переходных процессов в ряде БТ с одинаковой топологией ($S_3 = 2.5 \times 6$ мкм), но с различными профилями при-

месей в базе и коллекторе (рис. 7.31, а, б). Наименьшей инерционностью характеризуется БТ с профилем 3, для которого наимень-

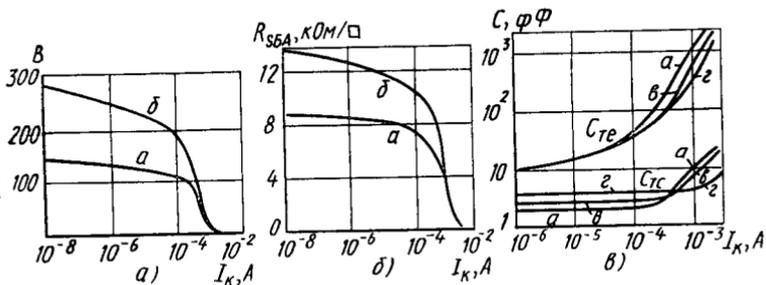
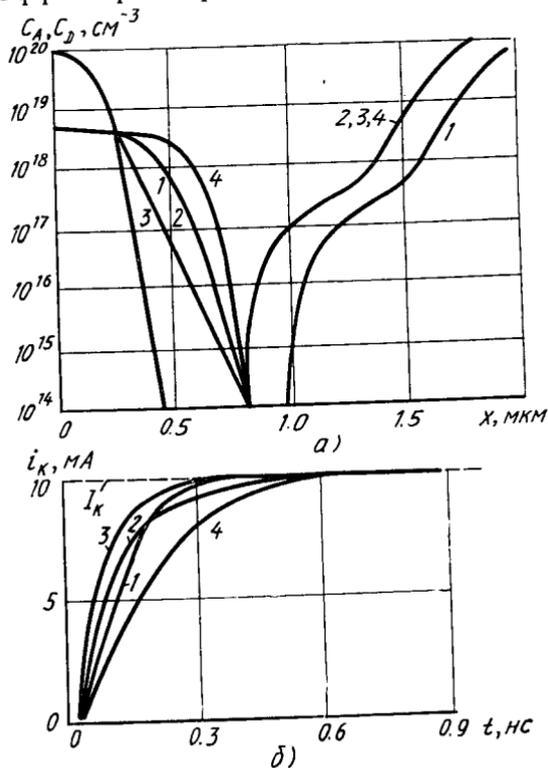


Рис. 7.30. Зависимости от режимного тока коллектора коэффициента передачи тока (а), поверхностного сопротивления активной базы (б), полных емкостей эмиттера и коллектора (в); $S_{\delta} = 1.25 \times 2$ мкм, $S_K = 4.7 \times 2$ мкм

шая степень легирования базы приводит к более высокой подвижности носителей заряда и соответственно к меньшему времени переноса носителей в базе. Эффект расширения базы очевиден из сравнения расчетов переходных процессов для БТ 2 и БТ 1 с «отодвинутым» скрытым слоем коллектора (рис. 7.31, б). Причем по уровню $0.1I_K$ меньшей задержкой характеризуются БТ с профилем 2, а по уровню $0.9I_K$ — БТ с профилем 1. Наибольшей инерционностью характеризуется профиль 4 с наиболее сильным легированием всей активной базовой области.

Введение в статистическое моделирование. Рис. 7.32 иллюстрирует выявленную корреляцию тока насыщения I_s , по-

Рис. 7.31. Расчетные профили примесей БТ (а); соответствующие зависимости коллекторного тока от времени, рассчитанные с помощью нестационарной ФТ-программы моделирования (б)



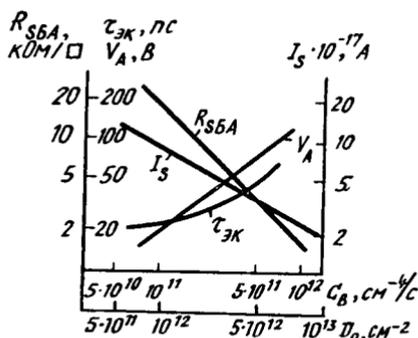


Рис. 7.32. Корреляция электрических параметров времени задержки эмиттер — коллектор, тока насыщения, напряжения Эрли, поверхностного сопротивления активной базы с числом Гуммеля в базе

верхностного сопротивления активной базы R_{SBA} , напряжения Эрли V_A , а также времени задержки эмиттер — коллектор $\tau_{ЭК}$ с «числом Гуммеля» в базе $G_B = \int C_A(x) dx$. Корреляции I_s , R_{SBA} , $\tau_{ЭК}$, V_A и C_B полезны при статистических исследованиях БИС и минимизации натуральных экспериментов для полного описания всех электрических характеристик и параметров БТ.

С точки зрения обработки технологического процесса наиболее эффективен контроль начального тока I_s по ВАХ $I_K(U_E)$ вследствие простоты его измерения.

Использование упрощенных моделей БТ, где на основе регрессионного анализа получены линейные аппроксимации всех статистических и динамических параметров БТ от токов насыщения I_s , имеет большие преимущества при проектировании БИС. Создавая подобные электрические модели полупроводниковых приборов, учитывающих разбросы основных технологических параметров, можно разработать реальные статистические модели для наилучшего бездефектного проектирования. Вследствие этого огромную перспективу имеет совместное статистическое физико-технологическое и электрическое моделирование БИС, в рамках которого формализуется метод связывания разбросов технологических параметров с соответствующими параметрами интегрального прибора и всей БИС. В этом случае можно применить математический аппарат случайных технологических и электрических параметров БТ, направленный на оптимизацию БТ, улучшение характеристик элементов БИС, увеличение выхода годных схем.

ОПТИМИЗАЦИЯ И МАКРОМОДЕЛИРОВАНИЕ ИНТЕГРАЛЬНЫХ БЫСТРОДЕЙСТВУЮЩИХ ЭЛЕМЕНТОВ И СХЕМ

8.1. СХЕМО-ТЕХНОЛОГИЧЕСКАЯ ОПТИМИЗАЦИЯ ЛОГИЧЕСКИХ ЭЛЕМЕНТОВ БИС. АДАПТАЦИЯ СТРУКТУРНО-ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ БТ К ПОТРЕБЛЯЕМОЙ МОЩНОСТИ ЭЛЕМЕНТОВ БИС

В инженерной практике расчет ИМС малой и средней степеней интеграции ранее был практически отделен от расчета топологии активного элемента — БТ и тем более от выбора технологических режимов изготовления БТ. Обычно, получив от технологов и разработчиков приборов основные электрические параметры модели активного элемента — БТ, разработчики схем подбирают электрические режимы (внешние параметры: сопротивления, емкости, напряжения питания и напряжения цепей смещения) для обеспечения требуемой по техническому заданию совокупности факторов: быстродействия, потребляемой мощности, помехоустойчивости и др. Однако условие достижения требуемого (или максимального) быстродействия, потребляемой мощности, особенно для схем на биполярных элементах, зачастую сопряжено с большой потребляемой мощностью ЛЭ и соответственно всей БИС. Основными путями уменьшения параметров t_3 и ρ ЛЭ являются системный и схемотехнический, когда за счет улучшения структурно-организационных принципов построения БИС на сложных ЛЭ с большой логической гибкостью удается уменьшить значения параметров t_3 и ρ на выполнение элементарной логической функции. Средством снижения t_3 и ρ служит также уменьшение логического перепада U_L внутри БИС до значений, определяемых запасом помехоустойчивости логических цепей, дающих возможность нормального функционирования в заданных эксплуатационных условиях.

Кроме проблемы уменьшения U_L ЛЭ центральное место в разработке быстродействующих БИС занимает проблема выбора оптимального элемента БТ и адаптации его структурно-технологических параметров к потребляемой мощности ЛЭ БИС. Под оптимальным физико-технологическим элементом понимают элемент с таким распределением примесей и конструктивно-топологическими особенностями в глубь и плане кристалла, которые обеспечивали бы оптимальные быстродействие, потребляемую мощность и другие необходимые характеристики для надежного функциони-

рования БИС, включая максимальный выход годных схем. Причем при разработке субнаносекундных БИС схемотехнический расчет схем и фрагментов не должен быть отделен от расчета элементов БТ, так как только совместная схемотехническая и технологическая их оптимизация позволяет за счет разработки оптимальных элементов БТ и выбора оптимального электрического режима функционирования ЛЭ создать высококачественные БИС.

Таблица 8.1

Влияние структурно-технологических факторов на электрические параметры БТ

Параметр	$V \uparrow$	$\tau_{ЭК} \downarrow$	$R_{БА} \uparrow$	$R_{БП} \downarrow$	$C_{je} \downarrow$	$C_{jc} \downarrow$	$J_B \uparrow$	$J_M \uparrow$	$J_D \uparrow$
$C_K \uparrow$		\downarrow							\uparrow
$C_E \uparrow$	\uparrow	\downarrow					\uparrow	\uparrow	
$\bar{C}_B \uparrow$	\downarrow	\uparrow	\downarrow	\downarrow			\uparrow	\uparrow	
$W_E \downarrow$	\downarrow	\downarrow							
$W_K \downarrow$	\uparrow	\downarrow					\uparrow		
$W_B \downarrow$	\uparrow	\downarrow	\uparrow						\uparrow
$a \uparrow$	\uparrow	\downarrow	\downarrow		\uparrow				
$h \downarrow$			\downarrow		\downarrow	\downarrow	\downarrow	\downarrow	\downarrow
$l \downarrow$			\uparrow	\uparrow	\downarrow	\downarrow	\downarrow	\downarrow	\downarrow
$S_K \downarrow$									\downarrow

Совокупность требований к электрическим характеристикам БТ для обеспечения максимального быстродействия БИС на переключателях тока (ПТ) и эмиттерных повторителях (ЭП) порождает противоречивые требования к его технологическим и конструктивно-топологическим параметрам. В табл. 8.1 показано влияние данной технологии (интегральные параметры вертикального профиля — концентрации примесей в эмиттере, базе, коллекторе C_E , C_B , C_K , ширины квазинейтральных областей W_E , W_B — и топологии — ширины и длины эмиттера h , l , площади коллектора S_K) на электрические параметры БТ для обеспечения его оптимального исполнения в субнаносекундной БИС. Для оценки влияния указанных параметров, направленных на минимизацию оптимальных времен задержки $t_{3\min}$, знаками \uparrow и \downarrow обозначены требования увеличения или уменьшения исходного параметра. Положительная тенденция отмечена совпадением знаков изменения входного и выходного параметров, отрицательная — несовпадением. Например, с одной стороны, уменьшение W_B приводит к увеличению V и уменьшению $\tau_{ЭК}$ (но есть ограничение на уменьшение W_B для достижения больших V и малых $\tau_{ЭК}$ вследствие рассмот-

ренных ранее эффектов высокого легирования и механизмов переноса в сверхтонкой области полупроводника). С другой стороны, уменьшение W_B приводит к росту $R_{сба}$, $R_{ба}$; следовательно, условия уменьшения W_B противоречивы, есть принципиальная возможность оптимизации данного параметра для достижения минимального времени задержки $t_{з\ min}$. Другой пример. Для увеличения напряжения пробоя коллектора БТ необходимо использовать эпитаксиальный слой с большим значением $W_{эп}$ и малым значением $S_{эп}$, в то же время ограничение на величину R_k и требование высокого быстродействия требуют уменьшения $W_{эп}$ и увеличения $S_{эп}$. Уменьшение ширины и длины эмиттера БТ приводит к уменьшению емкостей БТ, но и к нежелательному увеличению сопротивления базы, уменьшению плотностей токов начала эффектов вытеснения, модуляции проводимости и расширения базы (J_B , J_M , J_P).

Разработка неоптимального БТ БИС вызывает: 1) неоправданную потерю быстродействия ЛЭ; 2) перерасход потребляемой мощности, а следовательно, снижение качества и надежности БИС, увеличение ее стоимости за счет введения дорогих средств охлаждения; 3) неполное использование возможностей современной технологии; 4) пониженный процент выхода годных схем или увеличение брака.

Задачи и подходы к оптимизации БИС. Как правило, выбор оптимального (оптимальных) БТ и БИС и оптимизацию быстродействия самой БИС необходимо рассматривать как одну из сложнейших задач математического программирования — параметрической оптимизации. В общем случае задача параметрической оптимизации заключается в определении вектора параметров \mathbf{P} , минимизирующих целевую функцию $\Phi(\mathbf{P})$ при заданных ограничениях для систем алгебраических или дифференциальных уравнений, определяющих математическую модель схемы:

$$\min \Phi(\mathbf{P}), \quad r(\mathbf{P}) \leq 0, \quad \mathbf{d}_1 \leq \mathbf{P} \leq \mathbf{d}_2,$$

где $\Phi(\mathbf{P})$ — целевая функция, например быстродействие схемы; \mathbf{P} — вектор параметров оптимизации; неравенство $r(\mathbf{P})$ накладывает ограничения на выходные параметры схемы; \mathbf{d}_1 , \mathbf{d}_2 — векторы, ограничивающие область определения параметров оптимизации, определяемые из конструктивно-технологических возможностей реализации БИС.

Для различных формулировок задач оптимизации и методов их решения на каждом шаге поиска следует вычислять значения целевых функций $\Phi(\mathbf{P})$ и градиенты $d\Phi(\mathbf{P})/d\mathbf{p}$, $d\mathbf{r}/d\mathbf{p}$, т. е. в многоэтапном процессе оптимизации многократно решать задачу анализа при различных значениях вектора параметров. Однако вследствие высокой размерности вектора параметров оптимизации \mathbf{P} , обусловленных наличием в БИС большого количества БТ с общим количеством изменяемых параметров более 10^3 — 10^5 , решение задачи оптимизации БИС затруднительно и нецелесообразно

для большинства современных ЭВМ с ограниченными вычислительными ресурсами.

Основным практическим методом, открывающим перспективы реального оптимального расчета БИС и их элементов, является декомпозиция сепарабельных целевых функций, записываемая, например, в виде

$$\Phi(P) = \sum_{j=1}^M \varphi_j(P_j),$$

где M — количество фрагментов в БИС; j — номер фрагмента схемы; P_j — вектор параметров оптимизации j -го фрагмента.

Например, функции φ_j могут быть задержками в j -м фрагменте (либо потребляемая мощность j -м фрагментом). Эффективен подход к решению задач с сепарабельной целевой функцией, основанный на оптимизации переключательных характеристик $t_3(p)$ базовых ЛЭ БИС. При этом в принципе возможно решить задачу достижения оптимального быстродействия и потребляемой мощности БИС за счет общего распределения и дальнейшего дополнительного перераспределения ресурсов мощности по базовым ЛЭ фрагментов БИС, состоящих из последовательно включенных ЛЭ.

Выбор моделей БТ. Рассматриваемый ниже метод совместной и схемотехнической оптимизации быстродействия ЛЭ и БТ на основе неформальных многовариантных расчетов переключательных характеристик ЛЭ основан прежде всего на подтвержденной приемлемой адекватности электрической модели БТ, используемой для схемотехнических расчетов в широком режимном диапазоне (рис. 8.1, а—в). Ре-

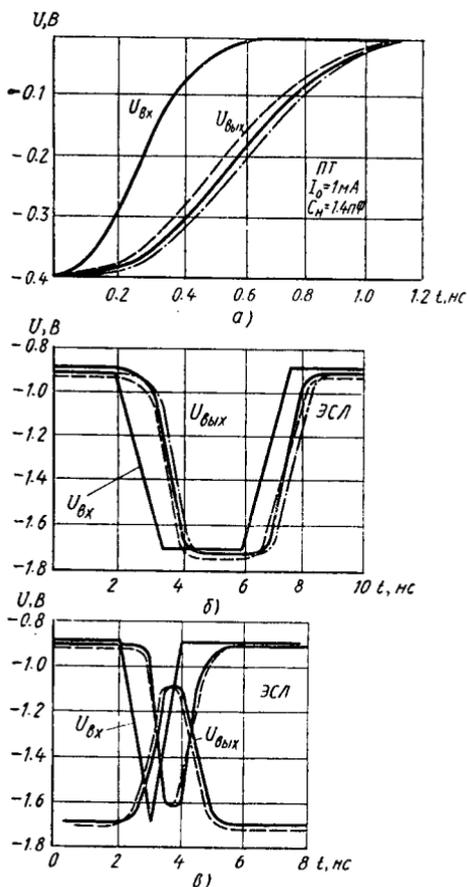


Рис. 8.1. Расчетные и экспериментальные зависимости переходного процесса в токовом переключателе (а), нагруженном ЭСЛ-элементе с выходными эмиттерными повторителями БИС (б); расчет реакции на импульсную входную треугольную помеху ЭСЛ-элемента БИС (в):

— эксперимент; — — — — —: расчет (двух- и трехсекционные модели БТ); — · — · — ·: расчет (односекционные эффективные модели БТ)

зультаты расчетов убедительно показывают достаточно малую погрешность моделирования переходных процессов 1—15% в элементах эмиттерно-связанной логики (ЭСЛ) на ПТ и ЭП (рис. 8.2) в широком диапазоне входных фронтов и токов переключения для рассмотренных эффективных распределенных моделей с двумя и тремя секциями по сравнению с

близким к идеальному расчету на основе многосекционной ($n=8\div 10$) модели БТ. Применение упрощенных моделей БТ с одним сосредоточенным сопротивлением базы $R_B=R/3$ при крутых фронтах ($t_{ф.вх} \leq 100$ пс) и больших токах переключения $I_0 > 1$ мА дает максимальную погрешность (80—100%); при $I_0 = 0.1\div 0.5$ мА, $t_{ф.вх} = 100\div 500$ пс погрешность для моделей БТ с $R_B=R/3$ составляет 30—60%. Проведенный обширный анализ переключательных характеристик ПТ и ЭСЛ-элементов в широком диапазоне потребляемых мощностей $10^{-6} \leq p \leq 10^{-2}$ Вт и фронтов входных сигналов 10^{-12} — 10^{-9} с позволяет заключить, что погрешность эффективной двухсекционной распределенной модели не превышает 10—15% (см. рис. 6.34, 8.3, а, б).

К практическим аспектам уменьшения нежелательных затрат на многовариантные задачи расчета и оптимизации ЛЭ БИС с обеспечением компромисса приемлемая погрешность расчета — сложность модели БТ следует отнести использование в расчетах односекционной электрической модели БТ с эффективным базовым сопротивлением (см. рис. 6.33, 8.3, а, б). С помощью подстроечного параметра — эффективного базового сопротивления в односекционной модели БТ — можно значительно (до 10—20%) уменьшить погрешность расчета ЛЭ при дополни-

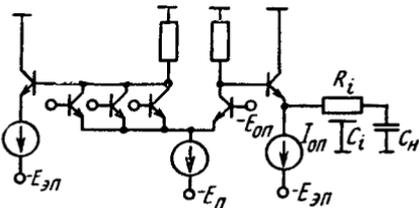


Рис. 8.2. Схема ЭСЛ-элемента, работающего на линию связи и емкость нагрузки в системе С³БИС

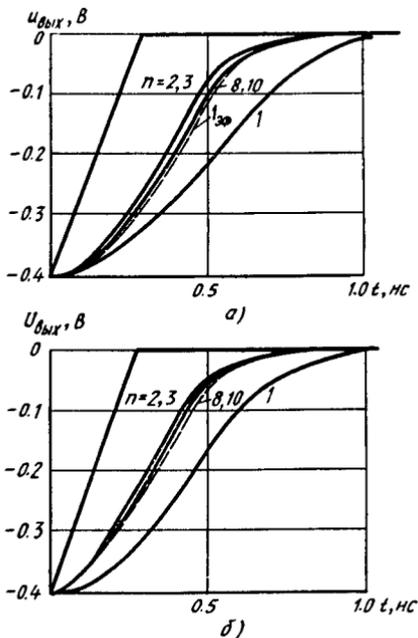


Рис. 8.3. Расчетные и экспериментальные зависимости переходного процесса в переключателе тока с помощью электрических моделей БТ, имеющих различное количество секций (пунктир соответствует расчету с одной эффективной секцией):

а— $t_{ф.вх}=0.3$ нс, $I_0=1$ мА; б— $t_{ф.вх}=0.3$ нс, $I_0=0.3$ мА

тельном сокращении вычислительных затрат. При этом для малых токов переключения $I \leq 0.1$ мА в модели БТ правомочно использование постоянного базового сопротивления $R_B \approx R/3$ (погрешность расчетов 10%); при средних токах переключения $0,3 \leq I_0 \leq 1$ мА в случае необходимости правомочно применение эффективного базового сопротивления $R/6 \leq R_{B\text{эф}} \leq R/3$ без базовой емкости с погрешностью $\leq 25\%$. Отметим, что выбор неадекватных моделей БТ со значительной погрешностью (30—50%) практически обесценивает результаты многовариантных расчетов и оптимизации БТ и ЛЭ.

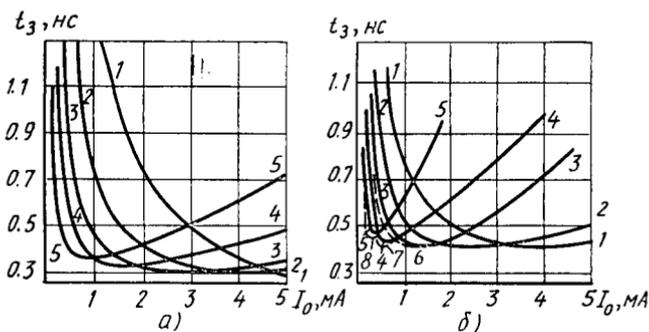


Рис. 8.4. Расчетные зависимости задержки ЭСЛ-элементов (в цепочке с коэффициентом разветвления, равным двум) от тока переключения:

а — для БТ с двумя базовыми выводами (1 — $l/l_N=4$; 2 — $l/l_N=2$; 3 — $l/l_N=1$; 4 — $l/l_N=0.5$; 5 — $l/l_N=0.25$); б — для БТ с одним базовым выводом (1—5 аналогичны рис. а; 6 — $l/l_N=1$, $C_{jc}=0.06$ пФ; 7 — $l/l_N=0.5$, $C_{jc}=0.028$ пФ; 8 — $l/l_N=0.25$, $C_{jc}=0.014$ пФ)

Рассмотрим возможности адаптации структурных и технологических параметров БТ с целью совместной схмотехнической и технологической оптимизации быстродействия субнаносекундных ЛЭ БИС. За основу возьмем рассчитанные с помощью программы схмотехнического моделирования переключательные характеристики ЛЭ (рис. 8.4, а, б), где в моделях БТ для уровня технологии изготовления с минимальным размером 2,5 мкм были использованы электрические параметры БТ в конструкциях с одним и двумя базовыми выводами $n_B = 1,2$ при $h, d = \text{const}$, а также при различных вариациях длины эмиттера $l/l_N = 0.25 \div 4$ ($l_N = 7$ мкм). Электрические параметры БТ получены с помощью выделения топологически инвариантных параметров и дальнейшего конструктивного расчета сопротивлений, емкостей, токов согласно элементарным правилам: $R_i = R_{i0} l_N / l$, $C_i \neq C_{i0} l / l_N$, $I_i = I_{i0} l / l_N$.

Для оценки степени влияния основных параметров модели БТ на ЛЭ применим рассчитанные на ЭВМ факторы чувствительности $s(t_3)$ к данным параметрам для БТ с $n_B = 1$ (табл. 8.2). Факторы чувствительности s -параметра БТ определялись как процентное изменение t_3 ЛЭ при изменении параметра БТ на 1% и неиз-

меньших значениях остальных параметров. С уменьшением тока I_0 уменьшается влияние на быстродействие ЛЭ параметров $R_{БЛ}$, $\tau_{ТН}$ и увеличивается влияние параметров C_K , C_n , I_0 .

Анализ переключаемых характеристик ЭСЛ-элементов позволяет заключить, что использование БТ с $n_B = 2$ дает максимальное быстродействие в области больших токов переключения

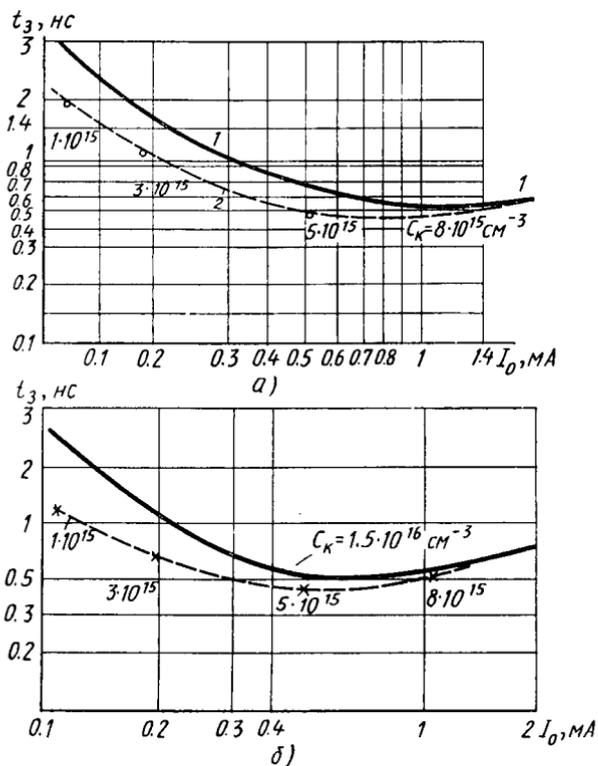


Рис. 8.5. Расчетные зависимости времени задержки ЭСЛ-элемента от тока переключения при нормированной длине эмиттера $l/l_N = 1$ (а) и $l/l_N = 0.5$ (б):
 1 — при постоянной емкости коллектора $C_{jс}$ ($C_K \approx 2 \cdot 10^{16} \text{ см}^{-3}$);
 2 — при переменной емкости коллектора $C_{jс}(C_K, I_0)$

и потребляемых мощностей ($I_0 > 1 \text{ мА}$, $p \geq 3 \div 5 \text{ мВт}$) с потерей в степени интеграции БИС. Требованию обеспечения достаточно высокого быстродействия и высокой степени интеграции ($p \leq 0,5 \div 1 \text{ мВт}$) для рассмотренной базовой технологии удовлетворяют только БТ с $n_B = 1$, имеющие минимальные площади переходов и соответственно барьерные емкости эмиттера, коллектора, подложки.

На рис. 8.5, а, б приведены расчетные зависимости $t_3(I_0)$ при $l/l_N=0.5$; 1.0 для двух случаев: 1) постоянной коллекторной емкости во всем диапазоне I_0 , $C_{jc}=\text{const}$, определяемой неизменным профилем легирования в коллекторе ($C_K \approx 2 \cdot 10^{16} \text{ см}^{-3}$);

Т а б л и ц а 8.2

Чувствительность времени задержки к параметрам БТ

Параметры	s, %	
	$l=l_N/2=7 \text{ мкм}$ $I_0=1,5 \text{ мА}$	$l=l_N/2=3,5 \text{ мкм}$ $I_0=0,2 \text{ мА}$
R_{BA}	0.37	0.29
$\tau_{тN}$	0.32	0.17
C_K	0.24	0.33
I_0	0.07	0.34
$C_{п}$	0.17	0.17

2) переменной коллекторной емкости $C_{jc}(C_K, I_0)$. Уменьшение емкости C_K при малых токах I_0 достигалось уменьшением концентрации C_K и увеличением толщины W_K эпитаксиального слоя (до n^+ -области коллектора), причем величина C_K выбиралась из условия отсутствия для данной плотности тока I_0/hl эффектов расширения

базы и квазинейтральности БТ таким образом, что $C_{K \text{ min}} \geq (I_0/hl) qv_s$.

Расчет коллекторных емкостей осуществлялся с помощью ФТ-программ ТРАН 1, ТРАН 2. Как видно из рис. 8.5, уменьшение концентрации C_K (и соответственно емкости C_{jc}) с убыванием тока I_0 приводит к некоторому уменьшению оптимальных токов переключения $I_{оп}$ и значительному снижению t_3 в области малых токов. Доминиру-

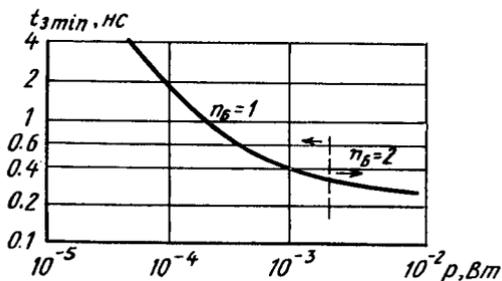


Рис. 8.6. Расчетная зависимость минимального времени задержки от потребляемой мощности для ЭСЛ-элемента на БТ с одним и двумя базовыми выводами и различной концентрацией примеси в коллекторе

рующий вклад коллекторных емкостей в t_3 ЛЭ при малых I_0 , как следует из рассмотренных графиков и результатов расчета чувствительностей параметров БТ (см. табл. 8.2), может быть учтен разработчиками БИС для синтеза БТ с оптимальными топологией и профилем распределения примесей. Исходя из ограничений на параметры топологии в рамках базового технологического маршрута $h_{\text{min}}=2.5 \text{ мкм}$, $l_{\text{min}}=3.5 \text{ мкм}$ и технологии — коллекторного профиля легирования $C_{K \text{ min}}=3 \cdot 10^{15} \text{ см}^{-3}$ базового процесса БТ — на основании расчетных зависимостей $t_3(I_0)$ для БТ с $n_B=1,2$, $l/l_N=1/4-4$ можно построить фундаментальную харак-

теристику $t_{3\min}(p)$ для конкретной базовой технологии (рис. 8.6) и в обобщенном виде (рис. 8.7). При достаточно больших потребляемых мощностях ($p > 3$ мВт) минимальным значениям $t_{3\min}$ соответствуют конструкции БТ с $n_B = 2$, $h_{\min} = 2.5$ мкм, $l = 7 \div 20$ мкм. При малых потребляемых мощностях ($0.8 \leq p \leq 2$ мВт) значения $t_{3\min}$ достигаются для БТ с $n_B = 1$ при уменьшении l/l_N и номинальных C_K . Для некоторой потребляемой мощности $p_1 \approx 1 \div 2$ мВт значения $t_{3\min}$ для двух БТ с $n_B = 1$ и 2 сравниваются. Если для БТ $n_B = 1$ достигнуты все ограничения топологии: $h_{\min} = 2.5$ мкм, $l_{\min} = 3.5$ мкм, т. е. невозможно влиять геометрически на сопротивление базы

$R_{БЛ}$, $R_{БП}$, то при дальнейшем снижении мощности p уменьшать доминирующую постоянную времени $\frac{U_n}{I_0} C_K$

коллекторного узла можно только за счет уменьшения уровня легирования C_K . При $C_K = C_{K\min}$ (а также h_{\min} , l_{\min}) некоторой потребляемой мощности $p_2 \approx 0.2 \div 0.3$ мВт выполняется соотношение $t_3 \sim 1/p$. Таким образом,

вследствие рассмотренной адаптации структурных (h , l , n_B) и технологических (C_K) параметров к потребляемой мощности ЛЭ в рамках базового технологического процесса достигаются минимальные времена задержки $t_{3\min} = 0.35 \div 0.9$ нс при $0.5 \leq p \leq 10$ мВт. Неформальный подход с использованием многовариантных схемотехнических расчетов и ФТ-моделирования подтверждает ощутимую пользу схемно-технологической оптимизации интегральных элементов и схем на основе самосогласования параметров профиля распределения примесей, геометрии и электрического режима БИС.

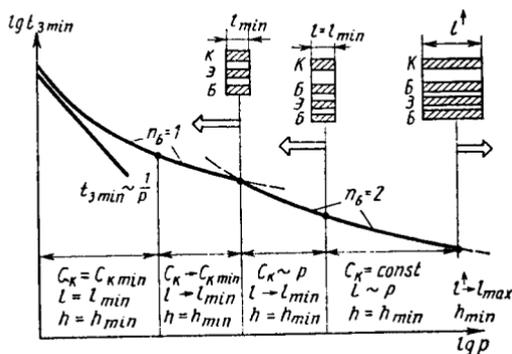


Рис. 8.7. Обобщенная кривая адаптации БТ логического элемента для обеспечения минимального времени задержки при заданной потребляемой мощности элемента

8.2. ВЫБОР ЭЛЕКТРИЧЕСКОГО РЕЖИМА И ОПТИМИЗАЦИЯ ЭЛЕМЕНТОВ БИС

Выражение для t_3 ЭСЛ-элемента, выведенное с помощью аппарата чувствительности ЭВМ, может быть записано в виде

$$t_3 \approx \frac{U_n}{I_0} \sum_{i=1}^4 k_{1i} C_i + (k_{21} R_{БЛ} + k_{22} R_{БП}) \frac{2\tau_N I_0}{U_n} + k_3 \tau_{TN} + k_4 R_{Б} C_{КА} + k_{лс} \tau_{лс}, \quad (8.1)$$

где $U_{л}, I_0$ — логический перепад и ток переключения;

$$\sum_{i=1}^4 k_{li}C_i = k_{11}C_{jc} + k_{12}C_{п} + k_{13}C_{je} + k_{14}C_{лс}; \quad C_{jc}, C_{п}, C_{je}$$

— барьерные емкости коллектора, подложки, эмиттера при нулевом смещении; $C_{лс}$ — эффективная емкость линии связи (ЛС); $k_{1i}, k_{21}, k_{22}, k_4, k_{лс}$ — коэффициенты чувствительности ЛЭ; $R_{БА}, R_{БП}$ — активное и пассивное сопротивления базы БТ; $k_2R_B = k_{21}R_{БА} + k_{22}R_{БП}$; $\tau_{ЭК} \approx \tau_{TN}$ — время задержки эмиттер — коллектор БТ, $C_{КА}$ — коллекторная емкость активного БТ (под дном эмиттера); $\tau_{лс}$ — постоянная времени распределенной ЛС: $\tau_{лс} \approx R_{лс}C_{лс}$ (см. рис. 8.2).

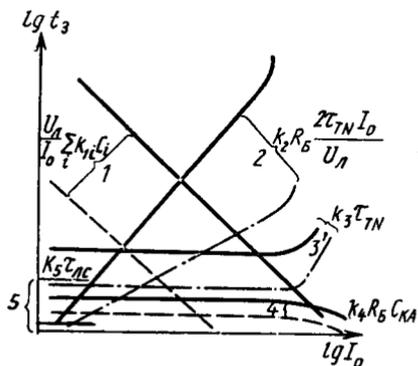


Рис. 8.8. Иллюстрация влияния уменьшения горизонтальных и вертикальных размеров БТ на уменьшение составляющих времени задержки логического элемента C^3 БИС

Масштабирование БТ. На рис. 8.8 представлены асимптотические зависимости от режимного тока I_0 основных составляющих t_3 ЛЭ в соотношении (8.1). Составляющая времени задержки (1 на рис. 8.8) уменьшается с ростом тока и с уменьшением емкостей $C_j, C_{п}, C_{лс}$ вследствие уменьшения h и минимизации пассивных областей БТ при использовании новых методов изоляции схем (в частности, глубокими канавками, заполненными

окислом). Влияние уменьшения минимального литографического размера сказывается также на уменьшении составляющей времени задержки $k_4R_{БА}C_{КА}$ (4 на рис. 8.8), которая при постоянных толщине и уровне легирования базы примерно пропорциональна квадрату ширины эмиттера БТ. Уменьшение минимального литографического размера с сопутствующими ему технологиями горизонтального сжатия БТ и ЛС, приводящие к уменьшению составляющих t_3 ЛЭ, показано на рис. 8.8 пунктиром.

Составляющая времени задержки $(k_{21}R_{БА} + k_{22}R_{БП}) \frac{2\tau_{TN}I_0}{U_{л}}$ (2 на рис. 8.8) увеличивается с ростом рабочего тока ЛЭ. Уменьшение минимального литографического размера влияет на изменение данной составляющей с ростом тока, однако это влияние значительно слабее против эффекта вертикального сжатия БТ с соответствующим сильным легированием базовых областей для предотвращения прокола БТ и минимизации базовых сопротивлений. Процесс вертикального сжатия БТ на рис. 8.8 показан штрихпунктирной линией для составляющих времен задержки $k_3\tau_{TN}$ и $k_2R_B \frac{2\tau_{TN}I_0}{U_{л}}$ (2 и 3 на рис. 8.8).

При небольших плотностях тока $R_{\text{БА}}$, $\tau_{\text{ТН}}$ слабо изменяются с ростом тока, при больших плотностях тока, когда проявляются эффекты высокого уровня инжекции, $R_{\text{БА}}$ уменьшается, а $\tau_{\text{ТН}}$ — возрастает. С ростом рабочего тока резко возрастает составляющая $k_2 R_{\text{Б}} \frac{2\tau_{\text{ТН}} I_0}{U_{\text{л}}}$ и уменьшается составляющая $k_4 R_{\text{БА}} C_{\text{КА}}$. Однако

при небольших рабочих токах $k_2 R_{\text{Б}} \frac{2\tau_{\text{ТН}} I_0}{U_{\text{л}}} \ll \frac{U_{\text{л}}}{I_0} \sum_i k_{1i} C_i$ и цепь

базовых сопротивлений оказывает слабое влияние на t_3 ЛЭ.

С уменьшением рабочих токов вследствие возрастания емкостной составляющей $\frac{U_{\text{л}}}{I_0} \sum_i k_{1i} C_i$ наблюдается значительная потеря

быстродействия ЛЭ С³БИС. Рис. 8.8 иллюстрирует взаимовлияние процессов уменьшения горизонтальных и вертикальных размеров БТ на изменение минимального времени задержки ЛЭ и оптимального тока переключения в зависимости от доминирования составляющих $k_2 R_{\text{Б}} \frac{2\tau_{\text{ТН}} I_0}{U_{\text{л}}}$ или $k_3 \tau_{\text{ТН}}$, поскольку обычно для БТ с

микротопологией выполняется условие $k_3 \tau_{\text{ТН}}, k_2 R_{\text{Б}} \frac{2\tau_{\text{ТН}} I_0}{U_{\text{л}}} >$

$> k_4 R_{\text{БА}} C_{\text{КА}}$.

Для тонкослойных БТ увеличение уровня легирования активной базовой области, необходимое для предотвращения прокола базы, значительно уменьшает $R_{\text{СБА}} (R_{\text{БА}})$; методом самосовмещения или сверхсамосовмещения при сильном легировании пассивной области базы значительно уменьшают $R_{\text{БП}}$ (или даже устраняют данную область в перспективных технологиях с боковыми поликремниевыми электродами, в так называемом симметричном БТ). Вследствие рассмотренных аспектов перспективных технологических методов при совместном уменьшении горизонтальных и вертикальных размеров БТ составляющие времени задержки, обусловленные сопротивлениями базы, минимизируются, а в перспективных С³БИС (особенно для невысоких рабочих токов), возможно, будут практически утрачивать свое значение. Тогда при

выполнении условий $k_3 \tau_{\text{ТН}} > k_2 R_{\text{Б}} \frac{2\tau_{\text{ТН}} I_0}{U_{\text{л}}}$ и $k_3 \tau_{\text{ТН}} > k_5 \tau_{\text{ЛС}}$ минимальное время задержки

$$t_{3 \text{ мин}} \approx 2k_3 \tau_{\text{ТН}}$$

(8.2)
285

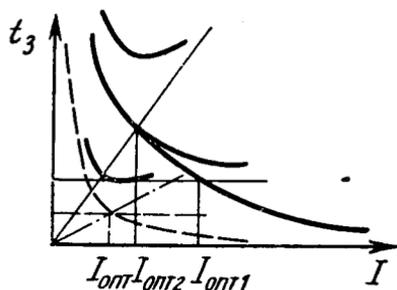


Рис. 8.9. Типы оптимальных по быстродействию токов логического элемента

достигается при оптимальном токе $I_{\text{опт}1}$ (рис. 8.9):

$$I_{\text{опт}1} = \sum_{i=1}^4 k_{1i} C_i U_n / k_3 \tau_{TN}. \quad (8.3)$$

Если же составляющие времени задержки, связанные с сопротивлением базы, доминируют в t_3 , как это имеет место для большинства ЛЭ современных СЗБИС, т. е. $k_2 R_B \frac{2\tau_{TN} I_0}{U_n} > k_3 \tau_{TN}$, $k_4 R_B C_{КА}$, то оптимальный ток $I_{\text{опт}2}$ смещается в сторону меньших значений, а минимальное время задержки $t_{3\text{min}2}$ значительно увеличивается (рис. 8.9). Оптимальный ток $I_{\text{опт}2}$ и минимальное время задержки $t_{3\text{min}2}$ определяют из условия

$$\frac{2k_2 R_B \tau_{TN} I_0}{U_n} = \frac{U_n}{I_0} \sum_{i=1}^4 k_{2i} C_i \left(\frac{\partial t_3}{\partial I_0} = 0 \right):$$

$$I_{\text{опт}2} = U_n \left(\sum_{i=1}^4 \frac{k_{1i} C_i}{2\tau_{TN} (k_{21} R_{БА} + k_{22} R_{БП})} \right)^{1/2}, \quad (8.4)$$

$$t_{3\text{min}2} = 2 [2\tau_{TN} (k_{22} R_{БА} + k_{22} R_{БП}) \left(\sum_{i=1}^4 k_{1i} C_i \right)]. \quad (8.5)$$

С точки зрения достижения максимального быстродействия и минимальной потребляемой мощности ЛЭ оптимальным следует признать электрический режим, когда для полученной совокупности параметров элемента СЗБИС вклады трех доминирующих задержек в полное время задержки практически равны:

$$(k_{21} R_{БА} + k_{22} R_{БП}) \frac{2\tau_{TN} I_0}{U_n} = \sum_{i=1}^4 k_{1i} C_i \frac{U_n}{I_0} = k_3 \tau_{TN}. \quad (8.6)$$

В этом случае обобщенный оптимальный ток

$$I_{\text{опт}} = \frac{\sum_i k_{1i} C_i}{k_3 \tau_{TN}} U_n = U_n \left[\frac{\sum_i k_{1i} C_i}{2(k_{21} R_{БА} + k_{22} R_{БП}) \tau_{TN}} \right]^{1/2} \quad (8.7)$$

при факторе γ , равном единице:

$$\gamma = \left[k_3 \tau_{TN} / 2 \left(\sum_i k_{1i} C_i \right) (k_{21} R_{БА} + k_{22} R_{БП}) \right]^{1/2} = 1, \quad (8.7a)$$

а минимальное время задержки

$$t_{3\text{min}} = 3k_3 \tau_{TN} = 3 \left[2 \left(\sum_i k_{1i} C_i \right) (k_{21} R_{БА} + k_{22} R_{БП}) \tau_{TN} \right]^{1/2}. \quad (8.8)$$

Графической иллюстрацией соотношений между токами $I_{\text{опт}1}$, $I_{\text{опт}2}$, $I_{\text{опт}}$ служит рис. 8.9, где при $I_{\text{опт}1} < I_{\text{опт}2}$ $\gamma > 1$, при $I_{\text{опт}2} < I_{\text{опт}1}$ $\gamma < 1$, при $I_{\text{опт}1} = I_{\text{опт}2} = I_{\text{опт}}$ $\gamma = 1$.

Согласно выражению (8.8), для уменьшения $t_{3\text{min}}$ требуется совместное уменьшение всех основных параметров БТ: C_i , $R_{\text{БА}}$, $R_{\text{БП}}$, τ_{TN} ; с другой стороны, для уменьшения $I_{\text{опт}}$ достижения требования увеличения быстродействия и интеграции С³БИС, согласно выражению (8.7), необходимо, чтобы уменьшение $\sum_i k_{1i}C_i$ происходило сильнее, чем $k_2R_{\text{Б}}\tau_{\text{TN}}$, т. е. чтобы темп «горизонтального сжатия» БТ превышал темп «вертикального сжатия» в рамках развивающихся перспективных технологий изготовления С³БИС.

Выражения для оптимальных энергии переключения $E = pt_3$ и фактора действия $D = pt_3^2$, характеризующих эффективность технологической и схемной реализации ЛЭ С³БИС, получим из (8.7), (8.8):

$$E_{\text{опт}} \approx U_{\text{и.п.}} I_{\text{опт}} t_{3\text{min}} \approx U_{\text{и.п.}} U_{\text{л}} \left(\sum_i k_{1i} C_i \right), \quad (8.9)$$

$$D_{\text{опт}} \approx U_{\text{и.п.}} U_{\text{л}} k_3 \tau_{\text{TN}} \left(\sum_i k_{1i} C_i \right) \approx \\ \approx 3U_{\text{и.п.}} U_{\text{л}} \left(\sum_i k_{1i} C_i \right)^{3/2} [(k_{21} R_{\text{БА}} + k_{22} R_{\text{БП}}) \tau_{\text{TN}}]^{1/2}, \quad (8.10)$$

где $U_{\text{и.п.}}$ — напряжение питания ЛЭ.

Оптимальные токи, определяемые по задержке $I_{\text{опт}}(t_3)$ и фактору действия $I_{\text{опт}}(D)$ для доминирующего влияния $C_{jc} \left(\sum_i k_{1i} C_i \approx \approx k_1 C_{jc} \right)$, можно записать в виде

$$I_{\text{опт}}(t_3) \approx U_{\text{л}} \left(\frac{C_{jc}}{R_{\text{Б}} \tau_{\text{TN}}} \right)^{1/2} \left(\frac{k_1}{2k_2} \right)^{1/2}, \quad (8.11)$$

$$I_{\text{опт}}(D) \approx U_{\text{л}} \left(\frac{C_{jc}}{R_{\text{Б}} \tau_{\text{TN}}} \right)^{1/2} \left(\frac{k_1}{6k_2} \right)^{1/2}. \quad (8.12)$$

При оптимальных токах $I_{\text{опт}}(t_3)$ и $I_{\text{опт}}(D)$ минимальное время задержки

$$t_3(I_{\text{опт}}(t_3)) \approx 2.8 (R_{\text{Б}} \tau_{\text{TN}} C_{jc} k_1 k_2)^{1/2}, \quad (8.13)$$

$$t_3(I_{\text{опт}}(D)) \approx 3.2 (R_{\text{Б}} \tau_{\text{TN}} C_{jc} k_1 k_2)^{1/2}. \quad (8.14)$$

Из выражений (8.11)—(8.14) следуют сравнительные соотношения

$$I_{\text{опт}}(t_3) = 1.73 I_{\text{опт}}(D), \quad t_3(I_{\text{опт}}(D)) = 1.15 t_3(I_{\text{опт}}(t_3)), \\ D(I_{\text{опт}}(t_3)) = 1.3 D(I_{\text{опт}}(D)); \quad t_3, \quad pt_3, \quad D \sim (R_{\text{Б}} \tau_{\text{TN}})^{1/2}, \\ t_3 \sim C_{jc}^{1/2}, \quad pt_3 \sim C_{jc}, \quad D \sim C_{jc}^{3/2}. \quad (8.15)$$

На рис. 8.10 в качестве иллюстрации рассмотренных аналитических соотношений приведены типовые зависимости времени задержки, энергии переключения и фактора действия ЭСЛ-элементов от тока переключения и логического перепада, рассчитанные с помощью программы схемотехнического моделирования на ЭВМ.

На основании машинных и аналитических расчетов для t_3 , pt_3 , D можно сделать следующие выводы:

1) минимизация t_3 в ряде случаев достижима для больших токов переключения. В то же время минимальные значения энергии переключения pt_3 достижимы только в области малых токов потребляемой мощности. Данные условия являются достаточно противоречивыми при определении критерия качества сверхбыстродействующих схем;

2) фактор действия $D = pt_3^{-2}$ является обобщенным ресурсным показателем, характеризующим эффективность технологической и схемной реализации сверхбыстродействующих ЛЭ в БИС. Оптимизация фактора действия по току переключения позволяет для заданного уровня логического перепада (определяемого требуемой помехоустойчивостью ЛЭ) найти минимальный ток $I_{\text{опт}}(D)$, для

которого как D , так и pt_3 будут минимальными. Отметим, что наблюдается устойчивый резкий экстремум функции $D(p)$ в отличие от $pt_3(p)$;

3) исходя из технического задания по требуемому быстродействию, потребляемой мощности и запасу помехоустойчивости оптимизация переключательных характеристик (по фактору действия или времени задержки) позволяет выбрать диапазон логического перепада для заданных диапазонов t_3 , p , т. е. оптимальный электрический режим ЛЭ и БИС.

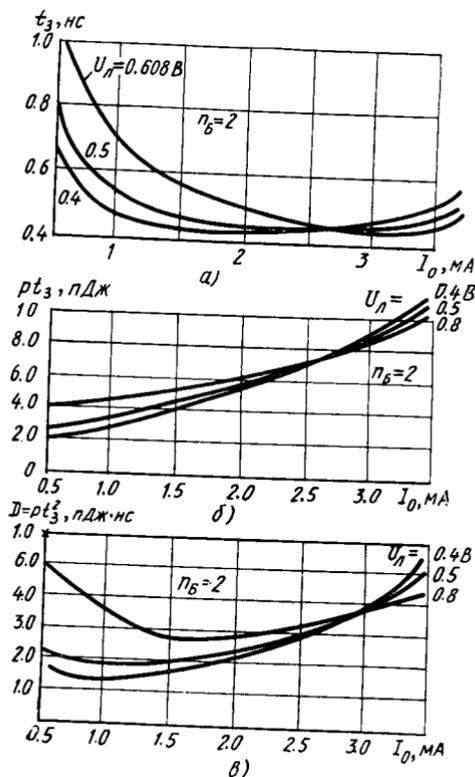


Рис. 8.10. Зависимости времени задержки (а), энергии переключения (б) и фактора действия (в) от тока переключения ЭСЛ-элемента при различных логических перепадах

Таким образом, задачей выбора оптимальных параметров БТ БИС является достижение минимальных времен задержки t_3 ЛЭ для минимальных токов переключения. Для этого схемными средствами надо уменьшать логический перепад U_{π} (и напряжение питания $U_{\text{п.п}}$), а также технологическими средствами значительно снижать горизонтальные размеры БТ и резисторов, что приводит к уменьшению емкостей C_{je} , C_{jc} , $C_{\text{п}}$. Затем, задавшись некоторыми ограничениями на параметры $R_{\text{Б}}$, $\tau_{\text{тн}}$, V (например, $V \geq \geq 50 \div 100$), необходимо минимизировать целевые функции $(R_{\text{Б}} \tau_{\text{тн}} C_{jc}^3)^{1/2}$, $(R_{\text{Б}} \tau_{\text{тн}} C_{jc})^{1/2}$ за счет вариаций вертикального профиля и геометрических горизонтальных размеров БТ.

Достаточно корректной формой оптимизации быстродействия ненасыщенных ЛЭ, как следует из машинных и аналитических расчетов, является выбор режима $I_0 \approx I_{\text{опт}}(D)$, например $I_{\text{опт}}(D) \leq \leq I_0 \leq I_{\text{опт}}(t_3)$, при этом $t_3(I_{\text{опт}}(D))$ не будет заметно больше $t_3(I_{\text{опт}}(t_3))$, в то же время будет ощутим выигрыш потребляемой мощности ЛЭ, что актуально для БИС высокой степени интеграции. С другой стороны, значение $t_3(I_0)$, взятое из предложенного диапазона токов, будет значительно меньше времен задержки, удовлетворяющих условию $p t_3 = \text{const}$ для малых p .

8.3. ОСНОВЫ ФИЗИКО-ТЕХНОЛОГИЧЕСКОГО СИНТЕЗА СУБМИКРОННЫХ ТОНКОСЛОЙНЫХ ЭЛЕМЕНТОВ БИПОЛЯРНЫХ С³БИС

Распределение примесей и конструкции оптимальных БТ С³БИС необходимо синтезировать таким образом, чтобы обойти большое количество физических пределов, ограничений и нежелательных электрофизических эффектов, возникающих при прогрессирующем уменьшении геометрических размеров компонентов С³БИС. При этом требуется использование ряда физических и технологических принципов создания высококачественных, тонкослойных, сильнолегированных БТ для: уменьшения толщины запирающих слоев; предотвращения прокола базы и уменьшения сопротивлений базовых областей БТ; предотвращения эффектов высокого уровня инжекции, в первую очередь расширения базы и квазинасыщения; обеспечения максимально высоких усиления и быстродействия БТ. Для осуществления рекомендаций и принципов, изложенных выше, нужно принимать меры для: повышения пробивного напряжения коллекторного перехода; увеличения пробивного напряжения эмиттерного перехода и минимизации туннелирования (что характерно для сильнолегированных переходов) и утечек в БТ.

На рис. 8.11 представлена расчетная зависимость максимальной концентрации примеси в базе $C_{\text{Б max}}$ от толщины базы $W_{\text{Б}}$ для БТ, показывающая, что при уменьшении $W_{\text{Б}}$ концентрация $C_{\text{Б max}}$ должна быть увеличена для поддержания одинакового напряжения прокола базы (≈ 5 В).

На основании расчетных данных увеличение концентрации легирующей примеси $5 \cdot 10^{17} - 3 \cdot 10^{18} \text{ см}^{-3}$ в базе БТ с $W_B = 0.10 \text{ мкм}$ является необходимой гарантией обеспечения приемлемых напряжений прокола базы. С учетом флуктуации базовой примеси ΔC_B , заданной соотношением $\Delta C_B \approx (2 \ln N_1 N_2 C_B / W_B^3)^{1/2}$, где N_1 — количество единиц объема W_B^3 в базовой области; N_2 — количество БТ в кристалле, рассчитывают величину $C_{B \max}(W_B)$, причем достаточно высокие напряжения прокола поддерживаются в БТ за счет увеличения $C_{B \max} \sim W_B^{-2}$ при уменьшении W_B .

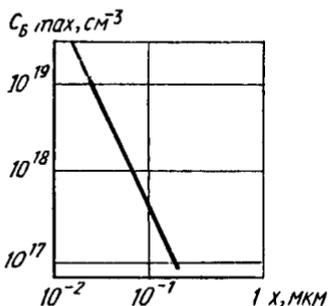


Рис. 8.11. Зависимость максимальной концентрации примеси в активной базе БТ от толщины базы для определения ограничения по напряжению прокола базы

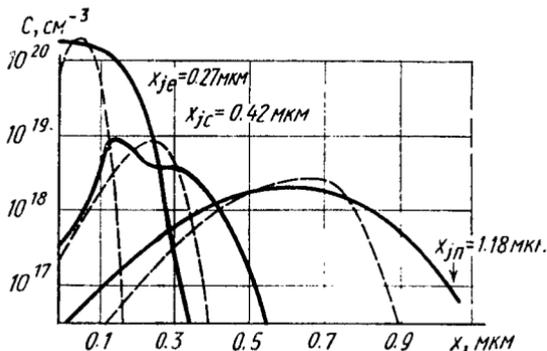


Рис. 8.12. Профиль распределения примесей тонкослойного сильнолегированного БТ с ионно-легированным коллектором

При концентрации легирующей примеси в базе более 10^{19} см^{-3} начинается эффект туннелирования носителей заряда между базой и эмиттером, что определяет фундаментальную ширину эмиттерного перехода — 15 нм. Поскольку прокол базы для БТ с эпитаксиальным коллектором происходит со стороны эмиттера, из-за туннелирования предельная ширина базы составляет 30 нм.

Так как достаточно высокие напряжения прокола тонкослойных БТ С³БИС с уменьшением толщины базы W_B должны обеспечиваться значительным увеличением концентрации акцепторной примеси $C_{B \max} \sim W_B^{-2}$, имеется реальная возможность ухудшения основных параметров БТ, определяющих качество разработки ЛЭ С³БИС: увеличение τ_{TN} из-за уменьшения коэффициента диффузии D_n ($\tau_{TN} \sim W_B^2 / D_n$) и спад коэффициента B_0 вследствие уменьшения эффективности эмиттера при росте дырочной составляющей тока из базы в эмиттер с учетом эффектов сильного легирования.

Рассмотрим принципиально важный механизм некоторой компенсации ухудшения параметров B и τ_{TN} , который следует иметь в виду при разработке перспективных сильнолегированных тонко-

слоистых БТ субмикронных С³БИС. При $C_B > 5 \cdot 10^{17} \text{ см}^{-3}$ в условиях сильного легирования происходит сужение ширины запрещенной зоны полупроводника в базе аналогично проявлению данного эффекта в сильнолегированном эмиттере, приводящего к сужению запрещенной зоны в эмиттере. Вследствие этого для тонкослойных высоколегированных БТ С³БИС можно ожидать не только интенсивной инжекции дырок из базы в эмиттер, но и дополнительного увеличения инжекции электронов из эмиттера в базу, что приводит к компенсации нежелательного изменения параметров.

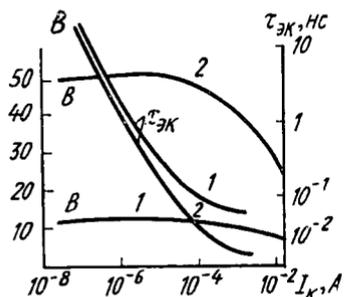


Рис. 8.13. Расчетные зависимости коэффициента передачи тока и времени задержки эмиттер — коллектор БТ от тока коллектора:

1 — сильное легирование в ФТ-модели учитывается только в эмиттере; 2 — сильное легирование учитывается одновременно в эмиттере и базе

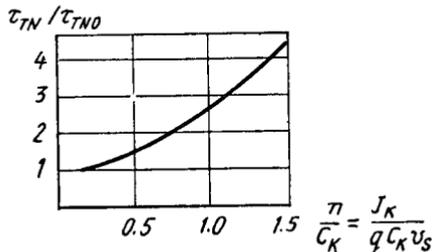


Рис. 8.14. Зависимость нормированного времени задержки эмиттер — коллектор БТ от нормированной плотности тока

Рассмотрим данный механизм компенсации на примере профиля распределения примеси в сильнолегированном тонкослойном БТ с ионно-легированным коллектором и мышьяковистым эмиттером (рис. 8.12). Профиль распределения примесей получен моделированием технологического маршрута изготовления тонкослойного субмикронного БТ с помощью Т-модели, включающего: 1) коллекторную имплантацию мышьяка с дозой $D_1 = 3 \cdot 10^{14} \text{ см}^{-2}$, энергией $E_1 = 900 \text{ кэВ}$; 2) базовую имплантацию бора с дозой $D_2 = 10^{14} \text{ см}^{-2}$, энергией $E_2 = 100 \text{ кэВ}$; 3) эмиттерную имплантацию мышьяка с дозой $D_3 = 3,5 \cdot 10^{16} \text{ см}^{-2}$, энергией $E_3 = 25 \text{ кэВ}$; 4) отжиг при температуре 950°C за время $t = 7 \text{ мкм}$. На рис. 8.12 пунктиром изображен профиль распределения примесей трех имплантаций до отжига, сплошными линиями — после отжига.

Отличительной особенностью данного профиля являются не только малые глубины залегания $x_{jc} = 0,27 \text{ мкм}$, $x_{jc} = 0,42 \text{ мкм}$ при толщине базы $W_B \sim 0,1 \text{ мкм}$, но и малая глубина залегания перехода коллектор — подложка $x_{jn} = 1,18 \text{ мкм}$.

На рис. 8.13 приведены рассчитанные с помощью программы ФТ-моделирования зависимости $B(I_K)$, $\tau_{ЭК}(I_K)$ для двух модель-

ных предположений: 1) эффекты сильного легирования учитываются только в эмиттере (кривые 1); 2) эффекты учитываются и в эмиттере, и в базе (кривые 2), что соответствует более реальному рассмотрению физических процессов в БТ с высоким легированием эмиттерной и базовой областей. Увеличение коэффициента B примерно в пять раз и уменьшение времени задержки τ_{TN} в два раза при токах $I_K = 0.1 \div 0.5$ мА второй модели являются позитивным подтверждением гипотезы о компенсации нежелательного ухудшения параметров B , τ_{TN} при создании тонкслоистых субмикронных БТ С³БИС.

Скоординированное сжатие горизонтальных и вертикальных размеров БТ С³БИС предъявляет серьезные требования к вопросам синтеза вертикального профиля БТ в целях не только физической реализации работоспособного прибора, но и эффективного режимного функционирования БТ с предотвращением влияния эффектов высокого уровня инжекции.

Концентрация примеси в коллекторе в целях предотвращения эффекта расширения базы при высоких плотностях тока $J > (1 \div 2) \cdot 10^4$ А/см² для БТ с микротопологией также должна быть увеличена до $C_K > 10^{16} \div 10^{17}$ см, причем увеличение C_K , как было выяснено из ФТ-расчетов, должно быть пропорционально плотности J_K во избежание деградации усиления и быстродействия БТ С³БИС. Данная рекомендация проиллюстрирована расчетной зависимостью рис. 8.14 по ФТ-программе: $\tau_{TN}/\tau_{TN0}(n/C_K)$, где $\tau_{TN0} = \tau_{TN}(J_K \rightarrow \infty)$, без учета эффектов высокого уровня инжекции); n — концентрация электронов в переходе коллектор — база; C_K — переменная величина легирующей примеси в эпитаксиальном коллекторе. В БТ с низкой концентрацией C_K при $n > C_K$ происходит расширение обедненной области коллекторного перехода (и базы) в сторону скрытого n^+ -слоя, причем изменения $\tau_{эк}/\tau_{TN0}$ наблюдаются уже при $n/C_K \approx 0.25$.

Рассмотрим способы конструирования субмикронных БТ с предотвращением эффектов высокого уровня инжекции, приводящих к ухудшению его быстродействия и усиления.

На рис. 8.15 представлены зависимости полных эмиттерных и коллекторных емкостей БТ, а также нормированных времен задержки ЭСЛ-элементов от плотности тока для транзисторов, отличающихся шириной эпитаксиального слоя и концентрациями. Транзистор БТ 4 выполнен с различными концентрациями в активной ($7 \cdot 10^{16}$ см⁻³) и пассивной (10^{15} см⁻³) областях.

Как видно из рис. 8.15, а, при прямых смещениях на эмиттерном переходе БТ его полная емкость определяется доминирующей диффузионной составляющей. Если при малых плотностях тока диффузионная емкость пропорциональна току, а наклон зависимости $C_{те}(J_K)$ зависит лишь от толщины базы, то при больших плотностях тока диффузионная емкость резко возрастает. Из рис. 8.15, а следует, что для БТ с одинаковой концентрацией, но различной толщиной эпитаксиального слоя (БТ 1, БТ 2) плотность

тока, при которой емкость $C_{те}$ начинает отклоняться от линейной зависимости, тем больше, чем больше толщина эпитаксиального слоя. Однако если при увеличении толщины эпитаксиального слоя повысить и концентрацию, то отклонение $C_{те}$ от линейной зависимости будет слабее, наступаая при больших плотностях тока (БТ 3, БТ 4). При низких плотностях тока полная емкость коллектора

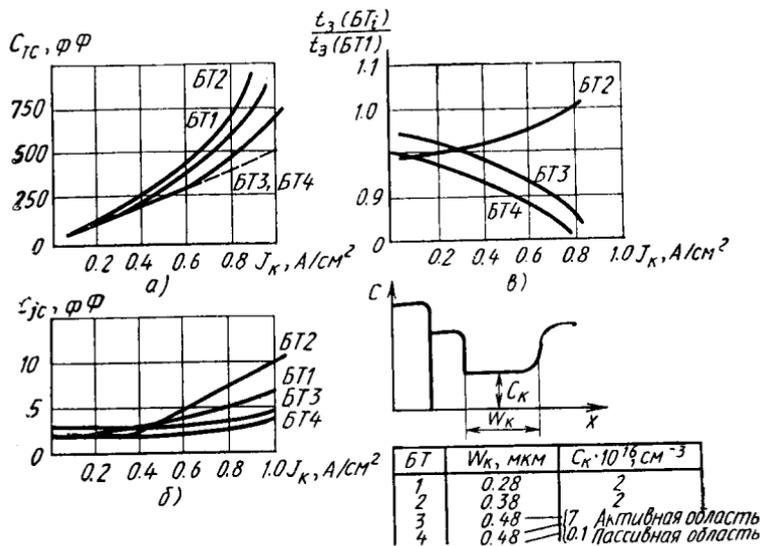


Рис. 8.15. Расчетные зависимости динамических параметров БТ C^3BIS от плотности режимного тока и концентрации примеси в коллекторе:
 а — полной емкости эмиттера; б — емкости коллектора; в — нормированного времени задержки ЭСЛ-элемента C^3BIS

определяется барьерной составляющей и пропорциональна корню квадратному из концентрации примесей коллектора. С увеличением же плотности режимного тока БТ отклонение от постоянной емкости коллектора будет больше для приборов с меньшими концентрацией и толщиной коллектора (рис. 8.15, б). Близкой к идеальной будет конструкция БТ со ступенчатым легированием: с сильным легированием в активной области и слабым — в пассивной. Такой БТ будет обладать одновременно минимальной емкостью коллектора и способностью работать при высоких плотностях тока без ухудшения рабочих характеристик. Зависимости нормированных времен задержки ЭСЛ-элементов C^3BIS от плотности коллекторного тока (рис. 8.15, в) показывают, что при малых плотностях тока задержка меньше у ЛЭ с меньшими толщ-

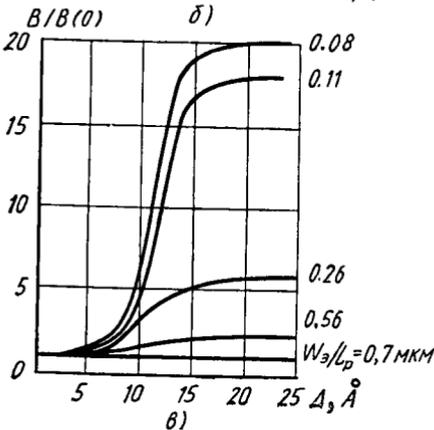
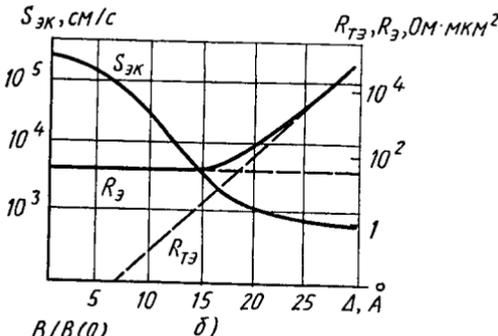
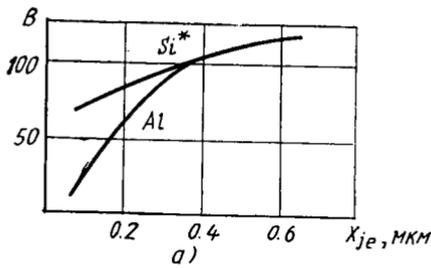


Рис. 8.16. Зависимости коэффициента передачи тока от глубины залегания эмиттерного перехода для БТ со стандартным и поликремниевым эмиттерами ($R_{БЭА} = 10 \text{ кОм}/\square$) (а), зависимости скорости поверхностной рекомбинации на эмиттерном контакте, полного и туннельного сопротивления эмиттера от толщины квазиокисного слоя (б): зависимости нормированного коэффициента передачи тока от толщины квазиокисного слоя для различных коэффициентов прозрачности поликремниевых БТ (в)

ной и концентрацией эпитаксиального слоя, а при больших плотностях тока — при использовании БТ с большей концентрацией эпитаксиального слоя. Лучшие динамические параметры, полученные для БТ со ступенчатой концентрацией примеси в эпитаксиальном слое (БТ 4), наглядно подтверждают ощутимое преимущество разработки таких БТ для С³БИС. При перспективных технологиях изготовления БТ с поликремниевыми контактами к эмиттеру и базе при минимизации пассивных областей БТ (особенно в технологии с боковыми поликремниевыми базовыми электродами) ступенчатое легирование коллектора нецелесообразно; выходом из положения может служить только увеличение концентрации примеси в коллекторной области БТ.

Поликремниевые БТ. В тонкослойных БТ с микротехнологией толщина эмиттера становится сравнимой или даже меньше диффузионной длины дырок в эмиттере: $W_3 \leq L_p$, $L_p(C_3 \approx (1 \div 2) \cdot 10^{20} \text{ см}^{-3}) \approx 0.15 \div 0.2 \text{ мкм}$; эмиттерная область такого БТ оказывается практически прозрачной для потока дырок. В тонкослойных БТ с «прозрачным эмиттером» скорость поверхностной рекомбинации в области эмиттерного контакта в значительной мере определяет коэффициент передачи тока. Если для БТ

с $W_{\text{э}} > 0.3 \div 0.4$ мкм базовый ток определяется механизмом объемной Оже-рекомбинации, то для более тонкослойных БТ коэффициент B определяется большей скоростью рекомбинации на контактирующей площади Al—Si ($s \approx 10^6$ см/с) и существенным разбросом скоростей рекомбинации. Рассматриваемые поверхностные явления и малая эффективность эмиттера из-за роста неконтролируемого базового тока для тонкослойных БТ с сильнолегированной базой и «прозрачным эмиттером» не позволяют получать приемлемых стабильных коэффициентов $B \approx 50 \div 100$. Это иллюстрирует рис. 8.16, а. Для БТ с различными глубинами залегания эмиттерного перехода $x_{je} = 0.1 \div 0.6$ мкм в рамках программы ФТ-моделирования рассчитаны зависимости коэффициента B с учетом разброса скорости поверхностной рекомбинации на эмиттерном контакте. Поэтому практическое масштабирование БТ, изготовленных по стандартным технологиям, неосуществимо вследствие резкой потери качества изготовления БТ. Поэтому в современных и перспективных конструктивно-технологических реализациях элементов С³БИС становится исключительной роль поликремниевых эмиттерных контактов, обеспечивающих в тонкослойных и сверхтонкослойных БТ высокие коэффициенты передачи тока наряду с малыми $\tau_{\text{тн}}$ и приемлемыми значениями напряжения прокола (≥ 5 В).

Для расчета поликремниевых БТ С³БИС использована расчетная в гл. 4 ФТ-модель, в которой влияние поликремниевogo контакта определяется эффективной скоростью рекомбинации $s_{\text{эк}}$ на границе раздела Si—Si* дополнительным туннельным сопротивлением эмиттера $R_{\text{тэ}}$. Из расчетных зависимостей $R_{\text{тэ}}$, $s_{\text{эк}}$ (рис. 8.16, б) видно, что $R_{\text{тэ}}$ увеличивается с ростом толщины квазиокисного слоя Δ (который может быть образован перед осаждением Si* на Si), в то же время сильный спад $s_{\text{эк}}$, наблюдаемый при $\Delta = 0 \div 10$ Å, становится незначительным при $\Delta \geq 15$ Å.

Результаты, приведенные на рис. 8.16, в для БТ с металлическим контактом, поликремниевым контактом без межфазного слоя, со слоем толщиной $\Delta = 5 \div 25$ Å, подтверждают возможность получения больших коэффициентов B поликремниевых БТ по сравнению со стандартными БТ, причем при $\Delta = 0$

$$\frac{B(\text{Si}^*)}{B(\text{Al})} = 2.2 \div 2.8, \frac{B(\text{Si}^*, \Delta = 10 \div 15 \text{Å})}{B(\text{Al})} = 5 \div 17, \text{ при } \Delta > 15 \text{Å} \text{ увеличение } B \text{ незначительно.}$$

Результаты расчетов косвенно согласуются с экспериментами, в которых при соответствующей обработке эмиттера со стравливанием квазиокисного слоя до осаждения на него поликремния ($\Delta = 0$) получалось соотношение $B(\text{Si}^*)/B(\text{Al}) = 2 \div 3$ в зависимости от толщины Si, а также с экспериментами для поликремниевых БТ с $\Delta = 10 \div 15$ Å, в которых $B(\text{Si}^*, \Delta)/B(\text{Al}) = 7 \div 12$.

Физико-технологическая оптимизация БТ для оценки максимального быстродействия логических элементов С³БИС. Основными конструктивно-технологическими методами, обеспечивающими одновременное увеличение степени интеграции и «технологического» быстродействия БТ кремниевых С³БИС, являются: 1) сверхсамосовмещение активных областей БТ; 2) использование поликремниевых мышьяковистых эмиттеров для создания тонкослойных сильнолегированных БТ; 3) локальная щелевая изоляция окислом и др. Методы физико-технологического моделирования дают уникальную возможность оптимизации БТ по критерию максимального «технологического» быстродействия логических элементов С³БИС за счет выбора «уровня тонкослойности» БТ с различными глубинами залегания эмиттерного, коллекторного и подложечного переходов при обеспечении приемлемых значений коэффициента B , потенциальном улучшении технологическими средствами выхода годных схем и уменьшении их стоимости.

Оптимальные параметры ЛЭ, характеризующие качество исполнения БТ в С³БИС, согласно (8.13), (8.14), можно выразить соотношением

$$I_{\text{опт}} \sim l \left[\frac{C_{\text{удк}} \eta^d(\eta)}{R_{\text{СБА}} \tau_{\text{ТН}}} \right]^{1/2},$$

$$t_{\text{з min}} \sim h \left[\frac{R_{\text{СБА}} \tau_{\text{ТН}} C_{\text{удк}} d(\eta)}{\eta} \right]^{1/2}, \quad (8.16)$$

$$I_{\text{опт}} t_{\text{з min}} \sim U_{\text{д}} h l d(\eta),$$

где $C_{\text{удк}}$ — удельная емкость активной коллекторной области; η — фактор конструкции БТ; $\eta = 1/3$ при $n_{\text{Б}} = 1$, $\eta = 1/12$ при $n_{\text{Б}} = 2$; $d(\eta)$ — фактор, отражающий зависимость площади коллекторного перехода от расстояния между контактами и $n_{\text{Б}}$ ($C_{\text{jc}} \approx C_{\text{удк}} h l d(\eta)$).

Уменьшение геометрических размеров БТ, согласно (8.16), приводит к уменьшению $I_{\text{опт}} t_{\text{з min}} \sim h l$, $I_{\text{опт}} \sim l$, $t_{\text{з min}} \sim h$.

Актуальной представляется возможность минимизации $t_{\text{з min}}$ за счет выбора профиля БТ (параметров технологического процесса), обеспечивающего при противоречивых предпосылках минимально возможные произведения $R_{\text{СБА}} \tau_{\text{ТН}} C_{\text{удк}}$, $R_{\text{СБА}} \tau_{\text{ТН}}$ при приемлемых значениях коэффициента передачи тока B . На рис. 8.17 приведены зависимости нормированного произведения $R_{\text{СБА}} \tau_{\text{ТН}} / (R_{\text{СБА}} \tau_{\text{ТН}})_0$ от нормированной дозы легирования в базе D/D_0 , рассчитанные с помощью программ Т- и ФТ-моделирования для двух технологических процессов (1 и 2 на рис. 8.17). Пунктиром отмечены ограничения для высоких D , дающих коэффициенты передачи тока во внутреннем БТ $B_0 \leq 80$.

На рис. 8.18 представлены расчетные зависимости B , $\tau_{\text{ТН}}$ и произведения $R_{\text{СБА}} \tau_{\text{ТН}}$ от поверхностного сопротивления активной

базы для различных БТ с «прозрачным» эмиттером ($x_{je} \approx 0.05 \div 0.08$ мкм). Отметим важную особенность выбора уровня тонкослойности БТ: реализуемый диапазон поверхностных сопротивлений $R_{S\text{BA}}$, с одной стороны, ограничен процессом резкого уменьшения коэффициента B (особенно при увеличении концентрации на границе кремний — поликремний): $R_{S\text{BA}} \leq (3 \div 4) 10^3$ кОм/□, с другой — потенциальным проколом базы; $R_{S\text{BA}} \geq 40 \div 60$ кОм/□.

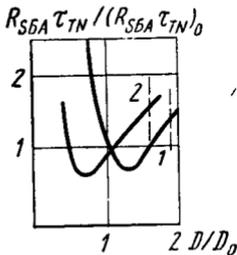


Рис. 8.17. Зависимости нормированного произведения поверхностного сопротивления активной базы и составляющей времени задержки эмиттер — коллектор БТ от нормированной дозы легирования базы для двух технологий изготовления БИС

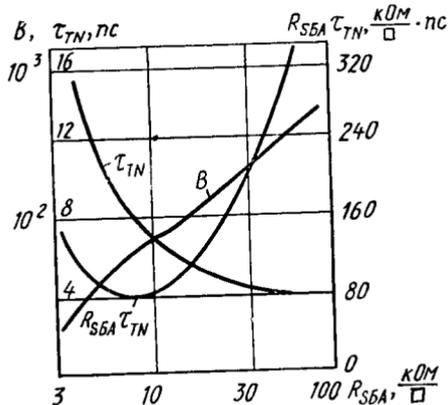


Рис. 8.18. Зависимости основных электрических параметров БТ от поверхностного сопротивления активной базы для оптимизации БТ и ЛЭ

С точки зрения оптимизации фактора $R_{S\text{BA}} \tau_{TN}$, определяющего в значительной степени максимальное «технологическое» быстродействие БТ, ЛЭ в С³БИС (при соответствующем выборе $C_{удк}$), наиболее приемлемы диапазон поверхностных сопротивлений $7 \leq R_{S\text{BA}} \leq 15$ кОм/□, обеспечивающих минимальное значение фактора $R_{S\text{BA}} \tau_{TN} \leq 90$ (кОм/□)пс при $B > 100$.

Примером физико-технологической оптимизации поликремниевое тонкослойного сильнолегированного БТ С³БИС с профилем примеси, рассчитываемым для определенного набора технологических операций и обеспечивающим параметры $R_{S\text{BA}} \tau_{TN} \leq 90$ (кОм/□)пс, $B \geq 200$, может служить синтез профиля с целенаправленной последовательностью «тестовых итераций» на минимизацию фактора $R_{S\text{BA}} \tau_{TN}$ и ограничением по B_{max} . Конечный в некотором смысле оптимальный профиль примеси в активной области БТ формировался такой последовательностью операций: 1) имплантацией бора с энергией 10 кэВ и дозой 4 мкКл/см² в эпитаксиальную пленку с удельным сопротивлением 0.1 Ом·см; 2) осаждением

поликремния толщиной 0.2 мкм; 3) имплантацией мышьяка с энергией 50 кэВ и дозой 800 мкКл/см²; 4) отжигом в течение 30 мин при температуре 950°C.

В результате физико-технологического расчета был сформирован профиль БТ с параметрами: глубины залегания эмиттерного и коллекторного переходов $x_{je}=0,075$ мкм, $x_{jc}=0,164$ мкм, поверхностное сопротивление $R_{SBA}=15$ кОм/□. Для такого профиля получены следующие топологически-инвариантные параметры: $C_{удк}=7,76 \cdot 10^4$ пФ/см², $c_{удэ}=4,2 \cdot 10^5$ пФ/см², $\tau_{тн}=6$ пс, $V_{max}=210$, причем произведение $R_{SBA}\tau_{тн}=90$ (кОм/□)пс и коэффициент B отвечают критериям оптимальности профиля БТ.

Для сверхсамосовмещенного БТ с двумя базовыми контактами, щелевой локальной изоляцией окислом, минимальным геометрическим размером 1 мкм, шагом металлизации 4 мкм, шириной межаобласти БТ 2 мкм площадь эмиттера $S_{э}=2$ мкм², площадь коллектора $S_{к}=4,4$ мкм², площадь подложки $S_{п}=18$ мкм². Активное сопротивление базы такого БТ $R_{БА} \approx (15/12)(1/2)=0,625$ кОм, пассивное сопротивление базы определяется практически сопротивлением слоя поликремниевого контакта базы $\rho=2 \cdot 10^{-3}$ Ом·см, т. е. $R_{бп}=75$ Ом. С учетом геометрии БТ (рассчитанных топологически-инвариантных параметров) полные барьерные емкости эмиттера и коллектора $C_{je}=8,4$ фФ, $C_{jc}=1,8$ фФ. В расчете емкости подложки по формуле $C_{кп}=C_{удт}S_{т}+C_{удб}^n\rho$, где был использован дополнительный параметр удельной боковой емкости подложки $C_{удб}^n=0,2$ фФ/мкм, получим $C_{кп}=7$ фФ.

Далее для оценки максимального быстродействия ЭСЛ-элементов БТ с оптимизированным профилем и щелевой изоляцией были рассчитаны с помощью программ схемотехнического моделирования переключательные характеристики одно-, двух- и четырехходовых стандартных ЭСЛ-элементов в режиме кольцевого генератора с логическим перепадом $U_{л}=0,4$ В при разных токах переключения I_0 . Для токов выходных эмиттерных повторителей был выбран диапазон 0.1—4 мА, обеспечивающий линейную работу повторителя при емкостной нагрузке и отсутствии заметных потерь быстродействия в выходных каскадах. Проведенные схемотехнические расчеты могут служить примером определения максимального «технологического» быстродействия ЛЭ, реализованного в рамках прогрессивных методов изготовления С³БИС при физико-технологической оптимизации БТ.

Однако по мере увеличения степени интеграции и размеров кристалла, а также логической гибкости ЛЭ системное быстродействие элементов, как показано на рис. 8.19, будет все больше определяться межэлементными ЛС: временем перезарядки нагрузочной емкости ЛС через выходное сопротивление элемента (ЭП), а также собственной RC-задержкой.

Достигнутое в рамках физико-технологической оптимизации БТ высокое «технологическое» быстродействие может быть практически сохранено с ростом степени интеграции С³БИС за счет ис-

пользования новых схмотехнических методов с выходными двухтактными переключателями напряжения комплементарного типа (в которых оба фронта выходного сигнала являются активными) при минимальном статическом потреблении мощности элементом, а также за счет секционирования ЛС.

Прогноз быстродействия и степени интеграции кремниевых биполярных С³БИС в зависимости от тока переключения (рис. 8.20) для различных минимальных литографических размеров $h = 2.5 \div 0.25$ мкм (для четырехвыходных ЛЭ и максимальной рассеиваемой мощности кристалла $P_{кр} \approx 2$ Вт) убедительно демонстрируют большие потенци-

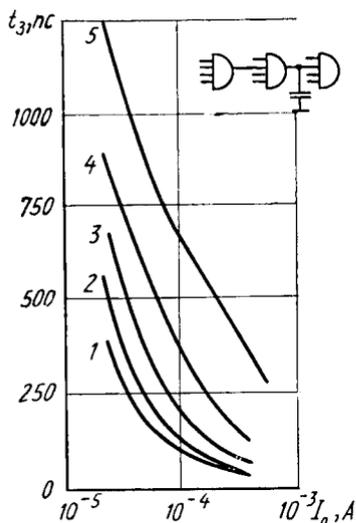


Рис. 8.19. Зависимости времен задержки ЭСЛ-элемента в цепочке ЛЭ, разработанных по 1 мкм-технологии, от тока переключения:

1 — одноходового, $C_H = 0$; 2 — двухходового, $C_H = 0$; 3 — четырехходового, $C_H = 0$; 4 — четырехходового, $C_H = 0.2$ пФ; 5 — четырехходового, $C_H = 0.8$ пФ

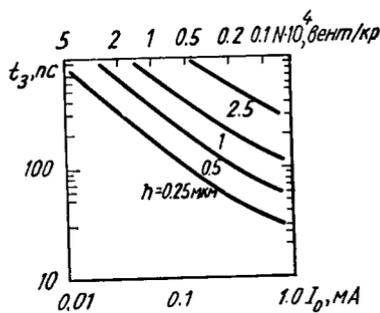


Рис. 8.20. Прогноз быстродействия и степени интеграции кремниевых биполярных С³БИС

альные возможности кремниевых С³БИС в реализации высокого системного быстродействия ЛЭ и дальнейшем росте степени интеграции.

8.4. ФОРМАЛЬНЫЕ И АДЕКВАТНЫЕ МАКРОМОДЕЛИ БЫСТРОДЕЙСТВУЮЩИХ ЛОГИЧЕСКИХ ЭЛЕМЕНТОВ И ФРАГМЕНТОВ БИС

Электрический расчет в рамках программ автоматизированного схмотехнического проектирования БИС, состоящих из 10^4 — $5 \cdot 10^5$ БТ, требует составления и описания системы из тысяч и десятков тысяч нелинейных обыкновенных дифференциальных уравнений, численное решение которой зачастую неосуществимо на со-

временных высокопроизводительных ЭВМ. Необходимость обращения к медленной внешней памяти ЭВМ приводит к значительному увеличению машинного времени расчета и делает моделирование неэффективным.

Достигнуть существенного увеличения производительности программ схемотехнического проектирования за счет улучшения методов численного решения и формирования систем дифференциальных уравнений нередко представляется невозможным. Одним из основных способов решения данной проблемы является использование в универсальных или специализированных программах схемотехнического проектирования упрощенных, но достаточно точных макромоделей ЛЭ и фрагментов БИС.

Макромодель ЛЭ или фрагмента БИС по сравнению с моделью на БТ или МОПТ является более простой, воспроизводящей с достаточной степенью адекватности поведение схемы для входных, передаточных и выходных характеристик и основные реакции схемы на внешние воздействия.

Основная цель, которую ставят и решают в процессе построения и использования макромоделей в универсальных или специализированных программах расчета схем, — упрощение и уменьшение размерности исходной схемы, что приводит к снижению порядка алгебродифференциальных уравнений, описывающих данный ЛЭ или фрагмент БИС. В свою очередь, этот фактор уменьшает время макромоделирования, сокращает оперативную память, необходимую для описания и расчета.

Формальные макромоделели основаны на формальной аппроксимации внешних характеристик цифровых ЛЭ и фрагментов, не имеют сходства с реальными ЛЭ и в целом инвариантны различным схемным базисам.

Адекватные макромоделели более полно учитывают специфику преобразования входных сигналов, переключения ЛЭ, отражают основные электрические процессы внутри и на выходе ЛЭ. Поскольку в рамках многоуровневого моделирования БИС возможно достаточно корректно идентифицировать параметры БТ, за основу адекватного макромоделирования принят синтез различных типов макромоделей ЛЭ и фрагментов БИС, параметрами которых являются основные электрические параметры БТ. При адекватном макромоделировании параметры БТ либо считают непосредственно параметрами макромоделей, либо из совокупности статических и динамических параметров БТ формируют новые переменные макромоделей ЛЭ и фрагментов на более высоком уровне описания БИС. При таком подходе имеется принципиальная возможность адекватного отражения в макромоделях физического объекта, учета разброса электрических параметров БТ и пассивных компонентов, связи технологических и электрических параметров БТ и ЛЭ, привлечения статистических методов анализа БИС.

Структура формальной макромоделели ЛЭ и фрагмента БИС. Макромодель строят для любого фрагмента цифровой БИС, удов-

летворяющего требованиям: 1) фрагмент имеет m входов, n выходов; 2) каждому выходу соответствует булева логическая функция от состояний на входах; 3) ток каждого входа зависит от напряжения на входе и скорости его изменения $dU_{\text{вх } i}/dt$; 4) ток каждого выхода в статическом режиме зависит только от напряжения и состояния на входе.

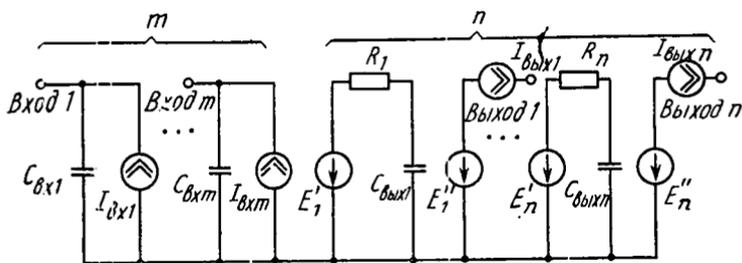


Рис. 8.21. Структура формальной макромоделли ЛЭ и фрагмента БИС

Фрагмент описывается статической $I_{\text{вх } i}(U_{\text{вх } i})$ и динамической $I_{\text{вх } i}(U_{\text{вх } i}, dU_{\text{вх } i}/dt)$ — входной характеристиками; статической передаточной характеристикой при холостом ходе на выходе $U_{\text{вых } j}(U)$, $i=1 \div m$, $j=1 \div n$. Обозначив $U_{\text{вых } j}$ напряжение на j -м выходе, U — одно из входных напряжений, конкретизируют статическую выходную характеристику в состоянии логического «0» $I^0_{\text{вых } j}(U_{\text{вых } j})$, статическую выходную характеристику в состоянии логической «1» $I^1_{\text{вых } j}(U_{\text{вых } j})$, а также переходную характеристику при холостом ходе на выходе $U_{\text{вых } j}(t)$, $j=1 \div n$.

Структура формальной макромоделли фрагмента представляет собой эквивалентную схему из m входных и n выходных блоков (рис. 8.21). Источник тока $I_{\text{вх } i}$ моделирует статическую, а емкость $C_{\text{вх } i}$ — динамическую входные характеристики фрагмента. Источник напряжения E_j моделирует статическую передаточную характеристику $U_{\text{вых } j}(U)$, где определяющее напряжение U есть функция входных переменных, реализуемая из логической функции, соответствующей данному входу. С помощью выходной $R_j C_{\text{вых } j}$ -цепочки моделируют переходную характеристику j -го выхода. Значение напряжения источника равно напряжению на емкости, а в статическом режиме — напряжению источника E_j . Источник тока $I_{\text{вых } j}$ позволяет учесть нагрузочную способность фрагмента, моделируя статические выходные характеристики в состояниях логических «0» или «1».

Все блоки, моделирующие статические характеристики, описываются кусочно-линейными функциями типа $y = a_l x + b_l$, $l=1 \div L$, где y — любая из функций, описывающих $I_{\text{вх } i}$, E'_j , $I_{\text{вых } j}$; x — соответствующие им аргументы $U_{\text{вх } i}$, $U_{\text{вых } j}$, U .

Инерционность перераспределения зарядов во входной цепи при подаче на вход изменяющихся во времени сигналов учитывают введением зависимости входных емкостей $C_{вх i}$ от $dU_{вх i}/dt$. Чтобы с достаточной степенью точности смоделировать переходную характеристику, на ней выделяют несколько характерных участков, в каждом из которых задают свою постоянную времени $\tau_j = R_j C_{вх j}$, причем обычно полагают $R_j \approx \text{const}$ и $C_{вх j} = \text{var}$.

Макромоделирование БИС большой размерности позволяет получить существенный выигрыш во времени (на порядок и выше), при этом точность формальной макромодели не зависит от объема моделируемой схемы, так как ее параметры определяются по принципу «черного ящика».

Адекватные (блочные и схемные) макромодели ЛЭ. В том случае, когда достигнутую с помощью формальных макромоделей точность расчета БИС нельзя считать удовлетворительной, применяют адекватные макромодели. Синтез адекватных макромоделей требует учета схемотехники и специфики используемых электрических режимов, функционирования нелинейных электрических моделей БТ, аппроксимативных линейных приближений данных моделей и других особенностей.

Синтез адекватных макромоделей (блочных и схемных) широкого класса ЛЭ и фрагментов БИС, например из различных комбинаций многоходовых ПТ и ЭП, проводят в общем случае на основе снижения порядка алгебродифференциальных уравнений, описывающих динамику и статику ПТ, ЭП и всего ЛЭ, усреднений, локальных линеаризаций, упрощений в базовых уравнениях с учетом соотношений между параметрами БТ и различными постоянными времени ЛЭ, а также особенностей схемотехники и электрического режима ЛЭ.

Для многоходового ЭП, работающего на RC -нагрузку, связь между входными и выходными напряжениями задается из решения системы уравнений Кирхгофа

$$U_{вых 1} \left[\frac{d^2}{dt^2} A_{11} + \frac{d}{dt} A_{21} + 1 \right] = U_{вх 1} \left[\frac{d^2}{dt^2} B_{11} + \frac{d}{dt} B_{21} + 1 \right] - U_{Э 1},$$

$$U_{вых n} \left[\frac{d^2}{dt^2} A_{1n} + \frac{d}{dt} A_{2n} + 1 \right] = U_{вх n} \left[\frac{d^2}{dt^2} B_{1n} + \frac{d}{dt} B_{2n} + 1 \right] - U_{Э n},$$
(8.17)

$$U_{вых} = (U_{вых 1} + \dots + U_{вых n})/n, \quad I_{Э 1} + \dots + I_{Э n} = (E - U_{вых})/R_n,$$

где $U_{Э n} = \varphi_T \ln I_{Э n}/I_{Э 0}$ — напряжение на эмиттерном переходе; $I_{Э n}$ — статический эмиттерный ток, $A_{1n} = \tau_{Э} (\tau_{М вх} + R_1 C_{М вх}) (1 + R_{М вх}/R_n)$; $A_{2n} = \tau_{Э} (1 + R_1/R_n) + C_{М вх} (r_{Э} + R_{М вх} (1 + r_{Э}/R_n))$; $B_{1n} = \tau_{Э} \tau_{М вх}$, $B_{2n} = \tau_{М вх} + \tau_{Э}$, $\tau_{Э} = r_{Э} (\bar{C}_{je} + \tau_{TN} \lambda_{Э})$; $\tau_{М вх} = C_{М вх} R_{М вх}$; $\lambda = 1/\varphi_T$, $\varphi_T/I_{Э} \neq r_{Э}$; $R_1 = R_{БП} + R_{БА} (C_{КП}/C_{je})$ — приведенное сопротивление базы; $C_{М вх}$, $R_{М вх}$ — емкость и сопротивление, моделирующие входное сопротивление ЛЭ, на который работает ЭП.

Для расчетов с повышенной точностью коэффициенты A_{1n} , A_{2n} , B_{1n} , B_{2n} вычисляются на каждом временном шаге решения системы (8.17).

При выводе уравнений, описывающих работу ПТ, были сделаны следующие допущения: 1) $I_0 = \text{const}$, $U_{\text{оп}} = \text{const}$; 2) параметры БТ ПТ идентичны; 3) ВАХ эмиттерных диодов представлены в виде кусочно-линейной аппроксимации $U = \varphi_T \ln \frac{I_{\text{Э}}}{I_{\text{Э}0}} + i r_{\text{Э}}$.

При аппроксимативном снижении порядка результирующего алгебродифференциального уравнения получены выражения, связывающие напряжение между внутренними базовыми точками входного и опорного БТ $U_{\text{ББ}i}(t)$ и напряжения на входе, а также на прямом и инверсном выходах ПТ:

$$U_{\text{ББ}i} = \frac{U_{\text{вх}i} - U_{\text{оп}} + \frac{d}{dt} C_1 R_2 (U_{\text{К}i} - U_{\text{н}})}{1 + g_{1i} + g_{2i}},$$

$$g_{1i} = (R_{\text{Э}} + R_1/B + 1) g_{3i},$$

$$g_{2i} = R_2 C_1 (1 + g_{1i}) + (R_1 + R_2) (\bar{C}_{je} + \tau_{\text{т}N} \lambda g_{3i}), \quad (8.18)$$

$$g_{3i} = I_0 \exp \lambda U_{\text{ББ}i} \left/ \left[1 + \sum_{i=1}^m \exp \lambda U_{\text{ББ}i} \right] \right. (1 + \exp \lambda U_{\text{ББ}i}),$$

где $U_{\text{вх}i}$ — напряжение на входе i -го переключающего БТ m -входного ЛЭ; $R_2 = R_{\text{БЛ}} C_{\text{КЛ}}/C_{j\text{с}}$.

Зависимости коллекторных токов i -го переключающего и опорного БТ от $U_{\text{ББ}i}$ представляются в виде

$$I_{\text{К}i} = I_0 - I_{\text{КП}}, \quad (8.19)$$

$$I_{\text{КП}} = I_0 \left/ \left(1 + \sum_{i=1}^m \exp \lambda U_{\text{ББ}i} \right) \right. . \quad (8.20)$$

Для коллекторных узлов, к которым подключены емкости нагрузки $C_{\text{н}}$ (включающие емкость коллектора $C_{j\text{с}}$), выполняются следующие соотношения:

для инверсного выхода

$$U_{\text{К}i} \left(1 + \frac{d}{dt} C_{\text{н}} \frac{U_{\text{л}}}{I_0} \right) = \frac{U_{\text{л}}}{I_0} \left[I_{\text{К}i} + \frac{d}{dt} C_1 (U_{\text{вх}i} - U_{\text{К}i}) \right], \quad (8.21)$$

для прямого выхода

$$U_{\text{н}} \left(1 + \frac{d}{dt} C_{\text{н}} \frac{U_{\text{л}}}{I_0} \right) = \frac{U_{\text{л}}}{I_0} \left(I_{\text{КП}} - \frac{d}{dt} C_1 U_{\text{н}} \right). \quad (8.22)$$

Система уравнений (8.18) — (8.22) описывает работу ПТ. На рис. 8.22 представлена адекватная блочная макро модель многовходового ЭСЛ-элемента (см. рис. 8.2) в виде набора блоков,

каждый из которых описывается соответствующим уравнением (8.17)—(8.22). Аналогично могут быть описаны элементы многоярусной, эмиттерно-диодной, эмиттерно-функциональной, гистерезисной, комплементарной и других типов перспективных ЛЭ на ПТ и ЭП.

При расчете статической, импульсной помехоустойчивости задержек и других параметров электрических цепей БИС в общем случае не интересуются процессами, происходящими внутри каждого ЛЭ, но детально анализируют сигналы на их входах и выходах. Выделим основные особенности рассматриваемого класса ненасыщенных ЛЭ: 1) большое входное и малое выходное сопро-

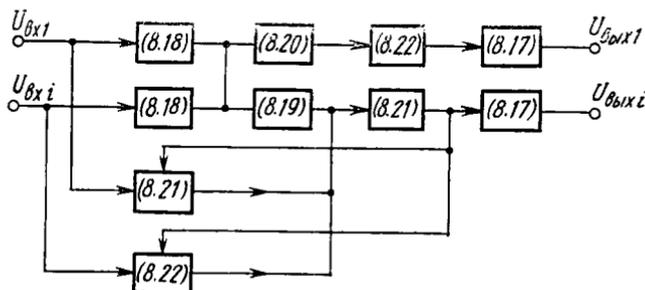


Рис. 8.22. Адекватная блочная макромоделль ЭСЛ-элемента

тивления; 2) плавные изменения напряжений и токов на выходах ЛЭ; 3) однонаправленный характер работы ЛЭ; 4) заметная инерционность ЛЭ. Большое входное и малое выходное сопротивления позволяют при расчете цепей ЛЭ учитывать влияние смежных частей фрагментов схемы за счет подсоединения к ЛЭ эквивалентов по входу и выходу (сопротивлений, нагрузочных емкостей). Плавные изменения токов и напряжений и инерционность ЛЭ дают возможность применять для решения дифференциальных уравнений простейший пошаговый итерационный метод Эйлера, а для решения алгебраических уравнений — метод последовательной итерации, разрывать обратные связи между входом и выходом ЛЭ (или целого узла) на интервале времени, определяющем интервал инерционности (порядка 0.1 длительности фронта ЛЭ). Однонаправленный характер работы ЛЭ позволяет в рамках специализированных программ адекватного макромоделирования при расчете цепей ЛЭ использовать эффективное блочное (поэлементное) моделирование БИС соответственно порядку прохождения электрических сигналов.

Расчет БИС, состоящей из подсхем с большими входными и малыми выходными сопротивлениями, может быть сведен к простейшему последовательному расчету, когда к внешним зажимам таких подсхем подключаются эквивалентные источники входных

ходов; $L_{эф}$ — эффективная индуктивность ЭП, моделирующая инерционные особенности выходного сопротивления: $L_{эф} = (R_{Б}^{ЭП} + R_{К})\tau_{ТН}$; $R_{эк} \approx R_{Б} + R_{К}$ — эквивалентное выходное сопротивление; $D_{эф}$ — эффективный диод смещения, описывающий ВАХ ЭП (в общем случае он шунтируется небольшой емкостью $C_{те}$); $R_{н}$ — сопротивление нагрузки ЭП; $E_{х}$ — соответствующее напряжения холостого хода со стороны ПТ; $E_{п}$ — напряжение питания схемы. Основная специфика схемного адекватного моделирования — описание нелинейных характеристик коллекторных токов прямого и инверсного выходов ПТ согласно соотношениям (8.19), (8.20) и входной емкости в виде

$$C_{вхi} = C_{je} + \lambda\tau_{тN}g_{zi} + C_{м}, \quad (8.23)$$

где g_{zi} определяется из (8.18); C_{je} — усредненная барьерная емкость эмиттерного перехода; $C_{м}$ — дополнительная емкость монтажа.

Значительное упрощение схемных макромоделей ЛЭ — ЭСЛ, многоярусных, ЭФЛ, ЭДЛ, гистерезисных, комплементарных — достигается за счет использования в (8.23) усредненных входных емкостей, дающих их оценку сверху, например $C_{вхi}^* \approx 3C_{je}(0) + 2\tau_{тN}I_0/U_{л}$, а также кусочно-линейной аппроксимации ВАХ диода ЭП.

Сравнение результатов моделирования на основе моделей БТ и схемных макромоделей ЛЭ дает выигрыш макромоделирования в 3—5 раз при приемлемой точности расчета 15—25%. При этом результаты схемного макромоделирования практически не отличаются от результатов блочного моделирования, но время подготовки данных существенно сокращается, а оперативность и эффективность расчета возрастают.

8.5. ЛОГИКО-ВРЕМЕННОЕ МАКРОМОДЕЛИРОВАНИЕ БИС

Использование адекватных блочных и схемных макромоделей логических элементов (ЛЭ) для расчета БИС, состоящих из нескольких сотен — тысяч ЛЭ, обуславливает не всегда оправданные большие вычислительные затраты оперативной памяти и машинного времени (например, 0,5—1 с для ЛЭ ЭВМ БЭСМ-6). Необходимость контроля логико-временного поведения быстродействующих матричных БИС, минимизации риска динамических сбоев, обоснования увеличения тактовых частот ЭВМ с разработкой соответствующих руководящих технических материалов (РТМ) на проектирование всех узлов ЭВМ делает крайне актуальной разработку моделей и специализированных программ логико-временного расчета БИС, содержащих 10^4 — 10^6 ЛЭ, на основе упрощенного, но достаточно точного (с погрешностью 20—25%) описания в них переходных процессов с минимальными вычислительными затратами.

Программа логико-временного расчета интегральных схем (ПЛОВРИС) предназначена для макро моделирования переходных процессов в БИС, содержащих 10^4 ЛЭ, с минимальными вычислительными затратами.

Формируется расширенная библиотека элементов с соответствующим описанием их логических функций и временных параметров сигналов, зависящих от основных режимных и эксплуатационных факторов. Входную информацию для ЛЭ — зависимости логических уровней и временных параметров выходных сигналов ЛЭ упрощенной трапецеидальной формы (с соответствующими за-

держками и линейными фронтами) — получают из расчета ЛЭ и их цепей с помощью программы схемотехнического моделирования. В зависимости от целей анализа и расчета БИС в программе ПЛОВРИС предусмотрено несколько вариантов описания ЛЭ, например уровни логических «1» и «0» зависят или не зависят от количества нагрузок ЛЭ, трапецеидальное или прямоугольное описание выходных сигналов с модификацией задержек переднего и заднего фронтов. В программе реализован алгоритм асинхронного событийного моделирования в семиричной логике. Все ЛЭ делятся на две группы: инерционные и безынерционные, которые обрабатываются

принципиально различным способом. Состояния безынерционных ЛЭ определяются на каждом временном шаге с процессе итераций. Для инерционных ЛЭ учитывается тот факт, что они могут изменить свое состояние только через некоторое время после изменения состояния входов и, следовательно, итерации по ним не производятся. Для инерционных ЛЭ динамически создается список переключений, в который заносятся переключающиеся элементы. По завершении процесса переключения ЛЭ исключаются из списка, освобождая отведенную под них память.

Структурная схема программ ПЛОВРИС представлена на рис. 8.24. Цифрами обозначены следующие блоки: 1 — считывание информации о схеме и выдача диагностики ошибок описания; 2 — считывание значений на входах; 3 — итеративное установление схемы; 4 — цикл по времени; 5 — установление наличия переключений в списке переключений; 6 — итеративное установление безынерционных ЛЭ; 7 — запись состояний, определение сбоев; 8 — печать результатов счета; 9 — определение следующего временного шага; 10 — считывание значений на входах; 11 — печать информации о сбоях ЛЭ.

Расчет схемы начинается со считывания входной информации

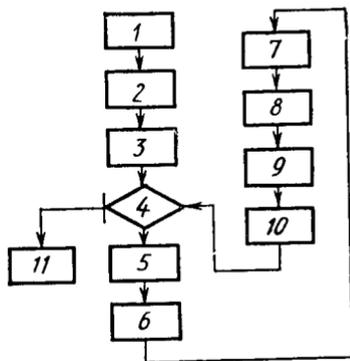


Рис. 8.24. Структура программы логико-временного моделирования

(блок 1) с незамедлительной выдачей диагностики ошибок, которые могут быть сразу же исправлены. По окончании приема информации о схеме и считывании значений на входах (блок 2) вычисляют состояния всех ЛЭ для нулевого значения времени (блок 3). Эта процедура позволяет исключить при счете режим установления схемы. Далее работа программы ПЛОВРИС производится с учетом реальных фронтов и задержек ЛЭ. Для вычисления фронтов и задержек ЛЭ (рис. 8.25, а, б) используется процедура ТАУ, которая по заданному времени T определяет параметры T_1, T_2, T_3 по известным логическому перепаду U_n , напряжению $E_{оп}$, количеству нагрузок n , ЛС и др. Фронты и задержки

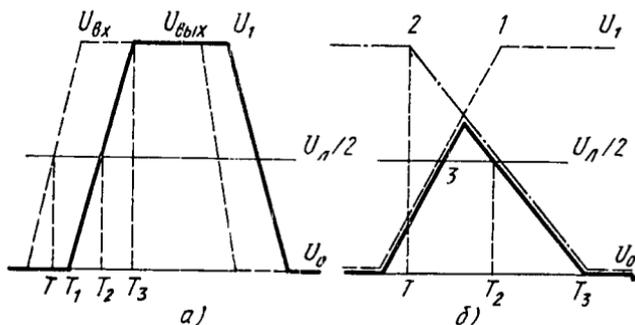


Рис. 8.25. Временные диаграммы ЛЭ при логико-временном моделировании

ки ЛЭ представлены суммами «собственных» (присущих m -входному ЛЭ без дополнительных нагрузок, связей, зависящих только от температуры) и «внешних» (с учетом функционирования в цепи ЛЭ) значений: $T_1 = T + T_3$, $T_2 = T_1 + T_{\Phi}(1 + k\kappa)/2$, $T_3 = T_1 + T_{\Phi}$, $T_{3, 0-1} = T_{3, 0, 0-1} + \alpha_{n, 3, 0-1}n + \alpha_{T, 3, 0-1}(t - t_0)$, $T_{\Phi, 0-1} = T_{\Phi, 0, 0-1} + \alpha_{n, \Phi, 0-1}n + \alpha_{T, \Phi, 0-1}(t - t_0) - k\alpha_{\kappa, \Phi, 0-1}(U_n/U_{n0})(1 + k\kappa) + \alpha_n(U_n - U_{n0})/I_0$, где $\kappa = 2(U_n/2 - U_0)/(U_n - 1)$; n — количество нагрузок; I_0 — ток переключения; α_i — коэффициенты, определяемые из компонентного расчета ЛЭ, согласуемые с результатами натурального эксперимента; $k = 1$ (-1) — для переднего (заднего) фронта; t — температура.

Обращение к процедуре ТАУ происходит каждый раз, когда ЛЭ заносится в список переключений, что позволяет за счет незначительных дополнительных временных затрат отказаться от использования большого по размерности массива задержек. По значениям времени T_1, T_2, T_3 однозначно определяют состояние «активного» ЛЭ.

Для определения связей между ЛЭ служит массив, в котором задаются для каждого ЛЭ номера ЛЭ, связанного с ним, и номер входа, к которому подсоединяется выход рассматриваемого ЛЭ.

В начале этого цикла (блок 5) просматривается список переключений и в случае перехода ЛЭ через уровень $U_{л/2}$ производится запись в начало списка переключений ЛЭ, которые должны переключаться. Следующим шагом определяются состояния всех безынерционных ЛЭ в процессе итераций (блок 6). По окончании этой процедуры все инерционные ЛЭ, переключающиеся в будущем, заносятся в список переключений. Таким образом, производится прогноз переключения. После этих процедур заново просматривается весь список переключений для записи значений в массив выходов ЛЭ и определения возможности сбоя (блок 7).

Сбой элемента определяют, как показано на рис. 8.25, б, обозначениями: 1 — старый прогноз, 2 — новый прогноз, 3 — действительный сигнал ЛЭ.

Для определения нового прогноза используют процедуру ТАУ с новыми значениями параметра T и фронта (передний фронт заменяется на задний, и наоборот). При превышении амплитуды выброса уровня $U_{л/2}$ сбой распространяется дальше по схеме, в противном случае просто фиксируется сбой ЛЭ.

Значительным этапом расчета на данном временном шаге является определение следующего момента $T_{\min} = \min\{T_1, T_2, T_3, T_{вх}\}$ (блок 9), где $T_{вх}$ — очередное значение времени задания входных данных, а \min отыскивается по всему списку переключений. В зависимости от T_{\min} происходит считывание входных значений (блок 10). Такой выбор временного шага позволяет отказаться от дискретизации задержек и фронтов, не пропускает момент (процесс) переключения в БИС и значительно экономит машинное время, затрачиваемое на расчет фрагментов БИС. Для удобства вывод на печать можно производить либо с фиксированным шагом по времени, либо все значения подряд. Завершает работу программы печать всех прошедших сбоев (блок 11).

Для разных типов фрагментов матричных БИС, состоящих из нескольких сотен ЛЭ, получены средние времена расчета статистики 0.1 мс и динамики 5—7 мс на элементарный ЛЭ на каждый шаг (при ограничении оперативной памяти ЭВМ БЭСМ-6 в 32 К бит). Количество обрабатываемых ЛЭ БИС может быть значительно увеличено, например до 10^5 — 10^6 , при уменьшении среднего времени расчета ЛЭ в 2—3 раза и расчете на более высокопроизводительных ЭВМ с принципиальным сохранением структуры и организации работы ПЛОВРИС.

Поскольку данные схемотехнического расчета ЛЭ и их цепей в БИС в широком диапазоне функционально-эксплуатационных факторов используются в логико-временной модели ЛЭ с минимальным огрублением формы сигнала, программа ПЛОВРИС позволяет эффективно, оперативно, с высокой степенью адекватности исследовать переходные процессы во фрагментах БИС.

Приемлемая адекватность макро моделирования и минимизация вычислительных затрат расчета БИС позволяют использовать методы и программные средства логико-временного макро моделиро-

вания в процессе проектирования узлов ЭВМ с расчетом статической и импульсной помехоустойчивости, обеспечением работоспособности БИС при широком изменении различных внешних и эксплуатационных факторов. В этой связи весьма привлекательной выглядит идея разработки принципиально новых РТМ с выявлением дополнительных возможностей повышения быстродействия и качества разработки БИС и узлов ЭВМ. Широко используемые в инженерной практике РТМ, составленные либо из расчета на наилучший случай, либо типовые, на основе схмотехнического анализа небольшого количества критических цепей ЛЭ с существенным ограничением ЛЭ вследствие высокой размерности анализируемых схем, накладывают довольно жесткие ограничения на ЛЭ с произвольными коэффициентами объединения и разветвления в подсистеме БИС. «Мелкоблочность» компонентного (схмотехнического) расчета и методов формального макро моделирования с существенным ограничением на количество анализируемых многовходовых ЛЭ, положенная в основу разработки традиционных инженерных РТМ первого поколения, не позволяет зачастую полностью реализовать возможности достижения более высокого быстродействия элементной базы БИС и ЭВМ. «Крупноблочные» методы логико-временного моделирования, использующие результаты сквозного многоуровневого расчета БИС (начиная от технологии их изготовления), электрические параметры приборов, ЛЭ и их оптимизацию, позволяют за счет более глубокого анализа многочисленных вариантов проектируемых фрагментов ЛЭ, исследования разнообразных электронно-логических ситуаций не только снять многие ограничения для стадии логического проектирования, но и максимально реализовать возможности технологии и конструктивные модификации разрабатываемых БИС.

8.6. СМЕШАННОЕ ЛОГИКО-ВРЕМЕННОЕ, СХМОТЕХНИЧЕСКОЕ И ФИЗИКО-ТЕХНОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ БИС

Фрагмент схемы или целиком БИС можно представить и описать на нескольких уровнях сложности и точности, зависящих от детализации формирования ее модели, которая может быть, например, функциональной, вентиляционной, приборной. Фактически полный цикл моделирования и оптимального проектирования БИС, изготавливаемой с помощью современных и перспективных технологий, с использованием того или иного приборно-технологического базиса, схмотехнических и системных решений, реализуется на всех иерархических уровнях, следовательно, целесообразность многоуровневого моделирования и проектирования БИС с целью улучшения всех ее функционально-эксплуатационных параметров не вызывает сомнения.

Основой сквозного многоуровневого моделирования БИС служит иерархическая структура расчетов приборов и схем с общей базой данных. Эффективным и перспективным методом, позволяю-

щим не только минимизировать вычислительные затраты машинных расчетов, но и улучшать качество вычислительного эксперимента, выявлять новые электрофизические закономерности в поведении прибора в схеме и схемы в реальных условиях ее функционирования в широком режимно-температурном диапазоне, является смешанное или комбинированное моделирование. В рамках данного метода различные по назначению программы работают на едином комбинированном уровне, но их функции при этом различны. Большинство современных программ моделирования реализует возможность совместимых логико-временных расчетов БИС на уровне вентилях и подсхем с расчетами на уровне приборов (БТ, МОПТ). Необходимость смешанного моделирования на уровне БТ и вентилях обусловлена разработкой новых конкурентоспособных БИС, которые практически невозможно эффективно моделировать и верифицировать только на уровне вентилях. При этом проектирование БИС на уровне БТ позволяет синтезировать больше логических функций, приходящихся на один БТ, чем проектирование на уровне вентилях, т. е. в рамках смешанного схемотехнического и рассмотренного выше событийного логико-временного моделирования имеется возможность резерва экономии машинного времени и памяти, а также улучшения качества проектирования БИС за счет использования передовых оптимальных схемных и системных решений.

Новую веху в развитии методов сквозного многоуровневого моделирования БИС открывает разработка программ смешанного схемотехнического расчета фрагментов БИС с использованием адекватных (двумерных и трехмерных) физико-топологических моделей БТ. Это наиболее актуально при стремительном совершенствовании субмикронных микротехнологий С³БИС на сверхсамосовмещенных тонкослойных БТ, работающих в режиме высоких плотностей тока в пикосекундном диапазоне, для выполнения цели одновременного увеличения системного быстродействия и уровня интеграции кремниевых биполярных С³БИС. Кардинальный метод объединения ФТ- и схемотехнического уровней для адекватного моделирования переходных процессов в фрагментах БИС в области малых времен позволяет резко увеличить адекватность, информативность расчетов за счет большей «физичности» описания процессов в приборах и схемах, отказа от ряда промежуточных операций синтеза и идентификации электрических распределенных моделей БТ на основе его упрощенного квазистатического поведения.

Математическая модель фрагмента БИС в схемотехнической части программы смешанного расчета формируется методом узловых потенциалов. Электрическая модель БТ представлена двумя источниками тока — коллекторного и базового, соединенными в общей эмиттерной точке, зависящими от напряжения на их выводах и некоторых внутренних переменных. Статический режим рассчитывается методом Ньютона, причем на каждой ньютонской

итерации производится обращение к ФТ-модели БТ, решение ФСУ полупроводника (с соответствующими начальными и граничными условиями), а также вычисление коллекторного I_K и базового I_B токов, которые, в свою очередь, используются на следующей ньютоновской итерации. Переходный процесс моделируется с использованием модифицированной схемы Эйлера.

Математическая модель фрагмента БИС для вектора узловых потенциалов φ записывается в виде

$$A \frac{d\varphi}{dt} = B(\varphi, t) + F(\varphi, t, I_n)$$

где вектор в правой части разбит на два слагаемых: вектор B формируется за счет «схемотехники», электрических моделей элементов, а вектор F — за счет «физики», ФТ-моделей БТ, для которых I_n — внутренние переменные. Расчет переходного процесса происходит так: решается разностная схема

$$A \frac{\varphi^{l+1} - \varphi^l}{\tau^{l+1}} = B(\varphi^l, t^l) + \frac{dB(\varphi^l)}{d\varphi} (\varphi^{(l+1)} - \varphi^l) + F(\varphi^l, t^l, I_n), \quad (8.24)$$

причем B линеаризуется в окрестности φ^l , а F берется с предыдущего временного шага, затем с помощью ФТ-моделей определяются токи I_K, I_B и вычисляется $F(\varphi^{l+1}, t^{l+1}, I_n)$. При выборе временного шага, диктуемом получением необходимой (заданной) точности расчета по ФТ-модели, схема (8.24) устойчива.

Нестационарная двумерная ФТ-модель БТ основана на использовании полунейвной абсолютно устойчивой разностной схемы (5.89)—(5.91), когда эти уравнения решаются последовательно на каждом шаге по времени, после чего производится вычисление токов I_K, I_B . Решение в базисе переменных экспонент квазипотенциалов позволяет симметризовать матрицы коэффициентов (5.90) и (5.91). В зависимости от типа используемой ЭВМ для решения используется либо метод ICSSG, либо прямые методы решения СЛАУ с сильно разреженными матрицами коэффициентов, которые значительно эффективнее итерационных, однако требуют большего объема оперативной памяти.

Время смешанного схемотехнического и ФТ-расчетов линейно зависит от числа БТ, описанных ФТ-моделями, и составляет в зависимости от типа используемой ЭВМ 10—30 мин на одну ФТ-модель (100 шагов по времени, 500—600 узлов СПР).

Применение методов смешанного моделирования в сквозном расчете БИС позволяет при взаимодействии уровней (см. рис. 1.6) реализовать уникальные дополнительные возможности оптимального проектирования микротехнологий, а также собственно приборов и схем как за счет совместного физико-технологического расчета БТ с автоматическим выбором СПР, так и за счет более эффективных приемов синтеза и идентификации моделей БТ, ЛЭ и фрагментов БИС.

Разработка новых технологий изготовления кремниевых С³БИС с субмикронными минимальными размерами приборных структур открыла перспективы значительного увеличения быстродействия и степени интеграции элементно-технологической и конструктивной базы высокопроизводительных ЭВМ высших поколений. В результате еще более усилилась взаимосвязь между технико-эксплуатационными параметрами С³БИС, собственно приборных структур и технологий их изготовления. Роль физико-технологического моделирования и последующих этапов электрического (схемотехнического), логико-временного и проектно-конструкторского моделирования С³БИС становится основополагающей в свете современных требований достижения лучших или даже предельно возможных для современного схемно-технологического уровня технико-экономических характеристик кремниевых С³БИС, высокой конкурентоспособности по отношению к другим перспективным вариантам реализаций аналогичных разработок.

Высокая эффективность автоматизации проектирования БИС и одного из главных его направлений — многоуровневого моделирования начиная от этапа физико-технологического расчета интегральных технологий и приборов — уже подтверждена в процессе эксплуатации САПР и конкретных разработок БИС. Однако имеются огромные резервы повышения эффективности проектирования и разработки элементно-технологической базы с учетом тенденций и перспектив развития методов и технических средств многоуровневого моделирования, позволяющих резко сократить сроки и стоимость проектирования, улучшить качество разработок.

Требования к сверхминиатюризации приборных структур, максимальному сжатию их горизонтальных и вертикальных размеров, сильному легированию полупроводниковых и диэлектрических слоев обусловили появление за небольшой отрезок времени нескольких поколений программно-алгоритмического обеспечения и моделей технологического и физико-топологического расчета элементов С³БИС. В зарождающемся третьем поколении программных средств и моделей технологического моделирования необходимо выделить современную тенденцию и актуальность построения эффективных численных многомерных моделей технологических процессов с совершенствованием кинетических механизмов в субмикронных приборных структурах и тесным увязыванием со спецификой топологии и особенностями профилей приборных структур.

Бурное развитие современных и перспективных поколений программных средств и физико-топологических моделей идет по пути

совершенствования эффективных численных методов решения фундаментальных уравнений для полупроводника в многомерном приближении с учетом реальных профилей и конструкций приборных структур, а также по пути использования новых более сложных кинетических моделей переноса заряда в тонкослойных структурах, более полного, адекватного описания в численных моделях эффектов сильного легирования, утончения и модификации электрофизических параметров для тонкослойных областей сильнолегированного кремния.

Совершенствование методов машинной идентификации статических и динамических электрических характеристик и параметров приборных структур, фрагментов С³БИС, синтеза их электрических моделей на основе использования результатов физико-технологического моделирования является краеугольным камнем при исследовании корреляций технологических и электрических параметров, оптимизации в бездефектном проектировании С³БИС.

Многоуровневое моделирование интегральных микротехнологий, приборов и схем развивается на базе достижений физики полупроводников и полупроводниковых приборов, вычислительной математики, программирования и смежных специальных технических дисциплин. При осуществлении вычислительного эксперимента над сложнейшим объектом — С³БИС — не только используются последние достижения этих дисциплин, но и оказывается интенсивное непреходящее влияние на их собственное развитие, инициируя возникновение новых проблем и методов их решения.

Крайне актуально решение вопросов адаптации и развития программных средств многоуровневого (в первую очередь физико-технологического) моделирования элементно-технологической базы для современных и перспективных высокопроизводительных ЭВМ с учетом того, что требования к производительности и оперативной памяти ЭВМ при проектировании С³БИС в ближайшее время возрастут на несколько порядков. Требования адекватного моделирования субмикронных С³БИС (технологий, приборов, фрагментов) будут сопряжены с вычислительными ресурсами, сравнимыми с ресурсами при решении задач точного прогнозирования погоды или проектирования авиалайнеров и космических кораблей. Однако эти затраты должны окупиться при создании новых С³БИС с уникальными технико-экономическими показателями, которые не могли бы быть в принципе созданы вручную или даже на базе современных ЭВМ третьего и четвертого поколений. В конечном итоге работа в области многоуровневого моделирования с основополагающим акцентом на совершенствование методов физико-технологического расчета позволит проектировать «из первых принципов» и разрабатывать уникальные по сегодняшним меркам С³БИС, применение которых в перспективных высокопроизводительных ЭВМ в конце XX — начале XXI вв. откроет новую страницу научно-технического прогресса.

СПИСОК ЛИТЕРАТУРЫ

Основная

1. Агаханян Т. М. Основы транзисторной электроники. — М.: Энергия, 1974.
2. Степаненко И. П. Основы микроэлектроники. — М.: Советское радио, 1980.
3. Бонч-Бруевич В. Л., Калашников С. Г. Физика полупроводников. — М.: Наука, 1977.
4. Зи С. Физика полупроводниковых приборов. — М.: Мир, 1984.
5. Бродуай И., Мерей Д. Физические основы микротехнологии. — М.: Мир, 1985.
6. Баталов Б. В., Егоров Ю. Б., Русаков С. Г. Основы математического моделирования больших интегральных схем на ЭВМ. — М.: Радио и связь, 1982.
7. Носов Ю. Р., Петросянц К. О., Шилин В. А. Математические модели элементов интегральной электроники. — М.: Советское радио, 1976.
8. Самарский А. А. Теория разностных схем. — М.: Наука, 1983.

Дополнительная

1. Бубенников А. Н., Гуляев Ю. В., Мельников В. А. и др. Новые технологии изготовления и концепции разработки сверхскоростных больших интегральных схем супер ЭВМ 4-го и 5-го поколений: Обзор // Зарубежная радиоэлектроника. 1985. № 9.
2. Бубенников А. Н., Садовников А. Д. Проблемы физико-технологического моделирования субмикронных элементов ССБИС: Обзор // Зарубежная радиоэлектроника. 1987. № 3, 4.
3. Валиев К. А., Раков А. В. Физические основы субмикронной литографии в микроэлектронике. — М.: Радио и связь, 1984.
4. Польский Б. С. Численное моделирование полупроводниковых приборов. — Рига: Зинатне, 1986.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Адсорбция 74
Активная базовая область БТ 177, 214
Активный режим работы БТ 176, 246, 254
Акцептор 48, 49, 77, 105, 137, 139
Арсенид галлия 112, 115
Арсин 74
- База транзистора 59, 177, 186
Барьерная емкость перехода 234, 235, 253, 257, 269, 270
Барьер Шотки 117, 118
Биполярный транзистор (БТ) 7, 20, 93, 177, 214, 267, 290, 294
Больцмана уравнение 39, 40, 42, 107, 109, 125
Большая интегральная схема 7, 19, 276, 299
Бора диффузия 51, 53
Быстродействие БТ, ЛЭ 17, 19, 280, 283, 288, 293, 297, 299
Быстрый термический отжиг 60, 61
- Время задержки эмиттер — коллектор БТ 253, 255, 260, 269, 270, 291
— — коллектор — эмиттер БТ 257, 259
Время жизни неосновных носителей заряда 134, 135, 142, 182, 208
Вольт-амперная характеристика БТ 169, 180, 209, 247, 255
Высокий уровень инжекции в базе БТ 185, 195, 196, 197, 249
— — в коллекторе БТ 197, 201, 202, 203, 270, 273, 291, 293
Вытеснение тока статическое 215, 217, 219
— — динамическое 222, 223, 226, 265, 273
- Генерационный ток 135, 207, 208
Геометрический предел 17
Геометрические размеры 19
Глубина перехода 93, 95, 265, 268, 272
Глубокий коллектор 290
Граница раздела областей 70, 71, 77, 79, 80
Граничные процессы 69, 71, 72, 77, 79, 80
— условия 116, 117, 118, 143
Граничная частота БТ 188, 205
Гуммеля метод 59, 160, 168
- Двумерное технологическое моделирование 82, 84, 85
Двумерный физико-технологический расчет 243, 244, 245
Десорбция 74
Дефекты кристалла 11, 15, 29, 31, 32
Диффузия 44
— из неограниченного источника 50
— из ограниченного источника 50
— по вакансиям 45, 50, 51
— по междоузлиям 45, 50, 51, 60
Диффузия модель 45, 47
Диффузионная емкость эмиттера 239, 254, 270, 273
— — коллектора 238, 259, 273
Диффузионно-дрейфовая модель БТ 104, 105, 110, 113, 115
Диффузионный процесс 47, 49, 51, 53, 55, 57, 68
Донор 49, 77, 105, 138, 139, 142
- Емкость полная БТ 257, 259, 260
— — эмиттера БТ 253, 270, 273
Емкость подложки БТ 13, 260
- Зона валентная 120, 121
— запрещенная 120, 121, 123, 126
— проводимости 120, 121, 126
— примесная 120, 121
- Излучательная рекомбинация 132, 134, 142
Инверсный режим 176, 257, 258, 259
Интегральная микросхема 7
Интегральное соотношение управления зарядом 184, 185, 187, 192, 195
Ионное легирование 28, 31, 34, 37, 38, 39, 41, 42, 43
— — аналитическая модель 33, 34, 36
— — модель КУВ 40, 41, 42
Искаженные энергетических зон 119, 120, 121, 126, 139
Итерационные методы решения 158, 159, 162, 165, 166
— — уравнения Пуассона 163, 164, 166
— — уравнения непрерывности 163, 164, 166
- Кинетическое уравнение Больцмана 39, 40, 42, 107, 109, 125
МОП-транзистор 7
Коллекторный переход 93, 195, 197, 201, 203
Компоненты тока базы 205
Контактная разность потенциалов 234, 255, 263
Корреляция технологических и электрических параметров 268, 272, 274
Коэффициент диффузии 47, 50, 51, 53, 56, 57, 68
— линейный окисления 62, 63, 65, 66, 91
— обратной связи 247
— параболический окисления 62, 63, 65, 66, 91
Коэффициенты влияния параметров 267, 268
Кремний 9, 14, 19, 27, 61, 69, 85
- Линхарда, Шарфа, Шинотта модель 30, 34
Литографии моделирование 96, 97, 99
Ловушка рекомбинационная 132, 133
Логико-временное моделирование 21, 306, 307, 309
Логический элемент 16, 19, 279, 281, 283, 284, 288, 299
- Макромодели ЛЭ. фрагментов БИС 299, 300, 302, 305, 308
Масса эффективная носителя заряда 110, 112, 139
Межзонная рекомбинация 132, 133
Методы дискретизации ФСУ 146, 149, 151, 153, 155, 157
— идентификации параметров БТ 246, 257, 263, 266
- Моделирование
— граничных процессов 69
— логико-временное 306
— многоуровневое 21
— технологическое 27, 75
— физико-топологическое 102, 146
— электрическое 173, 275
Мышьяка диффузия 53, 55
- Наращивание эпитаксиальное 71, 73, 80, 81
Начальный ток коллектора 193, 195, 247, 257
Непрерывности уравнения 47, 104, 106
Нестационарный режим моделирования 169, 171, 264, 265
Нитрид кремния 9
Нормальное включение транзистора 176, 246, 257
- Объединение программ физико-технологического моделирования 240, 242, 245
Ограничения при физико-топологическом моделировании 111, 113, 115

Одномерная технологическая модель БТ 77, 79, 80
— физико-технологическая модель БТ 243, 244
— физико-топологическая модель БТ 167, 168
Оже коэффициенты 134, 143
Окисления модель 61, 62, 63, 67, 68, 69, 91
Оптимизация быстродействия ЛЭ 276, 280, 283, 284, 287, 296
— физико-технологическая БТ 289, 291, 293, 296, 299
Пассивная база БТ 214, 228
Переходной процесс в БТ 237, 264, 266, 273
Поверхностная рекомбинация 116, 117, 249, 250, 258
Поверхностное сопротивление активной базы 168, 263, 296, 297
Подвижность основных носителей заряда 138, 139, 140, 144
Поликремниевый БТ 13, 14, 294, 297
Поликремний 87, 89, 91, 92, 143, 144
Предел интеграции 15, 17, 18
Пуассона уравнение 104, 105, 106, 164, 165
Размеры кристалла 11, 16, 17
Распределение Максвелла — Больцмана 122, 123, 130
Распределенная модель БТ 212, 214, 222, 232, 234
Резист негативный 97, 98
— позитивный 97, 98
Рекомбинация модель 132, 133, 134
Рекомбинационный ток 168, 169, 208, 209, 228, 247, 248, 257, 258
— член 104, 134
Самосовмещение областей 12, 13, 298
Сверхсамосовмещение 13, 289, 298
Сверхкэроновые сверхБИС—С²БИС 7, 19, 289, 299
Система линейных алгебраических уравнений 162, 163, 165, 172

Скорость насыщения носителей 198, 199, 203, 291
Смешанное моделирование БИС 310, 311
Соотношение Эйнштейна 47, 127, 129
Сопротивление базы БТ 217, 219, 225, 227, — коллектора 228, 229, 233, 256
— эмиттера 228, 238, 252
Степень интеграции 7, 10, 17, 19, 299
Сурьмы диффузия 55, 56

Термогенерация 145
Технологическое моделирование 21, 27, 75, 84, 87, 93, 95
Травление анизотропное 99, 100, 101
— изотропное 99, 100, 101

Формальные макромодели ЛЭ 301, 302
Фосфора диффузия 57, 58, 60

Численные методы решения окисления 71, 84, 85
— — — уравнения диффузии 77, 79, 83, 84
— — — уравнений полупроводника 158, 159, 165
— — — эпитаксии 80, 81

Ширина запрещенной зоны 120, 123, 126, 141

Эквивалентные схемы БТ 220, 228, 231, 234, 238, 247, 258
Электрофизические параметры 135, 137, 139, 140, 143
Эмиттерный переход 94, 269, 272
Эмиттерно-связанный логический (ЭСЛ) элемент 279, 281, 283, 288
Эффект статического вытеснения 215, 217, 219
— динамического вытеснения 222, 223, 226, 265, 273
— пробоя перехода 256
— прокола базы 290
— расширения базы 297, 201, 273, 291, 293
— сильного легирования 122, 123, 126, 129, 139, 121, 142
— эмиттерного выталкивания 59
— Эрди 194, 247

ОГЛАВЛЕНИЕ

Предисловие	3
Введение	4
Глава 1. Пути и перспективы развития кремниевой интегральной электроники и сквозного многоуровневого моделирования БИС	7
1.1. Кремниевая компьютерная технология	7
1.2. Перспективы повышения степени интеграции и быстродействия кремниевых БИС, их ограничения и пределы	10
1.3. Проблемы многоуровневого моделирования и особенности физико-технологического моделирования кремниевых БИС	20
Глава 2. Модели технологических процессов изготовления БИС	27
2.1. Поколения программ технологического моделирования БИС	27
2.2. Модели процессов ионного легирования	28
2.3. Модели диффузионных процессов	44
2.4. Модели термического окисления	61
2.5. Моделирование граничных процессов	69
2.6. Моделирование процесса эпитаксии	72
Глава 3. Программное обеспечение моделирования технологии изготовления БИС	75
3.1. Назначение технологических программ	75
3.2. Программная реализация одномерного моделирования технологических процессов изготовления БИС	77
3.3. Двумерное моделирование технологии изготовления БИС	82
3.4. Моделирование технологических процессов с использованием поликремния	87
3.5. Применение программ технологического моделирования	93
3.6. Моделирование процессов образования поверхностных конфигураций	95
Глава 4. Физико-топологические модели полупроводниковых структур	102
4.1. Назначение и проблемы численного моделирования полупроводниковых структур	102
4.2. Основные уравнения полупроводника	104
4.3. Ограничения физико-топологических моделей. Возможные подходы к моделированию	111
4.4. Граничные условия в численном моделировании полупроводниковых структур	116
4.5. Модель искажений энергетических зон в сильнолегированном кремнии	119
4.6. Синтез моделей переноса носителей заряда полупроводниковых структур с учетом эффектов сильного легирования	124
4.7. Модели процессов рекомбинации — генерации	131
4.8. Модели электрофизических параметров для физико-топологического расчета полупроводниковых структур	135
4.9. Модели электрофизических процессов в поликремнии	143
4.10. Модели процессов термогенерации в сильнолегированных полупроводниковых структурах	145

Глава 5. Методы и алгоритмы численного физико-топологического моделирования полупроводниковых структур	146
5.1. Методы дискретизации ФСУ	146
5.2. Схемы дискретизации ФСУ полупроводниковых структур	155
5.3. Общие методы решения систем конечно-разностных уравнений полупроводниковых структур	158
5.4. Итерационные методы решения линейаризованных уравнений Пуассона и непрерывности	162
5.5. Программы стационарного физико-топологического моделирования полупроводниковых структур	167
5.6. Программы нестационарного двумерного физико-топологического моделирования	169
Глава 6. Электрические модели интегральных приборов для машинного расчета БИС	173
6.1. Требования к электрическим моделям. Классификация электрических моделей интегральных приборных структур	173
6.2. Обобщенные выражения для токов и коэффициента передачи тока интегральных БТ	177
6.3. Интегральное соотношение управления зарядом для синтеза электрических моделей БТ	184
6.4. Основы построения электрических иерархических моделей БТ с управлением зарядом	188
6.5. Описание физических эффектов в БТ электрической моделью с управлением зарядом	193
6.6. Компоненты тока базы. Уменьшение коэффициента передачи тока в микрорежиме	205
6.7. Моделирование многомерных физических процессов в быстродействующих БТ	212
6.8. Синтез электрических моделей БТ. Библиотека иерархических электрических моделей БТ	227
Глава 7. Объединенное физико-технологическое моделирование. Идентификация и корреляция параметров элементов быстродействующих БИС	240
7.1. Методология машинного расчета и идентификации характеристик и параметров интегральных БТ с помощью программ технологического и физико-топологического моделирования	240
7.2. Объединение программ технологического и физико-топологического моделирования интегральных БТ в единый комплекс физико-технологического расчета	242
7.3. Идентификация параметров электрической модели БТ в активном нормальном режиме с помощью программ ФТ-моделирования	246
7.4. Идентификация параметров электрической модели БТ в инверсном и насыщенном режимах с помощью программ ФТ-моделирования	256
7.5. Расчет на ЭВМ статических и динамических параметров БТ БИС в различных температурных режимах	261
7.6. Расчет на ЭВМ переходных процессов в транзисторной схеме с общей базой с помощью программы двумерного нестационарного моделирования	263
7.7. Исследование на основе физико-технологического моделирования влияния изменения технологических параметров на электрические характеристики БТ БИС	267
Глава 8. Оптимизация и макро моделирование интегральных быстродействующих элементов и схем	275
8.1. Схемно-технологическая оптимизация логических элементов БИС. Адаптация структурно-технологических параметров БТ к потребляемой мощности элементов БИС	275

8.2. Выбор электрического режима и оптимизация элементов БИС	283
8.3. Основы физико-технологического синтеза субмикронных тонко- слойных элементов биполярных С ³ БИС	289
8.4. Формальные и адекватные макромодели быстродействующих логических элементов и фрагментов БИС	299
8.5. Логико-временное макромоделирование БИС	306
8.6. Смешанное логико-временное, схемотехническое и физико-тех- нологическое моделирование БИС	310
Заключение	313
Список литературы	315
Предметный указатель	316

Учебное издание

Бубенников Александр Николаевич

**МОДЕЛИРОВАНИЕ ИНТЕГРАЛЬНЫХ
МИКРОТЕХНОЛОГИЙ, ПРИБОРОВ И СХЕМ**

Зав. редакцией В. И. Трефилов. Редактор Т. И. Артемова. Мл. редактор С. А. Пацева. Художник В. В. Гарбузов. Художественный редактор Т. М. Скворцова. Технический редактор З. В. Нуждина. Корректор Г. И. Кострикова.

ИБ № 8029

Изд. № ЭР-475. Сдано в набор 22.02.89. Подп. в печать 18.07.89. Т-13914. Формат 60×88^{1/8}. Бум. офс. № 2. Гарнитура литературная. Печать офсетная. Объем 19,60 усл. печ. л. 19,60 усл. кр.-отт. 20,77 уч.-изд. л. Тираж 25 000 экз. Зак. № 1141. Цена 1 р.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Неглинная ул., д. 29/14.

Московская типография № 8 Союзполиграфпрома при Государственном комитете по делам издательств, полиграфии и книжной торговли. 101898, Москва, Центр, Хохловский пер., 7.