

HETEROSTRUCTURE LASERS

H. C. CASEY, JR.

Bell Laboratories
Murray Hill, New Jersey

M. B. PANISH

Bell Laboratories
Murray Hill, New Jersey

PART A

FUNDAMENTAL PRINCIPLES

Academic Press

New York San Francisco London 1978

621.343.8
#33

Х.Кейси, М.Паниш

ЛАЗЕРЫ НА ГЕТЕРОСТРУКТУРАХ

В 2-Х ТОМАХ

1
ОСНОВНЫЕ ПРИНЦИПЫ

83

88

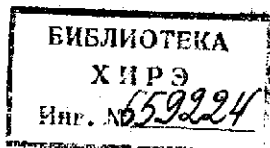
Перевод с английского

А. Е. ДРАКИНА

под редакцией

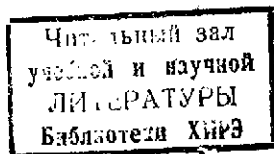
д-ра физ.-мат. наук

П. Г. ЕЛИСЕЕВА



198

2004



621.375.8

Лазеры на гетероструктурах — наиболее важная разновидность инжекционных лазеров, привлекающая все большее внимание специалистов благодаря перспективам их широкого применения в системах оптической связи, в вычислительной технике, голографии и других областях.

Настоящий том 1 монографии американских специалистов Кейси и Паниша содержит обзор фундаментальных физических явлений, на которых основана работа гетеролазеров.

Для научных работников, инженеров, аспирантов, студентов соответствующих специальностей.

Редакция литературы по физике

1704050000

К 20405-051 51-81, ч.1
041(01)-81

© 1978, by Bell Telephone Laboratories, Inc.
© Перевод на русский язык, «Мир», 1981

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

В свое время полупроводники помогли осуществить преобразование элементной базы электроники, что имело исключительно важные последствия для современной научно-технической революции: именно полупроводники позволили создать миниатюрные и экономичные транзисторы, а позднее и другие твердотельные приборы и интегральные схемы. Подобно этому применение полупроводников в лазерной технике привело к созданию особой разновидности лазеров, отличающихся малыми размерами, экономичностью и малой инерционностью, — инжекционных лазеров, которым предсказывается большое будущее, в особенности в технике оптической связи, основанной на использовании волоконных световодов. Инжекционные лазеры незаменимы в волоконно-оптических схемах связи, рассчитанных на передачу информации с большой скоростью и на большое расстояние. Как низковольтный полупроводниковый прибор инжекционный лазер совместим со схемами накачки и управления на основе полупроводников и, более того, может быть включен вместе с ними в монолитный оптоэлектронный модуль. Как оптический прибор лазер этого типа легко согласуется с волоконными световодами, причем выполняет одновременно функции и генератора несущей частоты, и ее модулятора. Интересны также другие сферы практических применений инжекционных лазеров — спектроскопия и спектроскопические датчики, системы оптической автоматики и памяти, устройства вычислительной техники и т. д.

В связи с вовлечением новых специалистов в работу по созданию и применению инжекционных лазеров и в связи с увеличением темпов этой работы возрастает потребность в книгах, излагающих на современном уровне принципы их действия, устройство и методы технологии. Двухтомная монография Кейси и Паниша призвана удовлетворить эту потребность. Она посвящена лазерам на основе гетероструктур (гетеролазерам). Это важнейшая разновидность инжекционных лазеров, доминирующая во всех областях применения, где неприемлемо глубокое охлаждение активного элемента. Использование гетероструктур дает ряд преимуществ: например пороговый ток для

начала генерации когерентного излучения в гетеролазерах удалось снизить примерно на два порядка величины.

Данную монографию отличает глубокая проработка всех основных вопросов и их подробное, близкое к исчерпывающему изложение. Путем отбора материала из обширной научной периодики, проведения большого числа иллюстративных численных расчетов авторам удалось создать целостную и однородную по глубине картину, в которой представлены и фундаментальные идеи и большое число конкретных подробностей. Эта книга пригодна и для учебных целей (для ознакомления новичков с физическими принципами, на которых основана работа гетеролазеров) и для каждодневного использования специалистами (как собрание полезнейшей информации — схем, методов расчета, графиков, сводных таблиц). Следует заметить, что авторы с успехом справились с задачей стройного и равноценного изложения далеких друг от друга областей физики, объединенных в проблеме гетеролазеров, таких, как, скажем, волноводы, фазовые диаграммы, дислокации и т. д.

Достоинством книги является наиболее полное описание основ современной технологии гетеролазеров, представляющее широкий интерес, поскольку оно во многом может быть распространено на технологию других приборов. Эта передовая технология стремительно развивалась в последнее десятилетие, причем заметный вклад в это развитие внесли авторы книги, активно работающие в известной лаборатории «Белл» (Мюррей-Хилл, США). В частности, д-р М. Паниш — один из создателей первого на Западе гетеролазера непрерывного действия при комнатной температуре.

В силу своих научных интересов и, возможно, вследствие ограниченного объема книги авторы исключили из нее некоторые важные разделы, например раздел о динамике лазерного излучения. Однако это вполне компенсируется сильными сторонами книги, о которых упомянуто выше.

Ссылки на литературу в подавляющем большинстве относятся к публикациям западных и японских авторов. Чтобы дать более полное представление о развитии теоретических и практических разработок в этой области, мы даем список литературы, включающий обзорные работы и монографии на русском языке. Советскими учеными внесен важный вклад в фундаментальные исследования, приведшие к созданию полупроводниковых лазеров и приборов на основе гетеропереходов (Ленинские премии за 1964 и 1972 гг.).

В русском издании исправлены опечатки, обнаруженные при переводе книги, а также замеченные авторами в американском издании и любезно присланные редактору.

ЛИТЕРАТУРА

1. Басов Н. Г. Полупроводниковые квантовые генераторы. — Усп. физич. наук, 1965, т. 85, вып. 4, с. 585.
2. Басов Н. Г., Попов Ю. М. Полупроводниковые лазеры. — Физ. техн. полупр., 1967, т. 1, вып. 11, с. 1593.
3. Басов Н. Г., Никитин В. В., Семенов А. С. Динамика излучения инжекционных полупроводниковых лазеров. — Усп. физич. наук, 1969, т. 97, вып. 4, с. 561.
4. Алферов Ж. И. Гетероструктуры и их применение в оптоэлектронике. — Вестник АН СССР, 1976, вып. 7, с. 28.
5. Алферов Ж. И., Гарбузов Д. З., Долгинов Л. М., Елисеев П. Г., Мильвидский М. Г. Многокомпонентные полупроводниковые твердые растворы и их применение в оптоэлектронике. — Вестник АН СССР, 1978, т. 9, вып. 4, с. 31.
6. Елисеев П. Г. Инжекционные лазеры на гетеропереходах. В кн.: Квантовая электроника. Сб. статей. Под ред. акад. Н. Г. Басова. — М.: Сов. радио, 1972, вып. 6.
7. Ривлин Л. А. Динамика излучения полупроводниковых квантовых генераторов. — М.: Сов. радио, 1976.
8. Елисеев П. Г. Полупроводниковые лазеры и преобразователи. В сб.: Итоги науки и техники, сер. Радиотехника, т. 14, ч. 1. — М.: ВИНТИ, 1978.
9. Богданкевич О. В., Дарзек С. А., Елисеев П. Г. Полупроводниковые лазеры. — М.: Наука, 1976.
10. Елисеев П. Г. Проблема надежности и физика деградационных явлений в полупроводниковых лазерах. В сб.: Итоги науки и техники, сер. Радиофизика, т. 14, ч. 2. — М.: ВИНТИ, 1978.
11. Долгинов Л. М., Елисеев П. Г., Исмаилов И. Инжекционные излучательные приборы на основе многокомпонентных полупроводниковых твердых растворов. В сб.: Итоги науки и техники, сер. Радиотехника, т. 21, с. 3.

ПРЕДИСЛОВИЕ

Эволюция полупроводниковых лазеров от ранних гомолазеров на основе GaAs к множеству разнообразных гетеролазеров и современному промышленному производству этих приборов потребовала сотрудничества ученых и инженеров, работающих в различных областях физики. Нужно было объединить усилия разных ученых, чтобы установить фундаментальные принципы, на которых основано действие полупроводниковых лазеров, освоить их изготовление и изучить рабочие характеристики. В дальнейшем при распространении исследований на другие материалы и структуры также потребуются широкие знания в областях, находящихся на стыках отдельных дисциплин. Настоящая книга представляет собой учебную монографию, причем большое внимание в ней уделяется именно комплексному характеру проблемы. Следует заметить, что мы включили в рассмотрение только те вопросы, которые достаточно хорошо изучены, чтобы сделать книгу удобной для преподавательской работы. Применения полупроводниковых лазеров не рассматриваются.

Каждый крупный раздел начинается с изложения основных законов, которым подчиняются рассматриваемые явления. Выражения, описывающие работу гетеролазеров, выводятся из фундаментальных законов, и приводятся конкретные численные примеры. Специалист по выращиванию кристаллов мог не изучать вопросов распространения электромагнитного излучения или усиления в лазере, в то время как физик, интересующийся этими предметами, ранее мог не встречаться с вопросами фазового равновесия или роста кристаллов. Поэтому выводы выражений даются с соответствующими определениями и достаточно подробно, чтобы позволить читателю изучить незнакомый предмет.

В книге приводятся как строгие, так и приближенные решения уравнений. В большинстве случаев на основе полученных выражений легко сделать численный расчет с помощью микрокалькулятора или простых программ для вычислительных машин. Наличие миникомпьютера с графическим выходом позволило нам проиллюстрировать численные результаты графиками.

Поэтому читатель может либо следовать детальным выкладкам, либо получить общее представление, пользуясь многочисленными иллюстрациями. Численные оценки основаны на данных для гетероструктуры GaAs—Al_xGa_{1-x}As. В настоящее время только для гетероструктур системы Al_xGa_{1-x}As имеется достаточно данных, чтобы проводить численные оценки полученных выражений.

При написании книги, посвященной быстро развивающемуся направлению, которое находится на стыке различных дисциплин, встречается ряд специфических трудностей. Одна из них связана с системой обозначений. Дело в том, что для представления различных величин применяются одни и те же символы. Чтобы не менять привычных обозначений, мы пытались, где это возможно, вносить лишь незначительные изменения. Например, буквой *n* обычно обозначают и концентрацию электронов и показатель преломления. Чтобы избежать путаницы, мы обозначаем показатель преломления символом \bar{n} . Вторая проблема состоит в том, что статьи по гетеролазерам появляются почти ежедневно. Они не только дают все новые данные, но часто меняют наше представление о разных вопросах. Например, изложение в гл. 4 пришлось изменить после того, как было установлено правильное относительное расположение Г-, L- и X-минимумов зоны проводимости в GaAs. Мы пытались сделать нашу книгу достаточно фундаментальной, чтобы появляющиеся публикации по гетеролазерам исходили из изложенных в ней принципов. Дальнейшие исследования наверняка изменят некоторые из наших сегодняшних представлений. Третья проблема — огромное число публикаций по полупроводниковым лазерам. Вместо того, чтобы пытаться включить в библиографию все статьи, мы даем ссылки на основные работы, чтобы читатель мог начать библиотечный поиск по интересующей его теме. И, наконец, отсутствие студентов в промышленной лаборатории лишило нас аудитории, на которой мы могли бы испытать различные варианты изложения материала. Однако «Белл лабораториз» предоставила нам возможность общаться с широким кругом экспертов по многим различным предметам. Благодаря этому в книгу включены темы, которые в противном случае были бы опущены.

Мы благодарим наших многочисленных коллег как из «Белл Лэборатриз», так и из других организаций за большую помощь, оказанную при подготовке и написании этой книги. Подход, используемый в гл. 2 при выводе выражения для расходимости луча, был предложен Л. Левином. Дискуссии с Дж. Л. Мерцем были полезны на начальном этапе вывода уравнений, описывающих распространение волн. В. Т. Цань помог нам достигнуть более полного понимания ряда выводов и сделал полезные замечания по поводу изложения вопроса о распределенной обратной связи. Ф. Стерн дал подробные рекомендации при написании части гл. 3, связанной с выводом оптического матричного элемента, а также предоставил нам вычисленные (неопубликованные) спектральные зависимости коэффициента усиления. Дискуссии с В. Б. Джойсом помогли при написании раздела об излучении абсолютно черного тела, а дискуссии с Дж. Р. Брюсом — при выводе выражения для длины экранирования. М. Лэкс сделал полезные замечания относительно правильного использования «золотого правила» Ферми. Обсуждения с В. Б. Джойсом оказались полезны при изложении в гл. 4 вопроса о задании уровня Ферми. Мы хотим поблагодарить Р. Дингла за предоставление нам неопубликованных данных по зависимости ширины запрещенной зоны $Al_xGa_{1-x}As$ от состава и Д. Е. Аспнеса за многочисленные обсуждения зонной структуры GaAs и $Al_xGa_{1-x}As$. Замечания М. Айлджемса были полезны при выводе в гл. 6 термодинамических соотношений для уравнения ликвидуса. Дискуссии с К. Д. Тэрмондом помогли при изложении вопроса о фазовом равновесии, а дискуссии с А. Ю. Чо внесли свой вклад в описание эпитаксии из молекулярных пучков. Обсуждения вопросов с В. Б. Джойсом и Т. Л. Паоли помогли уточнить многие идеи и представления в гл. 7. В. Т. Цань предложил приближение, используемое при объединенном рассмотрении растекания тока и поперечной диффузии в лазерах с полосковой геометрией. Многочисленные дискуссии, в которых приняли участие Б. К. Де Лоач, Р. В. Диксон, Р. Л. Хартман, Т. Л. Паоли, К. Х. Генри, Р. М. Петрофф и Д. В. Лэнг, помогли нам при изложении вопроса о деградации. М. Урбано внес полезные предложения по материалам всей книги.

Нам особенно хотелось бы выразить признательность г-же Кэрол Айгенброд и г-же Сильвии Липтон за перепечатку многочисленных черновишков книги.

ВВЕДЕНИЕ

§ 1. ВВОДНЫЕ ЗАМЕЧАНИЯ

В этой главе мы расскажем о некоторых ранних исследованиях по инжекционным лазерам и познакомим читателя с основными понятиями физики гетеролазеров. Следует заметить, что в первом томе основное внимание уделено описанию природы волноводного эффекта (гл. 2), усиления (гл. 3) и ограничения для носителей тока в гетероструктурах (гл. 4). В виде примера используется наиболее детально изученная система $\text{GaAs—Al}_x\text{Ga}_{1-x}\text{As}$. Во втором томе обсуждаются другие возможные химические системы (гл. 5) и методы эпитаксиального выращивания структур (гл. 6), а также соответствующая химическая термодинамика. Приводятся реальные рабочие характеристики и описываются методы изготовления гетеролазеров на основе $\text{GaAs—Al}_x\text{Ga}_{1-x}\text{As}$ (гл. 7). Обсуждаются свойства лазеров с широким контактом, в которых ограничение потока инжектированных носителей тока и волноводное распространение излучения происходят перпендикулярно плоскости $p-n$ -перехода, а также рассматриваются лазеры с полосковой геометрией, в которых ограничение тока происходит дополнительно еще и вдоль плоскости $p-n$ -перехода. И, наконец, обсуждаются физические процессы, влияющие на срок службы лазеров, а также возможность того, что он превысит 10^5 ч (гл. 8).

§ 2. ПОЛУЧЕНИЕ ВЫНУЖДЕННОГО ИЗЛУЧЕНИЯ В ИНЖЕКЦИОННЫХ ЛАЗЕРАХ И ВОЗМОЖНОСТЬ ИХ РАБОТЫ В НЕПРЕРЫВНОМ РЕЖИМЕ ПРИ КОМНАТНОЙ ТЕМПЕРАТУРЕ

**Вынужденное излучение, генерируемое
в $p-n$ -переходах**

В начале гл. 1 мы дадим читателю только ретроспективный обзор ранних работ, так как подробное описание физики и технологии полупроводниковых лазеров дано в последующих главах. Предположения о том, что полупроводники могут быть

использованы как лазерные материалы, были сделаны в 1958—1961 гг. [1—4]. Однако количественные соотношения, описывающие возможность достижения лазерного эффекта в полупроводнике, не были известны до тех пор, пока Бернар и Дюрафур [5] не сформулировали необходимое условие возникновения вынужденного излучения в полупроводнике. Они показали, что расстояние между квазиуровнями Ферми, соответствующими неравновесным концентрациям электронов и дырок, должно превосходить энергию излучаемого фотона. Это условие выводится в гл. 3. Бернар и Дюрафур предположили, что подходящими полупроводниками наряду с другими являются соединения элементов III и V групп GaAs и GaSb. Эти материалы были впервые описаны Велкером [6, 7] как потенциально полезные для полупроводниковых приборов.

В начале 60-х годов быстро увеличивался объем исследований по лазерам, и изучение приборов на $p-n$ -переходах представляло собой одну из наиболее быстро развивающихся областей электроники. Не удивительно, что эти исследования сосредоточивались на получении лазерного эффекта посредством инжекции неравновесных электронов через $p-n$ -переход. Интересно отметить, что по сообщению Бардина [8] в неопубликованных записках, относящихся к 1953 г., фон Нейман предлагал возможность усиления света путем использования вынужденного излучения в полупроводнике с накачкой, осуществляемой инжекцией через $p-n$ -переход.

Интересное описание некоторых ранних работ, приведших к первой серии сообщений о получении инжекционных лазеров, дано Холлом [9] из лаборатории «Дженерал электрик» в Скенектеди. До лета 1962 г. он очень скептически относился к возможности осуществления полупроводникового лазера. В то время причины для его скептицизма казались убедительными. Известные тогда лазеры требовали больших оптических путей, что казалось трудно совместимым с сильным поглощением на свободных носителях в полупроводниках. Оптические переходы в полупроводниках происходят в значительно более широком диапазоне энергий, чем в обычных лазерах, для которых характерны узкополосные дискретные переходы, и, что наиболее важно, излучательную рекомбинацию в полупроводниках всегда считали очень неэффективной. Холл отметил, что эти положения были пересмотрены после того, как Бернар и Дюрафур [5] внесли ясность в концепцию вынужденного излучения и была получена эффективная излучательная рекомбинация в GaAs. На конференции по твердотельным приборам в июле 1962 г. Кейес и Кунст [10] из лаборатории Линкольна доложили, что квантовая эффективность электролюминесценции в GaAs при 77 К может достигать значений, близких к 100%.

Между тем еще в 1961 г. в лаборатории ИБМ Лэндауэр, Лэшер, Думке и Кейес [9] рассматривали возможность получения лазерного эффекта в полупроводнике. Лэшер считал, что волновое распространение мод ведет к уменьшению потерь на свободных носителях. Думке [11] указывал на важность использования прямозонного полупроводника. О высокой эффективности электролюминесценции в GaAs при 77 К они узнали в марте 1962 г., когда лабораторию ИБМ посетил Мэйбург и рассказал о своих работах с диодами на GaAs. Мэйбург также доложил о своих результатах в сообщении на совещании Американского физического общества в марте 1962 г., однако, по-видимому, это сообщение прошло незамеченным научной общественностью. Холл [9] пишет, что только после того, как сотрудники ИБМ услышали доклад Кейеса и Куиста [10] на конференции по твердотельным приборам, у них сильно возрос интерес к использованию диодов на $p-n$ -переходах в качестве полупроводниковых лазеров.

Сообщения Панкова и Массули [12], Кейеса и Куиста [10], а также Мэйбурга сыграли, очевидно, основную роль в повышении интереса к полупроводниковым лазерам, хотя в этих работах рассматривалась всего лишь некогерентная электролюминесценция. В значительной степени такому повышению интереса способствовали также широкие дискуссии между участниками конференции по твердотельным приборам. Одним из участников конференции был Холоньяк из лаборатории «Дженерал электрик» в Сиракузах. Он изучал туннельный эффект в кристаллическом твердом растворе GaP_xAs_{1-x} и, кроме того, наблюдал при 77 К излучение в видимой области спектра из смещенных в прямом направлении $p-n$ -переходов, созданных в этом материале. Холоньяк вернулся в Сиракузы, пытаясь понять, как использовать внешний резонатор для создания обратной связи в инжекционном лазере [13].

В то время как многие исследователи изучали возможность создания инжекционных лазеров, Наследов и др. [14] в начале 1962 г. сообщили о незначительном сужении спектра электролюминесценции диода на арсениде галлия при 77 К и плотности тока $1,5 \cdot 10^3$ А/см². В исследуемой ими схеме резонатор отсутствовал, и не было ясно, возникало ли вынужденное излучение. Холл [9] решил использовать резонатор из зеркал, которыми являлись бы полированные грани кристалла GaAs, перпендикулярные плоскости диффузионного $p-n$ -перехода. К сентябрю 1962 г. Холл и др. [15] с полной определенностью получили когерентное излучение из смещенного в прямом направлении $p-n$ -перехода в GaAs при 77 К. Вывод о возникновении вынужденного излучения был основан на узости спектра излучения на длине волны около 0,84 мкм и поведении картины даль-

него поля излучения. Такие лазеры, состоящие из одного полупроводника, принято теперь называть лазерами на гомоструктурах, или гомолазерами. В фирме ИБМ Натан с сотрудниками [16] наблюдали сужение линии при 77 К в диодах на GaAs с травленной мезаструктурой. Как сообщал Натан Холлу [9], первый же диод, который они исследовали, обладал совершенно случайно свойствами резонатора, в котором наблюдалось характерное для инжекционного лазера сужение линии. Их статья [16] была представлена в октябре 1962 г.

Развивая обсуждавшиеся на конференции идеи, Холоньяк изготовил диоды с $p-n$ -переходом, полученным диффузией Zn в слой $\text{GaP}_x\text{As}_{1-x}$ n -типа, выращенные на подложках методом химического осаждения из газовой фазы. Следуя выводам Холла, он использовал резонатор, образованный самим кристаллом [13]. Холоньяк и Бивекуа [17] вскоре после Холла и др. получили инжекционный лазер, работающий при 77 К. [Новизна этой работы состояла в том, что впервые для получения инжекционного лазера был использован кристаллический твердый раствор элементов III и V групп и выбран такой состав этого раствора, который давал излучение в видимой области спектра (0,71 мкм). Вслед за статьей Холоньяка и Бивекуа, представленной на рассмотрение в октябре, в начале ноября последовала статья Куиста с сотрудниками из лаборатории Линкольна [18]. Они сообщали о получении лазерного эффекта как при 4,2 К, так и при 77 К. Отличительными признаками полученного эффекта являлись характерная зависимость выходной мощности излучения от тока и узость спектра излучения.]

За исключением диодов, исследованных Натаном и др. [16], эти первые инжекционные лазеры имели форму параллелепипеда с прямоугольным или трапецеидальным основанием. Чтобы приготовить их, нарезали пластинки GaAs (или $\text{GaP}_x\text{As}_{1-x}$) и полировали две их параллельные грани. Для получения этих пластинок использовали материал n -типа, в котором предварительно диффузией примеси p -типа был образован $p-n$ -переход. Плоскость $p-n$ -перехода была перпендикулярна полированным граням параллелепипеда. Таким способом создавался резонатор типа небольшого интерферометра Фабри — Перо. Обратная связь в резонаторе осуществлялась так же, как это обычно происходит в лазерах, поскольку полированные грани кристалла ведут себя как частично пропускающие зеркала. Лазерный эффект появляется при селективном усилении одной или более мод резонатора.]

Холоньяк [13] заметил, что лучшим способом формирования резонатора Фабри — Перо в полупроводниковом лазере может быть получение плоскопараллельных зеркал путем скалывания кристалла вдоль параллельных кристаллических плоскостей.

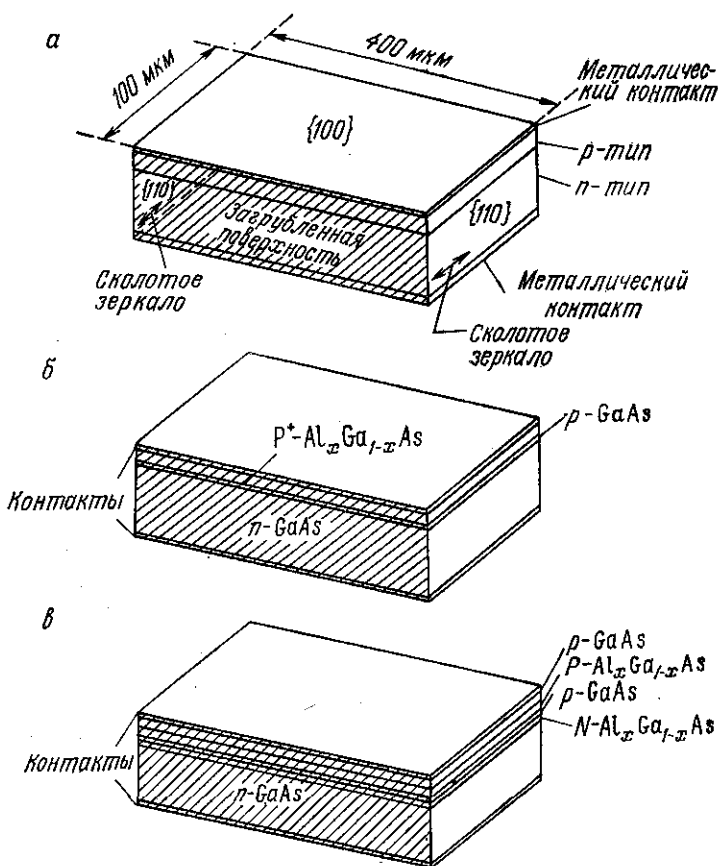


Рис. 1.2.1. Лазер на гомоструктуре (а); сколотые грани полупроводникового кристалла действуют как зеркала резонатора Фабри — Перо; типичные значения длины и ширины 400 и 100 мкм соответственно. Лазер на односторонней гетероструктуре (б); толщина слоя $p - GaAs$ составляет ~ 2 мкм. Лазер на двусторонней гетероструктуре (в); толщина активного слоя $p - GaAs$ обычно меньше 0,5 мкм.

Однако использовавшийся им материал плохо колелся. Поэтому Холоньяку пришлось работать с полированными кристаллами, чтобы в стремительном развитии исследований, происшедших летом 1962 г., стать одним из первых создателей инжекционных лазеров. Бонд и др. [19] первыми сообщили в 1963 г. о применении в качестве зеркал инжекционного лазера граней, полученных при скалывании кристалла. В настоящее время плоскость $p - n$ -перехода обычно параллельна грани {100} кристалла, так что плоскости естественного скола {110} пер-

пендикулярны плоскости $p-n$ -перехода. На рис. 1.2.1, *a* изображен гомолазер с зеркалами, образовавшимся при скалывании. Приведенные на рисунке размеры типичны как для гомолазеров, так и для других, более сложных инжекционных лазеров.

Сразу же после создания гомолазеров на $p-n$ -переходе наступило время повышенной активности в исследовании полупроводниковых лазеров. На конференции по твердотельным приборам в июне 1963 г. работала секция по лазерным диодам, на которой было доложено о получении лазеров на $p-n$ -переходах в InAs и $\text{Ga}_x\text{In}_{1-x}\text{As}$ [20], а также в InP [21]. Были представлены работы, в которых рассматривались вопросы об эффективности [22] и пороге генерации [23]. Влияние температуры на пороговый ток было рассчитано Мэйбургом [24] и исследовано Энгелером и Гарфинкелом [25], а также Пилкуном и др. [26]. Полученная экспериментально зависимость T^3 не согласовывалась с предсказанной зависимостью $T^{3/4}$. В 1964 г. было сообщено о создании лазера на $\text{InP}_x\text{As}_{1-x}$ [27]. В 1963 и 1964 гг. появилось уже девять обзорных статей [28—36]. С 1964 по 1966 г. были получены лазеры на ряде новых материалов, в том числе на соединениях элементов IV и VI групп [37]. В течение этого периода появились также многочисленные статьи о конструкциях приборов, влиянии различных параметров на оптические свойства и о возможных применениях инжекционных лазеров. Обзор этих работ содержится в статье Натана [38].

Общей и весьма неприятной чертой инжекционных гомолазеров являлись очень высокие значения плотности порогового тока ($\geq 50\,000$ А/см²) при комнатной температуре. Большинство исследований было проведено при температуре жидкого азота (77 К) или еще более низких температурах. Создание лазера, работающего в непрерывном режиме при комнатной температуре, оказалось невозможным, хотя в 1967 г. Даймент и Д'Азаро [39] при соответствующем выборе теплоотвода смогли получить непрерывный режим работы при температурах вплоть до 205 К. Инжекционный ток протекал через узкую полоску, проходящую вдоль всей длины лазера. Такая конструкция уменьшала общий ток, текущий через лазер, и улучшала теплоотвод. Как отмечается в гл. 7, полосковая геометрия в том или ином варианте стала позднее важной чертой, присущей почти всем гетеролазерам, имеющим практическое значение. Обычным режимом работы гомолазеров были повторяющиеся, очень короткие импульсы (≤ 1 мкс) с большой скважностью ($> 10^3$). Оглядываясь назад, мы сейчас понимаем, что очень высокие значения плотности порогового тока при комнатной температуре являлись характерным свойством гомолазеров.

Понижение плотности порогового тока при комнатной температуре

Интерес к исследованиям по гомолазерам, которые стали интенсивно проводиться в 1962 г., после 1965 г. начал падать, поскольку к этому времени не было достигнуто почти никакого снижения плотности порогового тока. В 1963 г. Кремер [40] предположил, что инжекционные лазеры с улучшенными характеристиками могут быть получены при использовании структуры, в которой слой полупроводника с относительно узкой запрещенной зоной находится между двумя слоями из более широкозонного полупроводника. Контакт между двумя такими различными полупроводниками называется гетеропереходом. Предполагалось, что эффективная излучательная рекомбинация возникает в узкозонном материале. (Статья Кремера, а также аналогичное предположение Алферова и Казаринова [41], высказанное в 1963 г., остались, по-видимому, незамеченными широкой научной общественностью.

К тому времени, когда появились первые предположения о возможности использования гетеропереходов в инжекционных лазерах, достигнутый уровень знаний не позволял выбирать пары полупроводников, подходящие для изготовления гетеропереходов. Кремер предложил наряду с другими сочетаниями делать широкозонные области из GaAs, а узкозонную область — из Ge. Алферов [42, 43] предложил использовать в инжекционных лазерах гетеропереходы GaAs—GaP_xAs_{1-x}. Однако в результате проведенных исследований он не смог получить низкопороговые лазеры на этих структурах. Вероятно, основной причиной этого было плохое качество кристалла из-за несовпадения периодов решеток двух различных полупроводников, из которых были изготовлены гетеропереходы. К сожалению, при концентрациях фосфора, достаточно больших для того, чтобы соединение GaP_xAs_{1-x} имело необходимую ширину запрещенной зоны, имеет место плохое совпадение периодов решеток этого соединения и GaAs.

В июне 1967 г. Вудолл и сотрудники [44] из ИБМ сообщили, что им удалось вырастить Al_xGa_{1-x}As на GaAs методом жидкофазной эпитаксии (ЖФЭ). Этому методу, впервые примененному Нельсоном [45] из РКА для выращивания GaAs на GaAs, суждено было в дальнейшем занять важное место в технологии полупроводников. Ширина запрещенной зоны Al_xGa_{1-x}As увеличивается при возрастании молярной доли AlAs x , поэтому излучение светодиодов из Al_xGa_{1-x}As является более коротковолновым, чем получаемое в светодиодах из GaAs. Эти таксимальные слои Al_xGa_{1-x}As могут быть высокого качества, так

БИБЛИОТЕКА
ХИФЭ

Изд. № 659824

учебной и научной
литературы
Библиотека ХИФЭ

как значения параметров решеток GaAs и AlAs почти одинаковы. Работа Руппрехта и др. [46] по светодиодам из $\text{Al}_x\text{Ga}_{1-x}\text{As}$ была опубликована несколькими месяцами позже работы Вудолла и др. [44], незадолго до конференции IEEE по полупроводниковым лазерам, состоявшейся в ноябре 1967 г. Ни одна из работ, представленных на конференции, не была посвящена проблеме уменьшения плотности порогового тока. Тем не менее Руппрехт и др. [47] в своем докладе распространили выводы своей ранней работы по светодиодам из $\text{Al}_x\text{Ga}_{1-x}\text{As}$ на гомолазеры на этом же соединении.

Несмотря на то что появление работ по $\text{Al}_x\text{Ga}_{1-x}\text{As}$ на двух конференциях не вызвало плодотворных дискуссий, как это имело место в случае со статьями по электролюминесценции, появившимися в 1962 г., исследования группы из ИБМ не остались незамеченными. На конференции по лазерам присутствовали Паниш и Хаяси из «Белл лэборатризи». Оба были новичками в работе с полупроводниками. Паниш ранее специализировался в химической термодинамике, а Хаяси — по приборам для ядерных исследований. Паниш изучал поведение примесей в GaAs, выращенном из раствора, и вместе с сотрудниками в 1966 г. сообщил [48] о наблюдении фотолуминесценции в GaAs, выращенном методом ЖФЭ, более эффективной, чем в GaAs, выращенном из расплава. Незадолго до этого Хаяси и Паниш объединили свои усилия в исследованиях инжекционных лазеров с целью добиться понижения плотности порогового тока при комнатной температуре $J_{\text{пор}}$ (300 К). Этому сотрудничеству способствовал Дж. К. Гэлт, который был директором исследовательской лаборатории электроники твердого тела в «Белл лэборатризи». Значительное снижение $J_{\text{пор}}$ (300 К) позволило бы получить лазеры, работающие в непрерывном режиме при комнатной температуре, и Гэлт был уверен в том, что эти лазеры будут играть важную роль в деле практического осуществления системы оптической связи на большие расстояния.

В 1967 г. за несколько недель до начала конференции по полупроводниковым лазерам Хаяси рассмотрел вопрос о том, какие преимущества по сравнению с гомолазерами могут быть получены от использования гетеропереходов для ограничения носителей тока в инжекционных лазерах. Был исследован [49] ряд соединений элементов III и V групп. Однако на тот факт, что GaAs и $\text{Al}_x\text{Ga}_{1-x}\text{As}$ имеют почти одинаковый период решетки, он не обратил должного внимания. Хаяси занимался также вопросом о влиянии формы хвостов плотности состояний зоны проводимости и валентной зоны на распределение инжектированных носителей в GaAs [50]. Хаяси и Паниш при рассмотрении проблемы понижения плотности порогового тока при комнатной температуре основное внимание уделили использова-

нию легирующих примесей для воздействия на форму хвостов зон.

Учитывая результаты исследований Паниша и др. [48], проведенных в 1966 г., Хаяси и Паниш решили использовать метод ЖФЭ. До начала конференции по лазерам они не были знакомы с опубликованной в 1963 г. работой Кремера и только что появившейся статьей Руппрехта. Руппрехт в своем выступлении на конференции подчеркнул возможность получения более коротких длин волн на $Al_xGa_{1-x}As$ -гомолазерах; при этом он отметил также почти полное совпадение периодов решеток $GaAs$ и $Al_xGa_{1-x}As$. Сообщение Руппрехта послужило причиной того, что Хаяси и Паниш решили пересмотреть свой подход к решению проблемы уменьшения плотности порогового тока лазера, поскольку были знакомы с методом ЖФЭ и уже имели ряд соображений о преимуществах, которые могли бы быть получены при использовании гетеропереходов. Разумно было предположить, что гетеропереход $GaAs-Al_xGa_{1-x}As$ с хорошим согласованием решеток обеспечит эффективный барьер для диффузии электронов. Из рассмотрения диффузии инжектированных носителей тока и их энергетического распределения Хаяси сделал вывод, что ограничение для носителей могло бы снизить порог генерации.

После участия в работе конференции по лазерам Паниш и Самски [51], воодушевленные идеями, обсуждавшимися на конференции, провели исследования с целью установить условия фазового равновесия между твердым раствором $Al_xGa_{1-x}As$ и жидкой фазой раствора $Al-Ga-As$ и использовать эти условия при разработке технологии выращивания эпитаксиальных слоев $Al_xGa_{1-x}As$ на подложках из $GaAs$. Результатом сотрудничества Хаяси и Паниша явилось создание лазера [52, 53] со значительно пониженной плотностью порогового тока при комнатной температуре $8,6 \cdot 10^3$ А/см². На рис. 1.2.1, б показана структура этого лазера. Он состоял из слоя p -типа $Al_xGa_{1-x}As$, сильно легированного Zn , выращенного на подложке из $n-GaAs$. При выращивании или при последующем отжиге Zn диффундировал в подложку из $GaAs$ и образовывал $p-n$ -переход, смещенный примерно на 2 мкм от гетероперехода. Гетеропереход создавал потенциальный барьер, который ограничивал диффузию электронов, инжектированных через $p-n$ -переход в $p-GaAs$. Более детально этот процесс описывается в последующих главах. Пороги генерации при комнатной температуре, значительно более низкие, чем в гомолазерах, были достигнуты в середине 1968 г., и к концу этого года Хаяси и Паниш убедились, что с имеющимися в их распоряжении материалами для подложек и изготовлении $p-n$ -переходов методом диффузии Zn они не смогут еще больше понизить плотность порогового

тока. Свои статьи Хаяси и Паниш послали в печать в январе 1969 г., будучи уверенными в том, что они являются первыми, кто предложил или изготовил гетеролазер. Их уверенность была сильно поколеблена, во-первых, рецензентом, который указал на отсутствие ссылки на статью Кремера 1963 г., и, во-вторых, публикацией в мартовском выпуске RCA Review статьи Крессела и Нельсона [54], в которой был описан по существу такой же лазер. Статьи Хаяси и Паниша [52, 53] были напечатаны в апреле 1969 г., а более детальная статья [55] — в 1970 г. В предвидении появления более сложных структур они предложили [55] называть такие лазеры лазерами на односторонней гетероструктуре (ОГС).

В то время как усилили Паниша и Хаяси, Крессела и Нельсона были сосредоточены на создании ОГС-лазера, Алферов и сотрудники из Ленинградского Физикотехнического института им. А. Ф. Иоффе также работали с системой $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$. В 1968 г. они сообщили [56] об исследованиях по инжекции через $p-n$ -гетеропереходы $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$, а в их работе [57], опубликованной в сентябре 1969 г., сообщалось о достижении плотностей порогового тока от $4 \cdot 10^3$ до $13 \cdot 10^3$ А/см² при температуре 300 К. На рис. 1.2.1, в показана исследованная ими лазерная структура. Она была выращена на подложке из GaAs и представляла собой слой p -GaAs, находящийся между слоями n - и p - $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Такая структура обеспечивает ограничение как для излучения, так и для носителей тока в узком слое арсенида галлия.

В заключительный период своей совместной работы по ОГС-лазерам Хаяси и Паниш в 1968 г. начали также работу по исследованию многослойных структур, аналогичных изображенной на рис. 1.2.1, в. Лазеры на таких структурах они называли лазерами на двусторонней гетероструктуре (ДГС). К началу 1970 г. плотность порогового тока была понижена ими до $2,3 \cdot 10^3$ А/см² при температуре 300 К [58], а к весне того же года плотность порогового тока была доведена до $1,6 \cdot 10^3$ А/см² при $T = 300$ К. Отводя от лазера тепло, они получили непрерывный режим генерации при температуре теплоотвода 311 К. Большая работа была проведена по снятию спектральной зависимости интенсивности излучения и зависимости интенсивности излучения от тока накачки. Характер этих зависимостей ясно свидетельствовал о получении вынужденного излучения в непрерывном режиме при комнатной температуре. О результатах проведенной работы Хаяси [59] впервые сообщил в июне 1970 г. на конференции по научным приборам в Сиэтле, штат Вашингтон. Статья [60] с результатами исследований была послана в Applied Physics Letters в начале июня и была напечатана в выпуске от 1 августа. И вновь Хаяси и Паниш были убеж-

дены в том, что они первые добились таких результатов. Однако в сентябрьском за 1970 г. выпуске журнала «Физика и техника полупроводников» была опубликована статья Алферова и др. [61], полученная редакцией в мае, в которой сообщалось о получении ДГС-лазера, работающего в непрерывном режиме при $T = 300 \text{ К.}$

[За время, прошедшее с момента создания первых непрерывных инжекционных лазеров, в лабораториях всего мира было проведено огромное количество работ по ДГС и более сложным гетеролазерам. Были проведены также системные исследования этих лазеров с целью использования их в системах волоконно-оптической связи. Хотя основные принципы остались теми же самыми, были предложены многочисленные модификации лазерных структур с целью оптимизировать их рабочие параметры. Новые материалы позволили создать лазеры, работающие на разных длинах волн. Усовершенствование процессов выращивания лазерных структур и изготовления лазеров привело к созданию на основе системы $\text{Al}_x\text{Ga}_{1-x}\text{As}$ лазеров с очень большим сроком службы. Кажется очевидным, что гетеролазеры еще долгое время будут оставаться в центре внимания исследователей.]

§ 3. ПРИНЦИП ДЕЙСТВИЯ ИНЖЕКЦИОННОГО ЛАЗЕРА

Волноводный эффект

[В полупроводниковых инжекционных лазерах ограничение светового поля и распространение излучения происходит в диэлектрическом волноводе.] На рис. 1.3.1, *a* показан трехслойный диэлектрический волновод с показателями преломления слоев \bar{n}_1 , \bar{n}_2 и \bar{n}_3 . В гл. 2 показано, что волноводный эффект возникает, когда

$$\bar{n}_2 > \bar{n}_1 \geq \bar{n}_3. \quad (1.3.1)$$

В этом случае угол падения луча θ_{12} на границу раздела между слоями 1 и 2 на рис. 1.3.1, *b* превосходит критический угол

$$\theta_c = \arcsin(\bar{n}_1/\bar{n}_2). \quad (1.3.2)$$

Аналогичное выражение может быть написано для границы раздела между слоями 2 и 3. При выполнении условий (1.3.1) электромагнитное излучение распространяется в направлении, параллельном границам раздела между слоями.

[В гомолазерах разность между показателем преломления центрального слоя волновода и более низкими показателями преломления соседних слоев составляет лишь $\sim 0,01$ — $0,03$. Такая малая величина этой разности обусловлена различием

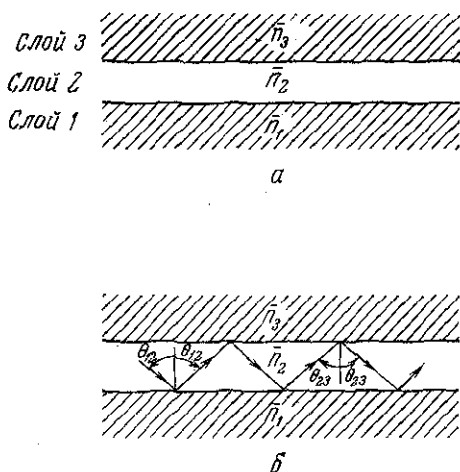


Рис. 1.3.1. Трехслойный плоский диэлектрический волновод (а). Траектории лучей при распространении в волноводе направляемой волны (б); угол падения луча на границу раздела θ отсчитывается от нормали к границе.

показателей преломления n - и p -областей, температурными градиентами, изменением показателя преломления вследствие инжекции носителей тока и усиления. В гетеролазерах скачки показателя преломления на каждом гетеропереходе могут быть значительно больше, и, следовательно, в гетеролазерах может быть сформирован четко выраженный волновод. В этом случае распределение оптического поля в волноводе может быть вычислено, если известны размеры системы и показатели преломления слоев.

Усиление и порог

Неосновные носители тока, инжектированные через $p-n$ -переход, могут рекомбинировать как излучательным, так и безызлучательным образом. Рекомбинационное излучение может взаимодействовать с валентными электронами и поглощаться ими или взаимодействовать с электронами зоны проводимости и вызывать испускание идентичного фотона. Когда концентрация инжектированных носителей становится достаточно высокой, уровень вынужденного излучения может превысить уровень поглощения, в результате чего возникает оптическое усиление. Для возникновения генерации необходима положительная обратная связь. Эта обратная связь обычно создается за счет пары сколотых граней, перпендикулярных оси волновода. Генерация возникает, когда усиление за один проход излучения

между такими полученными сколом зеркалами сравнивается с суммой всех потерь на том же пути. Внутренние потери α_i обычно обусловлены поглощением на свободных носителях и рассеянием на оптических неоднородностях. Коэффициент отражения зеркал R , меньший единицы, также дает вклад в потери, равный $(1/L) \ln(1/R)$, где L — длина резонатора.

В гл. 7 показано, что коэффициент усиления в GaAs при комнатной температуре может быть представлен в виде

$$g \text{ (см}^{-1}\text{)} = 5,0 \cdot 10^{-2} [(J\eta/d) - 4,5 \cdot 10^3], \quad (1.3.3)$$

где J — плотность тока, η — квантовая эффективность, а d — толщина равномерно накачанной области в микрометрах. На пороге генерации усиление должно равняться сумме всех потерь; следовательно,

$$g\Gamma = \alpha_i + (1/L) \ln(1/R). \quad (1.3.4)$$

Здесь Γ — коэффициент оптического ограничения, определяющий долю электромагнитного излучения, приходящуюся на равномерно возбужденную область. Объединяя выражения (1.3.3) и (1.3.4), получим выражение для плотности тока на пороге генерации

$$J_{\text{пор}} \text{ (А/см}^2\text{)} = 4,5 \cdot 10^3 d/\eta - (20d/\eta\Gamma) [\alpha_i + (1/L) \ln(1/R)]. \quad (1.3.5)$$

В некоторых гетеролазерах не весь ток идет на накачку однородно возбужденной области, которая называется активной областью. Часть плотности тока, не участвующая в процессе накачки, может быть названа током утечки J_L .

Гетеропереходы

Гетеропереход представляет собой переход между двумя различными полупроводниками, образованными в одном кристалле. Наиболее важное различие этих полупроводников в гетеролазерах заключается в разных значениях ширины запрещенной зоны и показателя преломления. В § 2 этой главы кратко описаны исследования, направленные на понижение $J_{\text{пор}}$ при комнатной температуре путем использования GaAs—Al_xGa_{1-x}As-гетероструктур. Уменьшение $J_{\text{пор}}$ происходит при уменьшении d и увеличении Γ ; [выражение (1.3.5)]. Это достигалось помещением активного слоя GaAs между более широкозонными слоями Al_xGa_{1-x}As, в результате чего одновременно получалось ограничение и для инжектированных носителей тока, и для оптического поля.

В гетеропереходах полупроводник с узкой запрещенной зоной обычно обозначают буквами n или p , а более широкозонный полупроводник — буквами N или P в зависимости от типа

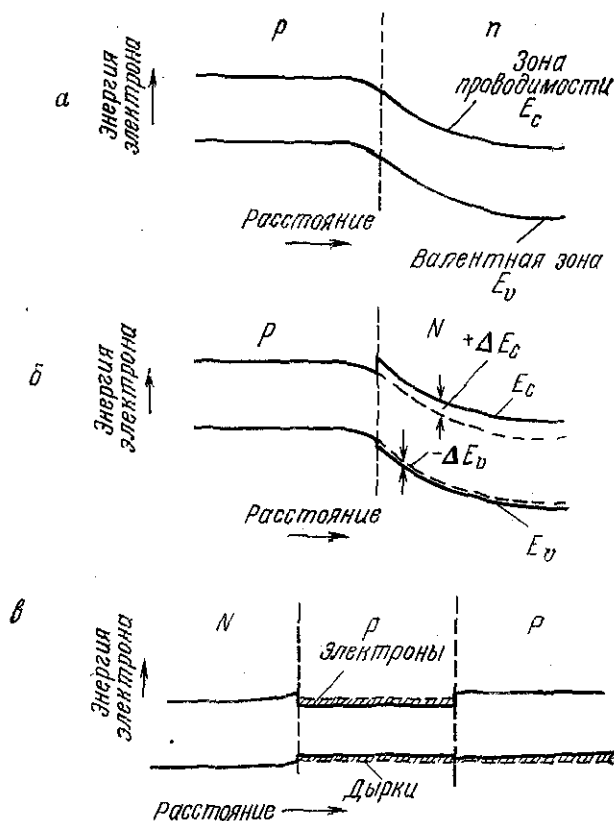


Рис. 1.3.2. Энергетическая зонная диаграмма $p-n$ -перехода в GaAs (а). Энергетическая зонная диаграмма GaAs— $Al_{0,3}Ga_{0,7}As$ $p-N$ -гетероперехода (б); получается при добавлении к энергии дна зоны проводимости n -GaAs величины ΔE_c и вычитании из энергии потолка валентной зоны n -GaAs величины ΔE_v , которые определяют разрывы соответствующих зон. Энергетическая зонная диаграмма GaAs— $Al_{0,3}Ga_{0,7}As$ двусторонней $N-p-p$ -гетероструктуры при большом прямом смещении (в).

основных носителей. Различие между гомопереходами и гетеропереходами может быть легко проиллюстрировано рассмотрением энергетических зонных диаграмм GaAs $p-n$ -перехода и GaAs— $Al_xGa_{1-x}As$ $p-N$ -гетероперехода. На рис. 1.3.2, а приведена энергетическая зонная диаграмма $p-n$ -перехода при термическом равновесии. Величина запрещенной зоны E_g в $Al_{0,3}Ga_{0,7}As$ больше, чем в GaAs. Перепад значений величины запрещенной зоны распределяется между зоной проводимости и валентной зоной. На рис. 1.3.2, б, изображающем $p-N$ -гетеропереход при тепловом равновесии, часть этого перепада, при-

ходящаяся на зону проводимости, обозначена через ΔE_c , а остальная часть, приходящаяся на валентную зону, — через ΔE_v . Как следует из рис. 1.3.2, в, на котором показана двусторонняя $N-p-P$ -гетероструктура при большом прямом смещении, $p-P$ -гетеропереход создает ограничение для электронов, инжектированных в p -слой, в то время как $N-p$ -гетеропереход препятствует утечке дырок из p -слоя в N -слой.

§ 4. ИНЖЕКЦИОННЫЕ ЛАЗЕРЫ

Гомолазер

В гомолазере, описанном в § 2 и изображенном на рис. 1.2.1, а, толщина d лежащего в p -области слоя, в котором происходит усиление, определяется длиной диффузии L_n инжектированных электронов. В сильно легированном высококачественном GaAs при комнатной температуре L_n может быть равной примерно 1—3 мкм. Концентрация инжектированных электронов в направлении, перпендикулярном плоскости $p-n$ -перехода,

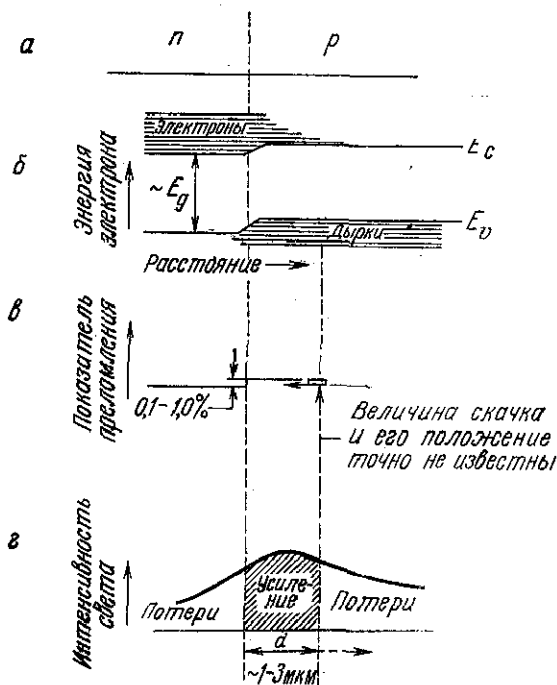


Рис. 1.4.1. Лазер на $p-n$ GaAs-гомоструктуре: а — схематическое изображение; б — энергетическая зонная диаграмма при большом прямом смещении и распределение носителей тока; в — профиль показателя преломления; г — распределение оптического поля.

уменьшается с расстоянием x как $\exp(-x/L_n)$. Поэтому накачка и, следовательно, усиление не будут однородными по объему. Кроме того, часть электронов продиффундирует за пределы области усиления и не будет участвовать в процессе усиления излучения. Пока еще нет детального описания волноводного эффекта в гомолазерах. Однако известно, что скачки показателя преломления на краях области усиления малы, что приводит к утечке световой энергии в окружающие области GaAs, в которых велики потери. Это проникновение оптического поля в области с большими потерями приводит к дополнительному вкладу в α_i . Поэтому в выражении (1.3.5) для $J_{\text{пор}}$ слабое оптическое ограничение приводит к малым значениям Γ и большим значениям α_i , в то время как толщина активной области d , определяемая диффузионной длиной, равна нескольким микрометрам. Эти факторы ограничивают $J_{\text{пор}}$ при комнатной температуре значениями, превышающими $26 \cdot 10^3$ А/см² [63]. На рис. 1.4.1, *в* и *г* схематически показаны профиль показателя преломления и распределение оптического поля для гомолазера (рис. 1.4.1, *а*) с распределением носителей тока, представленным на рис. 1.4.1, *б*.

Гетеролазеры

На примере двусторонней гетероструктуры (ДГС) могут быть проиллюстрированы наилучшим образом те свойства гетеропереходов, благодаря которым в гетеролазерах плотности порогового тока при комнатной температуре имеют гораздо более низкие значения, чем в гомолазерах. Как $N-p-P$, так и $N-p-P$ -структуры обладают одинаковыми свойствами. Гетеропереходы служат для ограничения инжектированных носителей в p (или n) активном слое, как это проиллюстрировано на рис. 1.3.2, *в*. Толщина активного слоя d может быть значительно уменьшена от неконтролируемых значений в несколько микрометров в гомолазерах до 0,1 мкм и меньше в GaAs—Al_xGa_{1-x}As ДГС-лазерах. Как следует из выражения (1.3.5), уменьшение d понижает $J_{\text{пор}}$. Для $N-p-P$ ДГС-лазера, изображенного на рис. 1.4.2, *а*, явление ограничения для носителей, проиллюстрированное рис. 1.3.2, *в*, вновь представлено на рис. 1.4.2, *б* для случая большого прямого смещения. Более широкозонные N - и P -области имеют к тому же более низкие значения показателя преломления, как это показано на рис. 1.4.2, *в*. Таким образом, гетеропереходы создают диэлектрический волновод, который ограничивает световое поле в активном слое (рис. 1.4.2, *г*). Поэтому d может быть значительно уменьшено при сохранении приемлемых значений Γ . Кроме того, часть излучения, распространяющаяся вне активного слоя, находится в более широкозонном полупроводнике и не поглощается, как это происходит

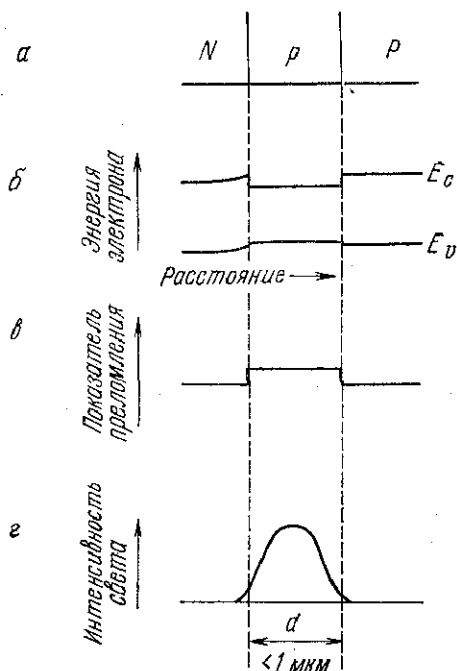


Рис. 1.4.2. Лазер на $N-p-p$ GaAs— $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ -двусторонней гетероструктуре: a — схематическое изображение; $б$ — энергетическая зонная диаграмма при большом прямом смещении; $в$ — профиль показателя преломления; $г$ — распределение оптического поля.

в гомолазере. Использование двусторонних гетероструктур позволило понизить $J_{\text{пор}}$ до значений порядка $1 \pm 0,5 \cdot 10^3 \text{ А/см}^2$ при температуре 300 К.

Большинство лазеров, предназначенных для работы вне лаборатории, имеют полосковую геометрию, при которой ограничение тока создается также в плоскости $p-n$ -перехода. Для осуществления бокового ограничения по току существует множество способов. Уменьшенная площадь при полосковой геометрии приводит к существенно более низким значениям токов в непрерывном режиме генерации, равным примерно 100—150 мА для ширины полоски 12 мкм и длины резонатора 400 мкм. В лазерах с полосковой геометрией может быть значительно улучшена также и стабильность излучения.

§ 5. МАТЕРИАЛЫ

Создание инжекционного лазера и получение непрерывного режима генерации при комнатной температуре были результатом не только открытия лазерного эффекта и достижений

физики полупроводников в 50-е годы, но также и развития технологии выращивания ряда новых полупроводников. В § 2 указано, что в полупроводниковом лазере нельзя использовать наиболее популярные полупроводники Si и Ge из-за того, что они обладают непрямой запрещенной зоной. Основными материалами для полупроводниковых лазеров служат соединения элементов III и V групп периодической системы Менделеева. Эти соединения изоэлектронны элементам IV группы Si и Ge. Кроме полупроводников типа $A^{III}B^V$ в гетеролазерах используются также соединения $A^{IV}B^{VI}$. Однако они обладают гораздо более узкой запрещенной зоной и излучают в области длин волн $\sim 5-15$ мкм.

В гетеролазерах на основе рассмотренных выше химических систем диапазон излучаемых длин волн существенно расширяется при использовании кристаллических твердых растворов, в которых часть атомов вещества заменена другими атомами из той же группы периодической системы, что и удаляемые атомы. Например, тройной твердый раствор соединений $A^{III}B^V$ может состоять из атомов элементов III группы A и B, произвольным образом расположенных в узлах подрешетки элементов III группы, и атомов элемента V группы C, расположенных во всех узлах подрешетки элементов V группы. Такой твердый раствор обозначается $A_xB_{1-x}C$, где x может меняться от нуля до единицы, за исключением возможной области несмешивания. Типичным примером такого раствора является твердый раствор $Al_xGa_{1-x}As$, который рассматривался в предыдущем параграфе этой главы и который широко используется в качестве примера на протяжении всей книги.

Четверной твердый раствор соединений $A^{III}B^V$ может состоять из атомов двух элементов III группы A и B, произвольным образом распределенных по узлам подрешетки элементов III группы, и атомов двух элементов V группы C и D, произвольным образом распределенных по узлам подрешетки элементов V группы. Этот твердый раствор описывается формулой $A_xB_{1-x}C_yD_{1-y}$. Другой четверной твердый раствор соединений $A^{III}B^V$ состоит из атомов трех элементов III (V) группы, произвольным образом распределенных по узлам подрешетки элементов III (V) группы, и атомов одного элемента V (III) группы; занимающих все узлы подрешетки элементов V (III) группы. Такой четверной твердый раствор описывается формулой $(A_xB_{1-x})_yC_{1-y}D$, где A, B и C — элементы III (V) группы, а D — элемент V (III) группы. Как и в случае тройного раствора, x и y меняются от нуля до единицы, за исключением областей несмешивания.

В ряде бинарных соединений $A^{III}B^V$ при возрастании атомного веса элемента III или V групп, как правило, уменьшается

ширина запрещенной зоны E_g , возрастает показатель преломления \bar{n} и, за исключением систем GaAs—AlAs и GaP—AlP, возрастает постоянная решетки a_0 . Кристаллические твердые растворы бинарных соединений обычно обладают свойствами, промежуточными по сравнению со свойствами граничных соединений. На рис. 1.5.1 и 1.5.2 схематически представлены зависимости E_g и a_0 от состава для систем $A_xB_{1-x}C_yD_{1-y}$ и $(A_xB_{1-x})_yC_{1-y}D$ соответственно. Рис. 1.5.1 соответствует случаю, когда атомный вес элемента А меньше атомного веса элемента В, а атомный вес элемента С меньше атомного веса элемента D; при построении диаграммы на рис. 1.5.2 предполагалось, что атомные веса элементов расположены в следующем порядке: $A < C < B$.

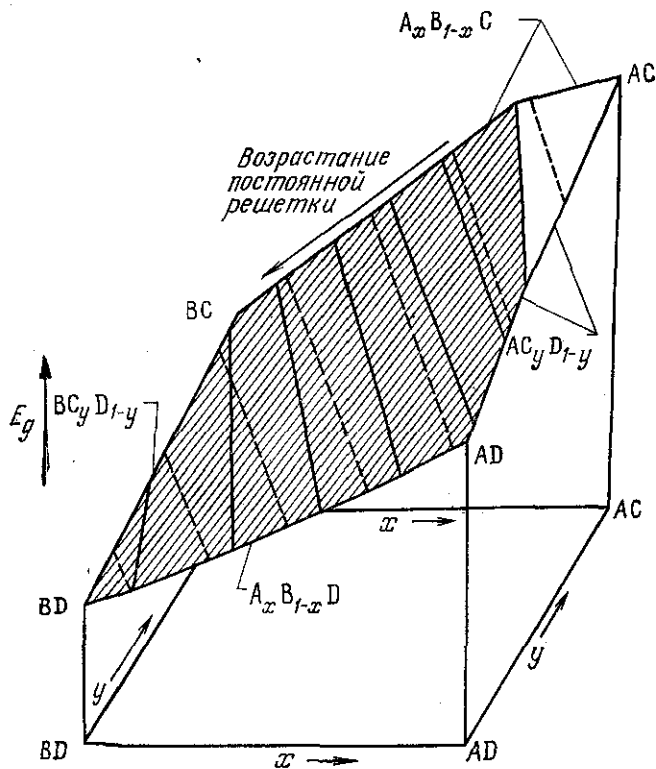


Рис. 1.5.1. Зависимость ширины запрещенной зоны от состава четверных твердых растворов элементов III и V групп типа $A_xB_{1-x}C_yD_{1-y}$ с атомными весами $A < B$ и $C < D$. Штриховые линии соответствуют изопериодическим составам, сплошные — изоэнергетическим составам, т. е. составам, имеющим одинаковую ширину запрещенной зоны. Заштрихованная и незаштрихованная поверхности определяют области прямозонных и непрямозонных составов соответственно. Углы диаграммы соответствуют четырем бинарным соединениям, а ее стороны — четырем тройным твердым растворам.

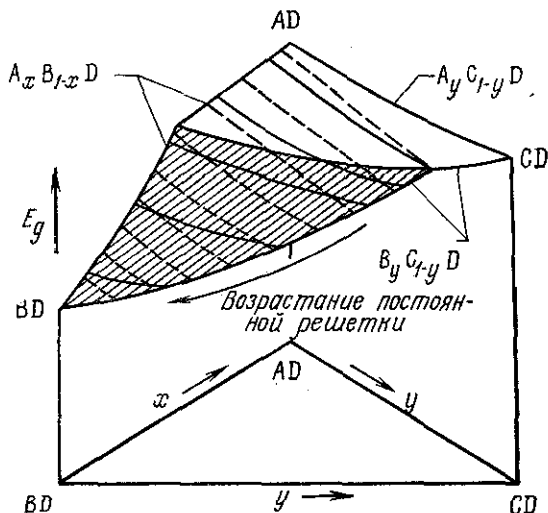


Рис. 1.5.2. Зависимость ширины запрещенной зоны от состава четверных твердых растворов элементов III и V групп типа $(A_x B_{1-x})_y C_{1-y} D$. Обозначение кривых, соответствующих изопериодическим и изоэнергетическим составам, а также поверхностей, определяющих области прямозонных и непрямозонных составов, то же, что и на рис. 1.5.1. Углы диаграммы соответствуют трем бинарным соединениям, а ее стороны — трем тройным твердым растворам.

Гетеролазеры представляют собой слоистые полупроводниковые структуры, в которых постоянная решетки остается неизменной при переходе от слоя к слою, а E_g и \bar{n} меняются. Системы твердых растворов обеспечивают непрерывное изменение E_g и \bar{n} для различных составов при постоянной a_0 , и представляют собой полупроводники, которые используются или, возможно, будут использоваться в гетеролазерах.

ЛИТЕРАТУРА

1. Aigrain P., в кн.: Proc. Conf. Quantum Electron, Paris, 1963, p. 1762.
2. Басов Н. Г., Вул Б. М., Попов Ю. М., ЖЭТФ, 37, 587 (1959).
3. Boyle W. S., Thomas D. G., US Patent 3059117, October 16, 1962, filed January 1960.
4. Басов Н. Г., Крохин О. Н., Попов Ю. М., ЖЭТФ, 40, 1879 (1961).
5. Bernard M. G. A., Duraffourg G., Phys. Stat. Solidi, 1, 699 (1961).
6. Welker H., Zs. Naturforsch., 7a, 744 (1952).
7. Welker H., Zs. Naturforsch., 8a, 248 (1953).
8. Bardeen J., в кн.: Collected Works of John Von Neumann, Pergamon Press, New York, 1963, vol. 5, p. 420.
9. Hall R. N., IEEE Trans. Electron Devices, ED-23, 700 (1976).
10. Keyes R. J., Quist T. M., Proc. IRE, 50, 1822 (1962).
11. Dumke P. W., Phys. Rev., 127, 1559 (1962).
12. Pankove J. I., Massoulié M. J., Journ. Electrochem. Soc., 109, 67C (1962).
13. Holonyak N. Jr., частное сообщение.
14. Наследов Д. Н. и др., ФТТ, 4, 1062 (1962).
15. Hall R. N. et al., Phys. Rev. Lett., 9, 366 (1962).

16. Nathan M. I. *et al.*, Appl. Phys. Lett., 1, 62 (1962).
17. Holonyak N., Jr., Bevacqua S. F., Appl. Phys. Lett., 1, 82 (1962).
18. Quist T. M. *et al.*, Appl. Phys. Lett., 1, 91 (1962).
19. Bond W. L. *et al.*, Appl. Phys. Lett., 2, 57 (1963).
20. Melngailis I., Rediker R. H., Solid-State Device Conf., Lansing, Michigan, June 1963; IEEE Trans. Electron. Devices, ED-10, 333 (1963).
21. Weiser K. *et al.*, Solid-State Device Conf., Lansing, Michigan, June 1963; IEEE Trans. Electron. Devices, ED-10, 334 (1963).
22. Broom R. F. *et al.*, Solid-State Device Conf., Lansing, Michigan, June 1963; IEEE Trans. Electron. Devices, ED-10, 334 (1963).
23. Hall R. N., Solid-State Device Conf., Lansing, Michigan, June 1963; IEEE Trans. Electron. Devices, ED-10, 334 (1963).
24. Mayburg S., Journ. Appl. Phys., 34, 1791 (1963).
25. Engeler W. E., Garjinkel M., Journ. Appl. Phys., 34, 2746 (1963).
26. Pilkuhn M., Rupprecht H., Blum S., Solid-State Electron., 7, 905 (1964).
27. Alexander F. B. *et al.*, Appl. Phys. Lett., 4, 13 (1964).
28. Hall R. N., Solid-State Electron., 6, 405 (1963).
29. Minden H. T., Semicond. Prod., 6, 34 (August 1963).
30. Vallese L. M., Semicond. Prod., 6, 25 (August 1963).
31. Rediker R. H., Solid State Design, 4, 19 (August 1963).
32. Lax B., Solid State Design, 4, 26 (November 1963).
33. Nathan M. I., Burns G., Electronics, 36, 61 (December 6, 1963).
34. Johnson C. M., Electronics, 36, 34 (December 13, 1963).
35. Lax B., Science, 141, 1247 (1963).
36. Burns G., Nathan M. I., Proc. IEEE, 52, 770 (1964).
37. Butler J. F. *et al.*, Appl. Phys. Lett., 5, 75 (1964).
38. Nathan M. I., Appl. Opt., 5, 1514 (1966).
39. Dymant J. C., D'Asaro L. A., Appl. Phys. Lett., 11, 292 (1967).
40. Kroemer H., Proc. IEEE, 51, 1782 (1963).
41. Алферов Ж. И., Казаринов Р. Ф., Авт. свид. СССР № 28448, 1963.
42. Алферов Ж. И. *и др.*, ФТП, 9, 279 (1967).
43. Алферов Ж. И., ФТП, 1, 436 (1967).
44. Woodall J. M., Rupprecht H., Pettit G. D., Solid-State Device Conf., June 19, 1967, Santa Barbara, Calif.; IEEE Trans. Electron. Devices, ED-14, 630 (1967).
45. Nelson H., RCA Rev., 24, 603 (1963).
46. Rupprecht H., Woodall J. M., Pettit G. D., Appl. Phys. Lett., 11, 81 (1967).
47. Rupprecht H. *et al.*, IEEE Semiconductor Laser Conf., November 29—30, 1967; IEEE Journ. Quantum Electron., QE-4, 110 (1968).
48. Panish M. B. *et al.*, Solid-State Electron., 9, 311 (1966).
49. Hayashi I., частное сообщение.
50. Hayashi I., IEEE Journ. Quantum Electron., QE-4, 113 (1968).
51. Panish M. B., Sumski S., Journ. Phys. Chem. Solids, 30, 129 (1969).
52. Hayashi I., Panish M. B., Foy P. W., IEEE Journ. Quantum Electron., QE-5, 211 (1969).
53. Panish M. B., Hayashi I., Sumski S., IEEE Journ. Quantum Electron., QE-5, 210 (1969).
54. Kressel H., Nelson H., RCA Rev., 30, 106 (1969).
55. Hayashi I., Panish M. B., Journ. Appl. Phys., 41, 150 (1970).
56. Алферов Ж. И. *и др.*, ФТП, 2, 1016 (1968).
57. Алферов Ж. И. *и др.*, ФТП, 3, 1328 (1969).
58. Panish M. B., Hayashi I., Sumski S., Appl. Phys. Lett., 16, 326 (1970).
59. Hayashi I., Panish M. B., Device Research Conference, Seattle, Washington, June 1970.
60. Hayashi I. *et al.*, Appl. Phys. Lett., 17, 109 (1970).
61. Алферов Ж. И. *и др.*, ФТП, 4, 1826 (1970).
62. Kressel H., Lockwood H. F., Ettenberg M., IEEE Spectrum., 10, 59 (1973).
63. Pilkuhn M. H., Rupprecht H., Journ. Appl. Phys., 38, 5 (1967).

ОПТИЧЕСКИЕ ПОЛЯ И РАСПРОСТРАНЕНИЕ ВОЛН

§ 1. ВВЕДЕНИЕ

Настоящая глава, посвященная описанию оптических полей и процесса распространения волн в гетероструктурах, предшествует рассмотрению вопросов о вынужденном излучении и гетеропереходах в гл. 3 и 4 соответственно, поскольку для их понимания необходимо знание уравнений Максвелла и вытекающих из них следствий. При написании главы о распространении волн в гетероструктурах должны быть решены две основные проблемы: во-первых, технический уровень изложения материала и, во-вторых, выбор вопросов, которые должны быть освещены. Вследствие того что предмет данной книги находится на пересечении многих дисциплин, мы намерены начать с изложения фундаментальных принципов и понятий для тех, кто не знаком с теорией электромагнитного поля, и провести выкладки до того уровня, когда станет возможным количественное описание процесса распространения волн в гетеролазерах. Изложение этих вопросов ведется с существенными подробностями, что помогает читателям, не знакомым с основами теории поля, проследить за выводом необходимых выражений. Вместе с тем более подготовленные читатели могут без ущерба для понимания излагаемых вопросов опустить те параграфы, в которых рассматриваются основы теории. Имеются книги, целиком посвященные оптическим волноводам [1—4], поэтому был проведен тщательный выбор материала, чтобы ограничить число излагаемых тем. Была сделана попытка дать полное описание процесса распространения волн в обычно встречающихся типах гетеролазеров.

В этой главе выводятся выражения для распределения оптического поля в направлении, перпендикулярном плоскости $p-n$ -перехода в гетеролазерах, и даны характерные численные примеры для $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ -гетероструктур. Рассмотрение распределения оптического поля вдоль плоскости перехода проводится после обсуждения в гл. 7 лазеров с полосковой геометрией. Уравнения Максвелла приведены в § 2 настоящей главы, где также выводятся соотношения, связывающие проводимость и диэлектрическую проницаемость с коэффициентом поглощения и показателем преломления. Вывод этих соотношений помогает лучше понять процесс распространения волн и позволяет

определить такие часто встречающиеся величины, как постоянная распространения, комплексный показатель преломления и коэффициент экстинкции.

В § 3 дано описание ДГС-лазера как диэлектрического волновода, а в § 4 рассматривается распространение волны в симметричном трехслойном плоском диэлектрическом волноводе. Центральный слой — это область в ДГС-лазере, в которой происходит генерация света и которая называется активным слоем. Трехмерное волновое уравнение для электрического поля оптической частоты выводится из уравнений Максвелла. Далее выводится дифференциальное уравнение, описывающее распространение электрического поля, поляризованного перпендикулярно направлению распространения, — поперечного электрического поля (ТЕ). Аналогичные уравнения описывают поперечные магнитные поля (ТМ), в которых магнитное поле поляризовано перпендикулярно направлению распространения. Эти поля зависят от двух пространственных переменных и времени, и решение волнового уравнения для них получается методом разделения переменных. Как следует из решений волновых уравнений, показатель преломления активного слоя должен быть больше показателей преломления прилегающих слоев, чтобы в трехслойной структуре происходило волноводное распространение излучения. Граничные условия для электрического и магнитного полей также выводятся из уравнений Максвелла. Применение этих граничных условий на границах раздела диэлектриков (гетеропереходах) приводит к дисперсионному уравнению, являющемуся уравнением на собственные значения, которое дает набор дискретных значений постоянной распространения. Получающиеся для этих дискретных значений конфигурации электрического и магнитного полей называются модами.

Обзор экспериментальных значений показателей преломления GaAs и $Al_xGa_{1-x}As$ сделан в § 5. Это позволяет путем численного решения получить распределение поля в симметричных трехслойных волноводах. Представлена зависимость показателя преломления от концентрации примесей, температуры и молярной доли AlAs. Приведено графическое решение дисперсионного уравнения, позволяющее наглядно показать, как получают моды и что такое условия отсечки мод в симметричных плоских трехслойных волноводах. В симметричной структуре основная мода никогда не отсекается, а условия отсечки для других мод зависят от разности показателей преломления и толщины активного слоя. Численные примеры, характеризующие распределения электрического поля и интенсивности света в гетероструктурах GaAs— $Al_xGa_{1-x}As$, даны с целью продемонстрировать зависимость этих распределений от толщины активного слоя и разности показателей преломления. Отношение

интенсивности той части излучения, которая распространяется в активном слое, к интенсивности всего излучения представляет собой коэффициент оптического ограничения и используется для выражения влияния волноводных параметров на пороговую плотность тока. Приводятся численные значения этого коэффициента для $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$, а также дано приближенное решение дисперсионного уравнения в случае малых толщин активного слоя. Влияние коэффициента оптического ограничения на пороговую плотность тока в ДГС-лазерах обсуждается в § 8 гл. 3, а соответствующие рисунки приведены в § 4 гл. 7.

В § 6 асимметричный трехслойный плоский волновод рассматривается с точки зрения модели зигзагообразных волн. Дисперсионное уравнение для распространяющихся волн выводится в этой модели из рассмотрения отражения волны на границе раздела диэлектриков. Это уравнение легко решается на ЭВМ как для симметричного, так и для асимметричного случаев. Приведенные примеры распределения электрического поля в симметричной структуре на основе $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ дополнены данными для асимметричного волновода. По мере того как волновод становится все более асимметричным, коэффициент оптического ограничения уменьшается, и при малом скачке показателя преломления на одной из границ будет существовать такое значение толщины активного слоя, соответствующее этому скачку, при котором будут выполняться условия отсечки и для основной моды.

Картина излучения в дальней зоне, представляющая большой интерес для исследования гетеролазеров, рассматривается в § 7. Показано, что вычисленная зависимость интенсивности излучения в дальней зоне от угла Θ , отсчитываемого от нормали к поверхности в направлении, перпендикулярном плоскости $p-n$ -перехода, содержит множитель $\cos^2 \Theta$, который часто опускают. В этом же параграфе проведено сравнение экспериментальной и рассчитанной картин излучения в дальней зоне для симметричных ДГС-лазеров.

Отражение на зеркальных гранях гетеролазеров обеспечивает для излучения обратную связь, необходимую для генерации, и влияет на ряд свойств гетеролазеров. В § 8 показано, что отражение на гранях приводит к выбору ТЕ-поляризации как преобладающей для излучения ДГС-лазеров. Коэффициент отражения на гранях входит в выражение для плотности порогового тока (§ 8 гл. 3), а также влияет (§ 4 гл. 7) на значение толщины активного слоя, при котором наблюдаются моды высшего порядка, определяющие распределение поля в направлении, перпендикулярном плоскости $p-n$ -перехода.

В § 9 описывается распределение оптического поля в четырех- и пятислойных волноводах и лазерах. В этих структурах

излучательная рекомбинация обычно ограничивается малой областью между гетеропереходами или между гомопереходом и гетеропереходом и имеется более широкая область, ограниченная гетеропереходами, которая и образует волновод. Поэтому такие лазеры называют двусторонними гетероструктурами с раздельным ограничением (ДГС — РО). Когда число слоев превышает три, дисперсионные уравнения становятся значительно сложнее и их решение должно быть получено на вычислительной машине.

В § 10 показано, что обратная связь для излучения может быть получена за счет периодического изменения показателя преломления внутри оптического резонатора, а не только путем скалывания или полировки зеркал диода. Структуры, использующие этот метод получения обратной связи, называются лазерами с распределенной обратной связью (РОС) или с распределенным брэгговским отражателем (РБО) и могут быть использованы как лазерные источники в интегральной оптике. Показано, что передача энергии между встречными волнами происходит тогда, когда периодичность изменения показателя преломления удовлетворяет известным условиям Брэгга, связывающим период изменения с длиной волны излучения. Коэффициент связи определяет количество передаваемой энергии. Его численные значения приводятся для GaAs—Al_xGa_{1-x}As-лазера с РОС.

§ 2. ОСНОВЫ ТЕОРИИ

Уравнения Максвелла

Уравнения Максвелла в соответствующей форме и с реальными граничными условиями служат отправной точкой при анализе распространения волн в гетероструктурах. Теория электромагнитного поля изложена во многих книгах. Весьма полезное изложение, более подробное, чем в настоящей монографии, дано, например, в книге Мэджида [5]. Два уравнения, особенно важные для анализа распространения электромагнитных волн, связывают вектор электрического поля \mathcal{E} и вектор плотности электрического потока \mathcal{D} с вектором магнитного поля \mathcal{H} и вектором плотности магнитного потока \mathcal{B} . Эти уравнения имеют вид

$$\nabla \times \mathcal{E} = -\partial \mathcal{B} / \partial t, \quad (2.2.1)$$

$$\nabla \times \mathcal{H} = \mathbf{i} + (\partial \mathcal{D} / \partial t), \quad (2.2.2)$$

где \mathbf{i} — вектор плотности тока. Для рассматриваемых здесь случаев берется линейная зависимость \mathcal{D} от \mathcal{E} , т. е.

$$\mathcal{D} = \varepsilon \mathcal{E}, \quad (2.2.3)$$

где ϵ — диэлектрическая проницаемость. Аналогично \mathcal{B} и \mathcal{H} связаны соотношением

$$\mathcal{B} = \mu \mathcal{H}, \quad (2.2.4)$$

где μ — магнитная проницаемость, для полупроводниковых гетероструктур равная вакуумному значению μ_0 . Вектор плотности тока i пропорционален \mathcal{E} :

$$i = \sigma \mathcal{E}, \quad (2.2.5)$$

где σ — проводимость. Уравнения (2.2.1) и (2.2.2) вместе с уравнениями

$$\nabla \cdot \mathcal{D} = \rho, \quad (2.2.6)$$

$$\nabla \cdot \mathcal{B} = 0 \quad (2.2.7)$$

называют уравнениями Максвелла. В уравнении (2.2.6) ρ обозначает плотность заряда.

В гетеролазерах некоторая путаница может возникнуть вследствие того, что в них наряду с электрическим и магнитным полями, связанными с приложенным напряжением смещения и током через переход, существуют также оптические электрическое и магнитное поля, связанные со спонтанным и вынужденным излучениями. Однако различие частот полей смещения \mathcal{E}_B и \mathcal{H}_B , с одной стороны, и оптических полей \mathcal{E}_O и \mathcal{H}_O — с другой, позволяет рассматривать эти поля отдельно. Максимальная частота напряжения смещения порядка $1 \cdot 10^9$ Гц, в то время как длина волны излучения гетеролазера равна примерно 0,9 мкм (инфракрасное излучение), что соответствует частоте порядка $3 \cdot 10^{14}$ Гц. Вследствие такого различия частот полные векторы электрического и магнитного полей можно представить в виде суммы векторов поля смещения и оптического поля:

$$\mathcal{E} = \mathcal{E}_B + \mathcal{E}_O, \quad (2.2.8)$$

$$\mathcal{H} = \mathcal{H}_B + \mathcal{H}_O. \quad (2.2.9)$$

Для линейных не взаимодействующих систем уравнения Максвелла для поля смещения и для оптического поля могут быть записаны отдельно. Например, уравнения (2.2.1) и (2.2.2) дают

$$\nabla \times \mathcal{E}_O = -\partial \mathcal{B}_O / \partial t = -\mu_0 \partial \mathcal{H}_O / \partial t, \quad (2.2.10)$$

$$\nabla \times \mathcal{H}_O = i + \partial \mathcal{D}_O / \partial t = \sigma \mathcal{E}_O + \epsilon \partial \mathcal{E}_O / \partial t \quad (2.2.11)$$

для оптических полей. Уравнения (2.2.10) и (2.2.11) будут использованы для описания распространения волн в гетероструктурах. Следует заметить, что в эти уравнения входят электрические постоянные σ и ϵ , в то время как распространение оптических волн описывают через показатель преломления \bar{n} и коэффициент поглощения α . Поэтому нашей следующей задачей будет вывод соотношений, связывающих σ и ϵ с \bar{n} и α .

Соотношения между электрическими и оптическими постоянными

Многие свойства распространения волн могут быть изучены при рассмотрении диэлектрической среды без потерь, в которой электрическая проводимость равна нулю. Простейшее решение соответствует однородной плоской волне, поляризованной в направлении оси x , так что $\mathcal{E} = \mathcal{E}_x$ и $\mathcal{E}_y = \mathcal{E}_z = 0$, а пространственное изменение происходит только в направлении оси z , так что $\partial/\partial x = \partial/\partial y = 0$. На рис. 2.2.1 показана используемая прямоугольная система координат. Расписывая ротор вектора \mathcal{E} в уравнении (2.2.10), получаем [6] (индекс O для оптического поля в дальнейшем мы опускаем)

$$\begin{aligned} a_x \left(\frac{\partial \mathcal{E}_z}{\partial y} - \frac{\partial \mathcal{E}_y}{\partial z} \right) + a_y \left(\frac{\partial \mathcal{E}_x}{\partial z} - \frac{\partial \mathcal{E}_z}{\partial x} \right) + a_z \left(\frac{\partial \mathcal{E}_y}{\partial x} - \frac{\partial \mathcal{E}_x}{\partial y} \right) = \\ = -\mu_0 \left[a_x \left(\frac{\partial \mathcal{H}_x}{\partial t} \right) + a_y \left(\frac{\partial \mathcal{H}_y}{\partial t} \right) + a_z \left(\frac{\partial \mathcal{H}_z}{\partial t} \right) \right]. \end{aligned} \quad (2.2.12)$$

где a_x, a_y, a_z — единичные векторы в направлении осей x, y и z соответственно. Сравнивая в уравнении (2.2.12) коэффициенты при одинаковых единичных векторах и учитывая, что $\mathcal{E}_y = \mathcal{E}_z = 0$ и $\partial/\partial x = \partial/\partial y = 0$, получаем

$$0 = -\mu_0 \partial \mathcal{H}_x / \partial t, \quad (2.2.13)$$

$$\partial \mathcal{E}_x / \partial z = -\mu_0 \partial \mathcal{H}_y / \partial t, \quad (2.2.14)$$

$$0 = -\mu_0 \partial \mathcal{H}_z / \partial t. \quad (2.2.15)$$

Уравнения (2.2.13) и (2.2.15) дают $\mathcal{H}_x = \mathcal{H}_z = 0$. Расписывая уравнение (2.2.11) и замечая, что $\sigma = 0$, получаем

$$-\partial \mathcal{H}_y / \partial z = \varepsilon \partial \mathcal{E}_x / \partial t. \quad (2.2.16)$$

Эти уравнения можно объединить, если продифференцировать уравнение (2.2.14) по z , а уравнение (2.2.16) — по t , полагая при этом, что ε и μ_0 не зависят от x, y, z и t . В результате

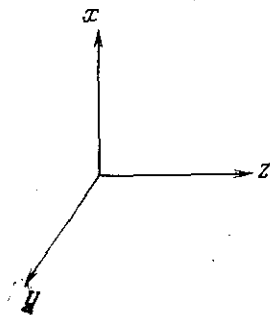


Рис. 2.2.1. Прямоугольная система координат.

этих операций будем иметь

$$\partial^2 \mathcal{E}_x / \partial z^2 = -\mu_0 \partial^2 \mathcal{H}_y / \partial z \partial t, \quad (2.2.17)$$

$$-\partial^2 \mathcal{H}_y / \partial z \partial t = \epsilon \partial^2 \mathcal{E}_x / \partial t^2. \quad (2.2.18)$$

Подставляя второе из этих уравнений в первое для устранения \mathcal{H}_y , получим волновое уравнение в виде

$$\partial^2 \mathcal{E}_x / \partial z^2 = \mu_0 \epsilon \partial^2 \mathcal{E}_x / \partial t^2. \quad (2.2.19)$$

Рассмотрим теперь представление величин, синусоидально изменяющихся во времени. В линейных системах, рассматриваемых здесь, волна любой формы может быть разложена на синусоидальные фурье-компоненты. В прямоугольных координатах эти синусоидальные электрическое и магнитное поля зависят от пространственных координат x , y , z и от времени t и обычно записываются в комплексной форме:

$$F(x, y, z, t) = \text{Re} [f(x, y, z) \exp(j\omega t)], \quad (2.2.20)$$

где Re обозначает реальную часть от выражения в квадратных скобках. Обычно Re не пишут, но всегда помнят, что рассматривается только действительная часть. Поэтому экспоненциальное и тригонометрическое представления используются попеременно. Круговая частота ω в выражении (2.2.20) определяется соотношением

$$\omega = 2\pi\nu, \quad (2.2.21)$$

где ν — частота.

В уравнении (2.2.19) \mathcal{E}_x является функцией от z и t и обычно записывается в виде

$$\mathcal{E}_x = \mathcal{E}_x(z, t) = \mathcal{E}_x(z) \exp(j\omega t). \quad (2.2.22)$$

В этом выражении индекс указывает на направление вектора, а аргумент z , являющийся пространственной переменной, часто опускают. Подставляя выражение (2.2.22) в уравнение (2.2.19), получаем

$$\partial^2 \mathcal{E}_x / \partial z^2 = -\mu_0 \epsilon \omega^2 \mathcal{E}_x. \quad (2.2.23)$$

Как будет показано ниже, величина $\omega(\mu_0 \epsilon)^{1/2}$ играет особую роль, и обычно ее обозначают символом β :

$$\beta^2 = \omega^2 \mu_0 \epsilon, \quad (2.2.24)$$

так что

$$\partial^2 \mathcal{E}_x / \partial z^2 = -\beta^2 \mathcal{E}_x. \quad (2.2.25)$$

Это уравнение имеет решение

$$\mathcal{E}_x(z, t) = [A \exp(-\beta z) + B \exp(\beta z)] \exp(j\omega t). \quad (2.2.26)$$

В дальнейшем мы увидим, что член $\exp(-\beta z)$ представляет волну, распространяющуюся в положительном направлении оси z ,

в то время как $\exp(j\beta z)$ определяет волну, распространяющуюся в отрицательном направлении той же оси. Оставляя только волну, распространяющуюся в направлении $+z$, решение для \mathcal{E}_x можем записать в виде

$$\mathcal{E}_x(z, t) = A \cos(\omega t - \beta z). \quad (2.2.27)$$

Подставляя это решение в уравнение (2.2.16), легко находим

$$\mathcal{H}_y(z, t) = (v\omega A/\beta) \cos(\omega t - \beta z). \quad (2.2.28)$$

Анализ выражений (2.2.27) и (2.2.28) даст ясное представление о свойствах распространяющихся волн и позволит установить соотношение между v и \bar{n} .

Наглядная картина распространения плоской волны может быть получена, если построить, используя выражения (2.2.27) и (2.2.28), координатную зависимость векторов \mathcal{E}_x и \mathcal{H}_y при разных фиксированных значениях времени t . Положим сначала $t = 0$. В этом случае аргумент косинуса будет просто равен βz . Соответствующая моменту времени $t = 0$ зависимость векторов электрического и магнитного полей от z изображена на рис. 2.2.2, а, на котором z отложено в единицах π/β . На рис. 2.2.2, б t положено равным $\pi/4\omega$, чтобы проиллюстрировать движение волны в положительном направлении оси z . При этом аргумент косинуса равен $\pi/4 - \beta z$, т. е. и максимум поля сместился в направлении z на расстояние $\pi/4\beta$. При возрастании t до $2\pi/\omega$ происходит перемещение волны на расстояние, равное одной длине волны, как это показано на рис. 2.2.2, в.

Из рис. 2.2.2 следует, что длина волны равна

$$\lambda = 2\pi/\beta, \quad (2.2.29)$$

откуда

$$\beta = 2\pi/\lambda. \quad (2.2.30)$$

Величина $2\pi/\lambda$ называется волновым числом, а β — постоянной распространения (она дает изменение фазы на единицу длины). Из рис. 2.2.2 следует также, что волна проходит расстояние λ за время $2\pi/\omega$, так что скорость ее распространения v определяется хорошо известным соотношением

$$v = \lambda/(2\pi/\omega) = \lambda v. \quad (2.2.31)$$

Подставляя в это соотношение $2\pi/\beta$ вместо λ и затем $\omega(\mu_0\varepsilon)^{1/2}$ вместо β [из (2.2.24)], получаем

$$v = 1/(\mu_0\varepsilon)^{1/2}. \quad (2.2.32)$$

В свободном пространстве $\varepsilon = \varepsilon_0$ и $v = c$, следовательно

$$c = 1/(\mu_0\varepsilon_0)^{1/2}. \quad (2.2.33)$$

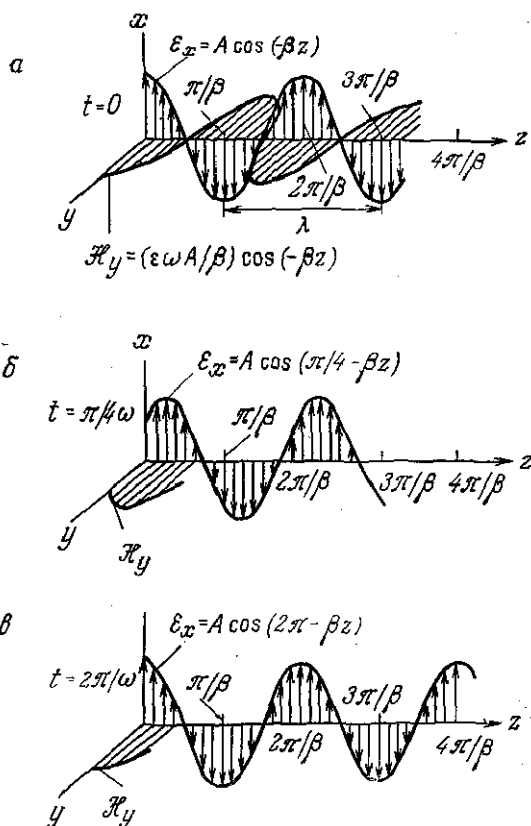


Рис. 2.2.2. Распространение бегущей волны в среде без потерь, определяемое уравнениями (2.2.7) и (2.2.8). а — $t=0$; б — $t=\pi/4\omega$; в — $t=2\pi/\omega$.

Отношение скорости распространения волны в свободном пространстве к ее скорости распространения в диэлектрической среде называется показателем преломления \bar{n} :

$$\bar{n} = (\epsilon/\epsilon_0)^{1/2}. \quad (2.2.34)$$

Частота излучения ν в диэлектрической среде имеет то же значение, что и в свободном пространстве, поэтому из формулы (2.2.31) получаем

$$\nu = v/\lambda = c/\lambda_0, \quad (2.2.35)$$

где λ_0 — длина волны в свободном пространстве. Из выражений (2.2.32) — (2.2.35) для длины волны в диэлектрической среде имеем

$$\lambda = \lambda_0 (\epsilon_0/\epsilon)^{1/2} = \lambda_0/\bar{n}, \quad (2.2.36)$$

причем равенство $\lambda = \lambda_0/\bar{n}$ используют довольно часто. Подставляя в соотношение (2.2.24) вместо ϵ величину $\bar{n}^2\epsilon_0$ и заменяя ϵ_0 на $1/c^2\mu_0$, получаем

$$\beta^2 = (2\pi\nu/c)^2 \bar{n}^2 = (2\pi/\lambda_0)^2 \bar{n}^2, \quad (2.2.37)$$

где учтено, что $\omega = 2\pi\nu$ по определению. Следовательно, если постоянную распространения в свободном пространстве определить как

$$k_0 = 2\pi/\lambda_0, \quad (2.2.38)$$

то соотношение (2.2.37) можно переписать в виде

$$\beta^2 = \bar{n}^2 k_0^2, \quad (2.2.39)$$

в котором постоянная распространения выражена через оптическую постоянную \bar{n} .

В среде с потерями электрическая проводимость уже не равна нулю. Как и в случае диэлектрика без потерь, мы рассмотрим однородную плоскую волну с $\mathcal{E} = a_x \mathcal{E}_x$, однако уравнение (2.2.11) теперь преобразуется к виду

$$-\partial \mathcal{H}_y / \partial z = \sigma \mathcal{E}_x + \epsilon (\partial \mathcal{E}_x / \partial t). \quad (2.2.40)$$

Дифференцируя это уравнение по t , а уравнение (2.2.14) — по z , получаем волновое уравнение

$$\partial^2 \mathcal{E}_x / \partial z^2 = \mu_0 \sigma (\partial \mathcal{E}_x / \partial t) + \mu_0 \epsilon (\partial^2 \mathcal{E}_x / \partial t^2). \quad (2.2.41)$$

Если считать, что временная зависимость $\mathcal{E}_x(z, t)$, как и в выражении (2.2.22), определяется множителем $\exp(j\omega t)$, то из уравнения (2.2.41) следует

$$\partial^2 \mathcal{E}_x / \partial z^2 = j\omega \mu_0 (\sigma + j\omega \epsilon) \mathcal{E}_x = \Gamma^2 \mathcal{E}_x. \quad (2.2.42)$$

В этом уравнении Γ обозначает комплексную постоянную распространения:

$$\Gamma = [j\omega \mu_0 (\sigma + j\omega \epsilon)]^{1/2} = \gamma + j\beta. \quad (2.2.43)$$

Волновое уравнение (2.2.42) имеет решение

$$\mathcal{E}_x(z, t) = [A \exp(-\Gamma z) + B \exp(\Gamma z)] \exp(j\omega t), \quad (2.2.44)$$

которое при подстановке Γ из (2.2.43) преобразуется к виду

$$\mathcal{E}_x(z, t) = \{A \exp(-\gamma z) \exp[j(\omega t - \beta z)] + B \exp(\gamma z) \exp[j(\omega t + \beta z)]\}. \quad (2.2.45)$$

Пренебрегая снова членом, который описывает волну, распространяющуюся в направлении $-z$, получаем

$$\mathcal{E}_x(z, t) = A \exp(-\Gamma z) \exp(j\omega t) = A \exp(-\gamma z) \cos(\omega t - \beta z) \quad (2.2.46)$$

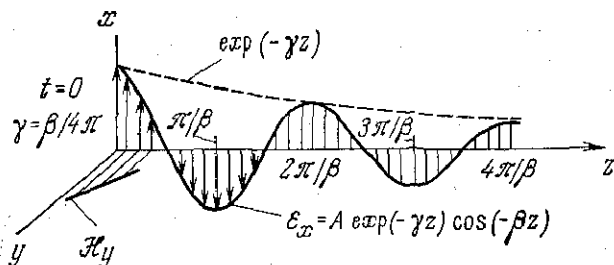


Рис. 2.2.3. Распространение бегущей волны в среде с потерями, определяемое выражениями (2.2.46) и (2.2.48).

Если взять решение (2.2.46) в экспоненциальной форме, продифференцировать его по t и проинтегрировать по z , то из уравнения (2.2.40) легко получить решение для \mathcal{H}_y :

$$\mathcal{H}_y = [(\sigma + j\omega\epsilon)/\Gamma] \mathcal{E}_x. \quad (2.2.47)$$

Из соотношения (2.2.43) следует, что $\sigma + j\omega\epsilon = \Gamma^2/j\omega\mu_0$, и решение (2.2.47) может быть переписано в виде

$$\mathcal{H}_y = (\Gamma/j\omega\mu_0) A \exp(-\gamma z) \cos(\omega t - \beta z). \quad (2.2.48)$$

Из выражений (2.2.46) и (2.2.48) видно, что γ характеризует затухание волны по мере ее распространения, как это показано на рис. 2.2.3.

В случае среды без потерь множитель при $\mathcal{E}_x(z, t)$ в волновом уравнении (2.2.23) представлялся в виде $\bar{n}^2 k_0^2$. Аналогичное выражение используют в случае среды с потерями:

$$\Gamma^2 = j\omega\mu_0(\sigma + j\omega\epsilon) = -\bar{N}^2 k_0^2, \quad (2.2.49)$$

где \bar{N} — комплексный показатель преломления с действительной частью \bar{n} и мнимой частью \bar{k} :

$$\bar{N} = \bar{n} - j\bar{k}. \quad (2.2.50)$$

(Следует заметить, что если бы временная зависимость выбиралась в виде $\exp(-j\omega t)$, то \bar{N} определялось бы как $\bar{n} + j\bar{k}$.) Объединяя выражения (2.2.49) и (2.2.50) и учитывая, что $k_0 = 2\pi/\lambda_0$, получаем

$$\epsilon/\epsilon_0 - j\sigma/\omega\epsilon_0 = \bar{n}^2 - \bar{k}^2 - j2\bar{n}\bar{k}. \quad (2.2.51)$$

Комплексная величина $(\epsilon - j\sigma/\omega)/\epsilon_0$ называется комплексной диэлектрической проницаемостью. Разделение действительной и мнимой частей в формуле (2.2.51) дает

$$\bar{n}^2 - \bar{k}^2 = \epsilon/\epsilon_0, \quad (2.2.52)$$

$$2\bar{n}\bar{k} = \sigma/\omega\epsilon_0. \quad (2.2.53)$$

Действительную часть комплексной диэлектрической проницаемости, определяемую выражением (2.2.52), и мнимую часть, определяемую выражением (2.2.53), часто обозначают ϵ_1 и ϵ_2 соответственно. Когда $\sigma = 0$, величина ϵ/ϵ_0 равна \bar{n}^2 , как это и следует из выражения (2.2.34) для диэлектрика без потерь. Из соотношений (2.2.43), (2.2.49) и (2.2.50) имеем

$$\Gamma = \gamma + j\beta = j(\bar{n} - j\bar{k})k_0, \quad (2.2.54)$$

откуда следует

$$\gamma = 2\pi\bar{k}/\lambda_0, \quad (2.2.55)$$

$$\beta = \bar{n}k_0. \quad (2.2.56)$$

Поэтому затухание связано с наличием проводимости, а постоянная распространения, как и для среды без потерь, зависит от \bar{n} . Величина \bar{k} называется коэффициентом экстинкции.

На практике обычно определяют коэффициент поглощения α , а не коэффициент экстинкции. Средний поток мощности через единицу площади определяется действительной частью комплексного вектора Пойнтинга и при экспоненциальной форме записи векторов \mathcal{E} и \mathcal{H} равен [7]

$$P = |\mathbf{S}| = \frac{1}{2} \operatorname{Re}(\mathcal{E} \times \mathcal{H}^*), \quad (2.2.57)$$

где звездочкой обозначено комплексное сопряжение. Используя выражения (2.2.46) для \mathcal{E} и (2.2.47) для \mathcal{H} , выражение для потока мощности можно записать в виде

$$P = I = I_0 \exp(-2\gamma z), \quad (2.2.58)$$

где все сомножители, стоящие перед $\exp(-2\gamma z)$, объединены в I_0 . Коэффициент поглощения α представляет собой относительную потерю интенсивности dI/I на бесконечно малой длине dz , отнесенную к этой длине:

$$\alpha = -\frac{dI/I}{dz}. \quad (2.2.59)$$

Из этого выражения с учетом (2.2.58) получаем

$$\alpha = 2\gamma \quad (2.2.60)$$

или, подставляя γ из выражения (2.2.55), находим

$$\alpha = 4\pi\bar{k}/\lambda_0. \quad (2.2.61)$$

Таким образом, соотношения (2.2.52), (2.2.53) и (2.2.61) связывают электрические постоянные ϵ и σ с оптическими постоянными \bar{n} и α . Следует заметить, что эти величины зависят от частоты в области оптических частот.

§ 3. ПЛОСКИЕ ДИЭЛЕКТРИЧЕСКИЕ ВОЛНОВОДЫ

✓ Гомолазеры

Гомолазер, показанный на рис. 1.4.1, имеет тонкую излучающую свет область, лежащую с p -стороны $p-n$ -перехода. Эта область называется активной областью и ограничивается с обеих сторон пассивными областями, которые характеризуются значительными потерями. Для гомолазера на GaAs потери при генерации обычно составляют 60 см^{-1} [8]. Такие потери малы по сравнению с ожидаемыми потерями для генерируемого излучения с энергией фотонов, близкой к энергии запрещенной зоны, которые являются следствием объемного поглощения в GaAs. Поэтому амплитуда оптического поля должна уменьшаться в направлении, перпендикулярном плоскости $p-n$ -перехода, вследствие отражения на границе диэлектрика, а не из-за потерь. Таким образом, электромагнитное излучение в гомолазере распространяется параллельно лазерной оси. Структура, которая осуществляет направленное распространение потока электромагнитной энергии параллельно своей оси, называется волноводом. Описанию волноводного эффекта в гомолазерах посвящены работы [9—13].

Для того чтобы в трехслойном плоском диэлектрическом волноводе волна распространялась главным образом в центральном слое, он должен иметь диэлектрическую проницаемость, более высокую, чем диэлектрические проницаемости наружных слоев [13]. В гомолазере трудно количественно оценить разность показателей преломления центральной излучающей области и прилегающих пассивных областей. Концентрации примесей, распределение температуры, концентрация инжектированных свободных носителей — все это влияет на профиль показателя преломления. Однако эти влияния малы и, как показано ниже, приводят к изменениям показателя преломления на величину порядка 10^{-3} . В итоге показатель преломления в центральном слое гомолазера больше показателей преломления окружающих его пассивных областей на величину $\sim 10^{-3}$.

В GaAs при равновесных концентрациях электронов n_0 и дырок p_0 до $1 \cdot 10^{18} \text{ см}^{-3}$ при комнатной температуре значения \bar{n} на длине волны генерации примерно одинаковы для материала как p -, так и n -типа [14]. При n_0 и p_0 больше $1 \cdot 10^{18} \text{ см}^{-3}$ \bar{n} в p -GaAs становится немного больше, чем в n -GaAs. Сравнение этих величин проводится в § 5 этой главы и иллюстрируется на рис. 2.5.3, где приведены значения \bar{n} для излучения с энергией фотонов 1,38 эВ при комнатной температуре, взятые из работы [14]. При больших токах инжекции влияние температурных градиентов, концентраций инжектированных носителей и уси-

ления становится значительным. Температура в области $p-n$ перехода из-за безызлучательной рекомбинации и поглощения на свободных носителях может возрастать. Согласно данным Марпла [15], изменение показателя преломления с температурой можно записать в виде

$$\delta\bar{n}_T \approx 4 \cdot 10^{-4} \Delta T, \quad (2.3.1)$$

где ΔT — изменение температуры. Следовательно, для увеличения показателя преломления на величину $\delta\bar{n}_T = 10^{-3}$ необходимо повышение температуры всего на несколько градусов. Инжектированные свободные электроны уменьшают показатель преломления на величину [16]

$$\delta\bar{n}_{\text{св. н}} = -(nq^2/2m_n\omega^2\epsilon)\bar{n}. \quad (2.3.2)$$

Здесь n — концентрация свободных электронов, q — заряд электрона, m_n — эффективная масса электрона. Аналогичное выражение может быть написано в случае материала p -типа с концентрацией дырок p и их эффективной массой m_p . В GaAs имеем $q^2/\epsilon = 2,21 \cdot 10^{-19}$ эрг·см, $\epsilon = 13,1\epsilon_0$ [17], $m_n = 0,067m_0$, где m_0 — масса электрона, равная $9 \cdot 10^{-28}$ г, и для $\lambda_0 = 0,9$ мкм $\omega^2 = 4,4 \cdot 10^{30}$ (рад/с)². Вблизи $\lambda_0 = 0,9$ мкм при комнатной температуре $\bar{n} \approx 3,6$ [15]. Эти значения дают $\delta\bar{n}_{\text{св. н}} = -1,5 \cdot 10^{-21} n$. В § 7 гл. 3 показано, что для усиления необходима концентрация носителей $\sim 1 \cdot 10^{18}$ см⁻³, поэтому $\delta\bar{n}_{\text{св. н}} \approx -1,5 \cdot 10^{-3}$. Поскольку $m_p > m_n$ [18], значение $\delta\bar{n}_{\text{св. н}}$ для электронов будет больше, чем для дырок.

Мнимая часть комплексного показателя преломления в выражении (2.2.50) определяется коэффициентом экстинкции, связанным с коэффициентом поглощения соотношением (2.2.61). В § 8 гл. 3 показано, что коэффициент поглощения становится отрицательным в условиях, когда происходит генерация лазерного излучения. В GaAs-гомолазерах коэффициент усиления ($-\alpha$) порядка 100 см⁻¹, что дает после подстановки в (2.2.61) значение мнимой части \bar{N} , равное $\sim 10^{-3}$. Это изменение \bar{N} может обеспечить волноводное распространение излучения за счет усиления [16, 19, 20]. Таким образом, большие токи инжекции в силу разных эффектов одновременно приводят и к увеличению, и к уменьшению показателя преломления GaAs, причем примерно на одинаковую величину. В гомолазере разность между показателями преломления центральной области, излучающей свет, и прилегающими пассивными областями мала, и поэтому оптическое ограничение мод в гомолазере слабо.

Гетеролазеры

В ОГС-лазере формируется более сильный волновод, чем в гомолазере [21, 22]. На границе раздела между p -GaAs и

слоем $\text{Al}_x\text{Ga}_{1-x}\text{As}$ существует резкий скачок показателя преломления \bar{n} по крайней мере на 10^{-2} , в то время как изменение \bar{n} на $p-n$ -переходе такое же, как и в гомоструктуре. Таким образом, волновод состоит из центрального активного слоя $p\text{-GaAs}$, граничащего с $n\text{-GaAs}$ с одной стороны и с $p\text{-Al}_x\text{Ga}_{1-x}\text{As}$ — с другой. Распределение оптического поля в ОГС-лазере было рассмотрено Байером и Батлером [23]. Поскольку скачки показателя преломления на $p-n$ -переходе и на границе $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ различаются между собой по меньшей мере в 10 раз, распределение оптического поля несимметрично, и область, занимаемая излучением, расширена в сторону $n\text{-GaAs}$. В § 6 этой главы мы рассмотрим асимметричный трехслойный плоский волновод.

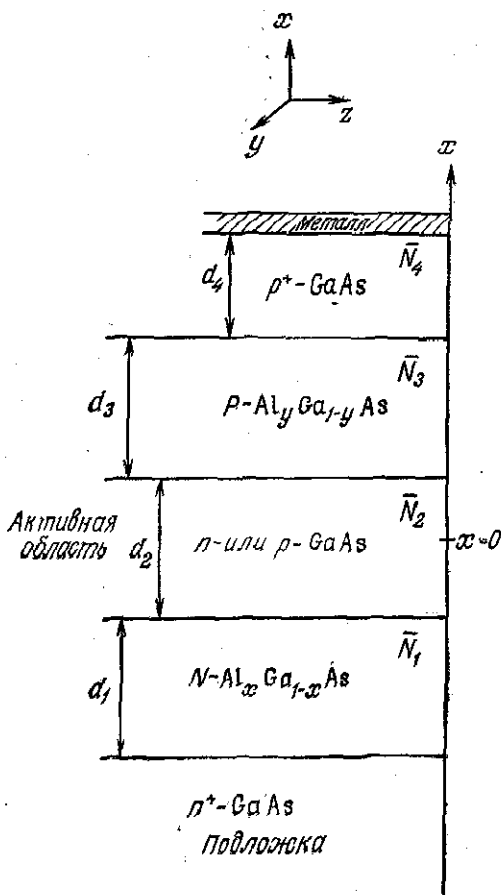


Рис. 2.3.1. Схематическое изображение слоев $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ ДГС-лазера.

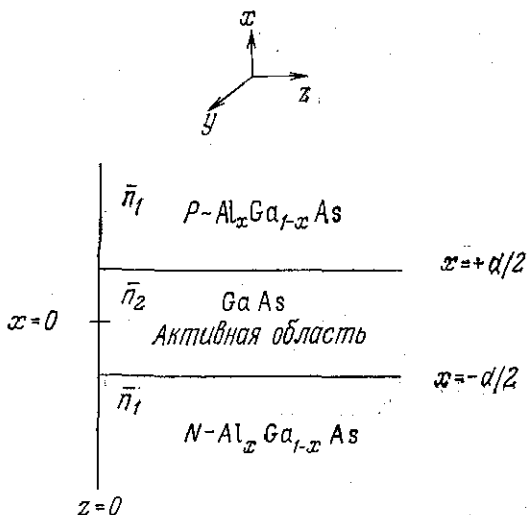


Рис. 2.3.2. Схематическое представление GaAs—Al_xGa_{1-x}As ДГС-лазера как симметричного трехслойного плоского диэлектрического волновода.

В ДГС-лазере образуется еще более сильный диэлектрический волновод. Лазеры такого типа были первыми инжекционными лазерами, которые заработали в непрерывном режиме при комнатной температуре [24, 25]. ДГС-лазеры количественно описать легче, чем гомолазеры или ОГС-лазеры, так как изменения \bar{n} на гетерограницах достаточно велики, чтобы не учитывать влияния концентрации инжектированных носителей, волноводного распространения за счет усиления и температурных градиентов. Анализ распространения волн в GaAs—Al_xGa_{1-x}As ДГС-лазерах проведен в работах [26—28]. Поскольку точные данные, описывающие зависимости показателя преломления GaAs от концентрации в нем примеси и показателя преломления Al_xGa_{1-x}As от молярной доли AlAs x , не были известны, в этих ранних вычислениях для получения численных результатов использовались оценочные значения скачков \bar{n} . Наличие в настоящее время экспериментальных значений \bar{n} позволило численно рассчитать распределения оптического поля в GaAs—Al_xGa_{1-x}As ДГС-лазерах. В § 4—10 настоящей главы мы используем эти экспериментальные данные как при анализе различных вопросов, так и в численных примерах.

На рис. 2.3.1 показана обычная конфигурация слоев в ДГС-лазере. Удобно обозначать слои n - и p -Al_xGa_{1-x}As буквами N и P , а слои n - и p -GaAs — буквами n и p . Молярная доля AlAs обозначается через x или y . Наличие верхнего слоя p -GaAs способствует созданию омических контактов. Комплекс-

ные показатели преломления обозначены через N_1 , N_2 , N_3 и N_4 . Батлер и др. [29], а также Стрейфер и др. [30] провели анализ волноводного распространения излучения в такой многослойной структуре. В правильно рассчитанном ДГС-лазере толщины слоев d_1 и d_3 достаточно велики для того, чтобы исключить влияние границ с контактным слоем p^+ -GaAs или с n^+ -подложкой на распределение оптического поля. На самом деле для описания волноводных свойств ДГС-лазеров нет необходимости изучать столь сложную структуру, как изображенная на рис. 2.3.1. Более подходящей для этой цели структурой является показанный на рис. 2.3.2 симметричный трехслойный плоский диэлектрический волновод. Слои $Al_xGa_{1-x}As$ с каждой стороны от активной области имеют одинаковую молярную долю AlAs. Эта структура рассматривается в следующих двух параграфах.

§ 4. РАСПРОСТРАНЕНИЕ ВОЛН В СИММЕТРИЧНОМ ТРЕХСЛОЙНОМ ПЛОСКОМ ВОЛНОВОДЕ

Волновое уравнение

В этой части параграфа из уравнений Максвелла будет получено приведенное волновое уравнение. Маркузе [1, 3] дал полное описание конфигураций поля (мод) в плоском волноводе, и мы, где это возможно, будем придерживаться терминологии и методов вывода, использованных в работах [1, 3]. При ознакомлении с теорией плоского диэлектрического волновода очень полезна также работа Капани и Бурке [2]. Мы рассмотрим как ТЕ-моды, в которых электрическое поле поляризовано перпендикулярно направлению распространения, так и ТМ-моды, в которых перпендикулярно направлению распространения поляризовано магнитное поле.

Волновое уравнение для вектора электрического поля можно получить, взяв ротор от обеих частей уравнения (2.2.10). Это дает

$$\nabla \times \nabla \times \mathcal{E} = -\mu_0 (\partial/\partial t) \nabla \times \mathcal{H}. \quad (2.4.1)$$

Используя уравнение (2.2.11) для исключения члена с $\nabla \times \mathcal{H}$, для случая $\sigma = 0$ получим

$$\nabla \times \nabla \times \mathcal{E} = -\mu_0 \epsilon \partial^2 \mathcal{E} / \partial t^2. \quad (2.4.2)$$

Левую часть уравнения (2.4.2) можно преобразовать, применив тождество [6]

$$\nabla \times \nabla \times \mathcal{E} = \nabla (\nabla \cdot \mathcal{E}) - \nabla^2 \mathcal{E}, \quad (2.4.3)$$

где ∇^2 — оператор Лапласа, определяемый выражением

$$\nabla^2 = (\partial^2/\partial x^2) + (\partial^2/\partial y^2) + (\partial^2/\partial z^2). \quad (2.4.4)$$

Из уравнения (2.2.6) следует, что $\nabla \cdot \mathcal{D} = \nabla \cdot \varepsilon \mathcal{E} = 0$, когда плотность заряда для оптического поля равна нулю. Тогда [31]

$$\nabla \cdot \varepsilon \mathcal{E} = \mathcal{E} \cdot \nabla \varepsilon + \varepsilon \nabla \cdot \mathcal{E} = 0, \quad (2.4.5)$$

$$\nabla \cdot \mathcal{E} = -(\mathcal{E}/\varepsilon) \cdot \nabla \varepsilon. \quad (2.4.6)$$

Для рассматриваемых здесь случаев ε считается постоянной внутри данного слоя, так что $\nabla \varepsilon = 0$. Следовательно, $\nabla \cdot \mathcal{E} = 0$, и из тождества (2.4.3) имеем

$$\nabla \times \nabla \times \mathcal{E} = -\nabla^2 \mathcal{E}. \quad (2.4.7)$$

Тогда уравнение (2.4.2) может быть записано в виде

$$\nabla^2 \mathcal{E} = \mu_0 \varepsilon \partial^2 \mathcal{E} / \partial t^2. \quad (2.4.8)$$

Уравнение (2.4.8) является трехмерным векторным уравнением для вектора электрического поля $\mathcal{E} = a_x \mathcal{E}_x + a_y \mathcal{E}_y + a_z \mathcal{E}_z$. Оно распадается на три отдельных независимых скалярных волновых уравнения:

$$\nabla^2 \mathcal{E}_x = \mu_0 \varepsilon \partial^2 \mathcal{E}_x / \partial t^2, \quad (2.4.9)$$

$$\nabla^2 \mathcal{E}_y = \mu_0 \varepsilon \partial^2 \mathcal{E}_y / \partial t^2, \quad (2.4.10)$$

$$\nabla^2 \mathcal{E}_z = \mu_0 \varepsilon \partial^2 \mathcal{E}_z / \partial t^2. \quad (2.4.11)$$

Аналогичным образом применяя операцию ротора к обеим частям уравнения (2.2.11) и используя уравнение (2.2.10) для устранения члена, содержащего \mathcal{E} , для вектора магнитного поля получаем уравнение

$$\nabla^2 \mathcal{H} = \mu_0 \varepsilon \partial^2 \mathcal{H} / \partial t^2. \quad (2.4.12)$$

Волновое уравнение для ТЕ-мод

Для структуры, показанной на рис. 2.3.2, значение $z = 0$ соответствует границе раздела полупроводник — воздух, а $x = 0$ соответствует центру активного слоя. Волновод считается бесконечным и однородным в направлении оси y , так что $\partial/\partial y = 0$. Для ТЕ-волн $\mathcal{E}_z = 0$ по определению, и, учитывая условие $\partial/\partial y = 0$, из уравнений (2.2.10) и (2.2.11) получаем $\mathcal{H}_y = \mathcal{E}_x = 0$. Следовательно, из уравнений (2.4.9) — (2.4.11) остается только уравнение (2.4.10), которое принимает вид

$$\partial^2 \mathcal{E}_y / \partial x^2 + \partial^2 \mathcal{E}_y / \partial z^2 = \mu_0 \varepsilon \partial^2 \mathcal{E}_y / \partial t^2. \quad (2.4.13)$$

Это уравнение представляет собой дифференциальное уравнение, описывающее распространение ТЕ-волны в плоском волноводе. Из него следует, что \mathcal{E}_y зависит от двух пространственных переменных x и z , а также от времени t .

Уравнение (2.4.13) легко решить, используя метод разделения переменных [32]. Решения ищутся в виде

$$\mathcal{E}_y(x, z, t) = XZT, \quad (2.4.14)$$

где X представляет собой функцию только x , Z зависит только от z , а T — только от t . Подставляя выражение (2.4.14) в уравнение (2.4.13), получаем уравнение

$$X''ZT + XZ''T = \mu_0\epsilon XZT'', \quad (2.4.15)$$

которое после деления обеих частей на XZT дает

$$X''/X + Z''/Z = \mu_0\epsilon(T''/T), \quad (2.4.16)$$

где двумя штрихами обозначены вторые производные функций X , Z и T по соответствующим аргументам.

Левая часть уравнения (2.4.16) не содержит t и, следовательно, не может меняться при изменении t . Поэтому правая часть уравнения (2.4.16) также не зависит от t и должна равняться некоторой константе, которую мы обозначим через $-b^2$:

$$X''/X + Z''/Z = \mu_0\epsilon(T''/T) = -b^2, \quad (2.4.17)$$

или

$$T'' + (b^2/\mu_0\epsilon)T = 0. \quad (2.4.18)$$

Решение этого уравнения имеет вид

$$T = A \exp[j(b/(\mu_0\epsilon)^{1/2})t] + B \exp[-j(b/(\mu_0\epsilon)^{1/2})t]. \quad (2.4.19)$$

При синусоидальной зависимости от времени $\exp[j(b/(\mu_0\epsilon)^{1/2})t] = \exp(j\omega t)$ и

$$b^2 = \omega^2\mu_0\epsilon. \quad (2.4.20)$$

Как и при рассмотрении плоских волн, мы будем использовать только решение с $+j\omega t$. Из уравнения (2.4.17) и соотношения (2.4.20) следует

$$(X''/X) + (Z''/Z) = -\omega^2\mu_0\epsilon. \quad (2.4.21)$$

Проводя те же рассуждения, что и при выводе уравнения (2.4.17), из уравнения (2.4.21) получаем

$$Z''/Z = -(X''/X) - \omega^2\mu_0\epsilon = -\beta^2, \quad (2.4.22)$$

где β^2 — постоянная разделения. Решение этого уравнения дает

$$Z = C \exp(j\beta z) + D \exp(-j\beta z). \quad (2.4.23)$$

Это общее решение будет использовано при дальнейшем анализе распространения волн в волноводе.

Как и при рассмотрении вопроса о распространении плоских волн, мы сохраним член с $\exp(-j\beta z)$, представляющий волну, распространяющуюся в положительном направлении оси z . Уравнение (2.4.21) теперь принимает вид

$$(X''/X) + \omega^2\mu_0\varepsilon - \beta^2 = 0, \quad (2.4.24)$$

или
$$(\partial^2 \mathcal{E}_y / \partial x^2) + (\omega^2\mu_0\varepsilon - \beta^2) \mathcal{E}_y = 0, \quad (2.4.25)$$

где
$$\mathcal{E}_y(x, z, t) = \mathcal{E}_y(x) \exp[j(\omega t - \beta z)]. \quad (2.4.26)$$

Из выражений (2.2.33) — (2.2.35) и (2.2.38) следует

$$\omega^2\mu_0\varepsilon = \bar{n}^2 k_0^2. \quad (2.4.27)$$

Учитывая это соотношение, из уравнения (2.4.25) получаем волновое уравнение в форме, наиболее удобной для исследования ТЕ-мод:

$$(\partial^2 \mathcal{E}_y / \partial x^2) + (\bar{n}^2 k_0^2 - \beta^2) \mathcal{E}_y = 0. \quad (2.4.28)$$

Решение этого уравнения имеет вид

$$\mathcal{E}_y(x) = A_e \cos \kappa x + A_o \sin \kappa x, \quad (2.4.29)$$

где
$$\kappa^2 = \bar{n}^2 k_0^2 - \beta^2. \quad (2.4.30)$$

Непосредственной подстановкой легко убедиться, что выражение (2.4.29) действительно является решением уравнения (2.4.28), если κ определяется соотношением (2.4.30).

Граничные условия

Чтобы получить решения для электрического и магнитного полей, необходимо установить соотношения, которым должны удовлетворять электромагнитные поля на границах раздела диэлектриков. Эти соотношения называются граничными условиями и могут быть получены из уравнений Максвелла. Интегрирование обеих частей уравнения (2.2.1) по поверхности s дает

$$\int_s \nabla \times \mathcal{E} \cdot da = - \frac{\partial}{\partial t} \int_s \mathcal{B} \cdot da, \quad (2.4.31)$$

где da — векторный элемент площади поверхности. Теорема Стокса [33] позволяет переписать левую часть уравнения (2.4.31) в виде

$$\int_s \nabla \times \mathcal{E} \cdot da = \oint_c \mathcal{E} \cdot dl, \quad (2.4.32)$$

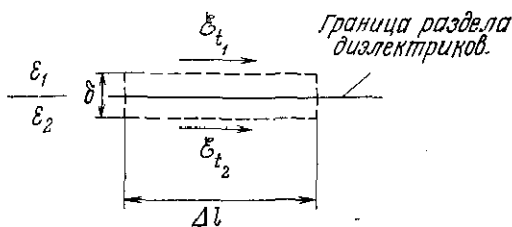


Рис. 2.4.1. Картина, поясняющая вывод граничных условий для тангенциальной составляющей электрического поля на границе раздела диэлектриков с диэлектрическими проницаемостями ϵ_1 и ϵ_2 . Вдоль границы интегрирование ведется по отрезкам длины Δl , перпендикулярно границе — по бесконечно малым отрезкам δ .

где c — контур, ограничивающий поверхность s , а $d\mathbf{l}$ — векторный элемент длины этого контура.

На рис. 2.4.1 показана граница раздела двух диэлектриков с диэлектрическими проницаемостями ϵ_1 и ϵ_2 . Тангенциальные составляющие электрических полей в средах 1 и 2 обозначены через \mathcal{E}_{t1} и \mathcal{E}_{t2} соответственно. Замкнутый контур интегрирования образуется при движении на расстояние Δl вдоль прямой, лежащей с одной стороны границы между ϵ_1 и ϵ_2 , и возврате вдоль прямой, лежащей с другой стороны границы, причем расстояние δ между двумя слоями, в которых лежат эти прямые, считается бесконечно малым. Линейный интеграл в уравнении (2.4.32) дает

$$\oint_c \mathcal{E} \cdot d\mathbf{l} = (\mathcal{E}_{t1} - \mathcal{E}_{t2}) \Delta l = 0, \quad (2.4.33)$$

поскольку левая часть уравнения (2.4.32) равна нулю, когда площадь, ограничиваемая контуром интегрирования, стремится к нулю при бесконечно малом δ . Поэтому граничное условие требует, чтобы тангенциальные составляющие электрического поля были равны, т. е.

$$\mathcal{E}_{t1} = \mathcal{E}_{t2}. \quad (2.4.34)$$

Применение теоремы Стокса к интегральной форме уравнения (2.2.2) дает граничное условие для тангенциальных компонент магнитного поля:

$$\mathcal{H}_{t1} = \mathcal{H}_{t2}. \quad (2.4.35)$$

Граничные условия (2.4.34) и (2.4.35) означают, что тангенциальные компоненты электрического и магнитного полей должны быть непрерывны на границе раздела диэлектриков.

Граничные условия для нормальных компонент полей могут быть получены из уравнений (2.2.3) и (2.2.4), однако при даль-

нейшем анализе нам понадобятся только условия для тангенциальных компонент. Подробное обсуждение граничных условий для нормальных компонент дано в работе [34].

Решения для четных ТЕ-мод

Как было отмечено Маркузе [35], анализ распространения волн в волноводе упрощается, если рассматривать отдельно четные и нечетные решения. В этой части параграфа мы будем близко следовать изложению данного вопроса в работе [35]. Внутри активной области структуры, показанной на рис. 2.3.2, выражение для четной моды получается из формул (2.4.26) и (2.4.29) в виде

$$\mathcal{E}_y = A_e \cos(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.36)$$

где x меняется в пределах от $-d/2$ до $+d/2$, а

$$\kappa^2 = \bar{n}_2^2 k_0^2 - \beta^2. \quad (2.4.37)$$

Магнитное поле может быть найдено из a_z -компоненты уравнения (2.2.10), которая имеет вид

$$\partial \mathcal{E}_y / \partial x = -\mu_0 \partial \mathcal{H}_z / \partial t. \quad (2.4.38)$$

При временной зависимости, выражаемой множителем $\exp(j\omega t)$, решение уравнения (2.4.38) имеет вид

$$\mathcal{H}_z = (j/\omega\mu_0) (\partial \mathcal{E}_y / \partial x), \quad (2.4.39)$$

или, учитывая выражение (2.4.36),

$$\mathcal{H}_z = (-j\kappa/\omega\mu_0) A_e \sin(\kappa x) \exp[j(\omega t - \beta z)]. \quad (2.4.40)$$

Компонента \mathcal{H}_x может быть получена аналогичным образом, но для дальнейшего использования граничных условий она не понадобится.

Для существования волноводных мод необходимо, чтобы поле затухало снаружи активного слоя. В уравнении (2.4.28) величина $\bar{n}_1^2 k_0^2$ должна быть меньше, чем β^2 , чтобы экспоненциальные решения этого уравнения имели действительный, а не мнимый аргумент. Тогда снаружи активного слоя, где $|x| > d/2$, компоненты поля будут выражаться следующими формулами:

$$\mathcal{E}_y = A_e \cos(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)] \quad (2.4.41)$$

и, согласно (2.4.39),

$$\mathcal{H}_z = (-x/|x|) (j\gamma/\omega\mu_0) A_e \cos(\kappa d/2) \times \\ \times \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (2.4.42)$$

где

$$\gamma^2 = \beta^2 - \bar{n}_1^2 k_0^2. \quad (2.4.43)$$

Из соотношений (2.4.37) и (2.4.43) следует, что для существования волноводных мод требуется выполнение условий $\bar{n}_2^2 k_0^2 > \beta^2$ и $\beta^2 > \bar{n}_1^2 k_0^2$, или

$$\bar{n}_2 > \bar{n}_1. \quad (2.4.44)$$

Компоненты поля в области $|x| > d/2$, экспоненциально затухающие без фазового сдвига (или потерь) в поперечной плоскости, называются «хвостами» моды (evanescent waves). В этом случае затухание является реактивным, а не происходит вследствие потерь, что приводило бы к фазовому сдвигу.

Дисперсионное уравнение, являющееся уравнением на собственные значения, вытекает из граничных условий, которым должно удовлетворять электромагнитное поле на границе раздела диэлектриков. Из выражений (2.4.40) и (2.4.42) при учете граничного условия (2.4.35) для \mathcal{H}_z при $x = \pm d/2$ следует

$$\operatorname{tg}(\kappa d/2) = \gamma/\kappa = (\beta^2 - \bar{n}_1^2 k_0^2)^{1/2} / (\bar{n}_2^2 k_0^2 - \beta^2)^{1/2}. \quad (2.4.45)$$

Неизвестной величиной в уравнении (2.4.45) является β . Решить это уравнение относительно β аналитически невозможно, поэтому оно должно решаться графически или численно на машине.

Мощность, переносимая модой в направлении оси z , может быть связана с амплитудным коэффициентом A_e при помощи вектора Пойнтинга (2.2.57). Эта мощность выражается формулой

$$P = \frac{1}{2} \int_{-\infty}^{\infty} (\mathcal{E} \times \mathcal{H}^*) dx = -\frac{1}{2} \int_{-\infty}^{\infty} \mathcal{E}_y \mathcal{H}_x^* dx, \quad (2.4.46)$$

где учтено, что $\mathcal{E}_x = \mathcal{E}_z = \mathcal{H}_y = 0$. Аналогично тому как было получено выражение для \mathcal{H}_z , из уравнений (2.2.10), (2.4.38) и (2.4.39) получаем

$$\mathcal{H}_x = -(j/\omega\mu_0) (\partial \mathcal{E}_y / \partial z). \quad (2.4.47)$$

Тогда из формул (2.4.36), (2.4.47) и (2.4.46) легко получить следующее выражение для мощности:

$$P = \frac{\beta}{\omega\mu_0} \int_0^{\infty} |\mathcal{E}_y|^2 dx. \quad (2.4.48)$$

Подставляя вместо \mathcal{E}_y в (2.4.48) выражения (2.4.36) при $0 < x < d/2$ и (2.4.41) при $d/2 < x < \infty$ и выражая A_e через P , получаем

$$A_e = \left[\frac{2\omega\mu_0}{\beta(d/2) + (\beta/\gamma)} P \right]^{1/2}. \quad (2.4.49)$$

где использовано уравнение (2.4.45).

Решения для нечетных ТЕ-мод

Из выражений (2.4.26) и (2.4.29) следует, что для нечетных мод при $|x| < d/2$ справедливы формулы

$$\mathcal{E}_y = A_0 \sin(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.50)$$

$$\mathcal{H}_z = (j\kappa/\omega\mu_0) A_0 \cos(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.51)$$

а снаружи активного слоя

$$\mathcal{E}_y = (x/|x|) A_0 \sin(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (2.4.52)$$

$$\mathcal{H}_z = (-j\gamma/\omega\mu_0) A_0 \sin(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)]. \quad (2.4.53)$$

Дисперсионное уравнение, соответствующее условию непрерывности \mathcal{H}_z на границе раздела, принимает вид

$$\operatorname{tg}(\kappa d/2) = -\kappa/\gamma. \quad (2.4.54)$$

Можно показать, что A_0 определяется тем же выражением (2.4.49), что и A_e .

Волновое уравнение для ТМ-мод

Для ТМ-мод $\mathcal{H}_z = 0$ по определению, и при учете условия $\partial/\partial y = 0$ уравнения (2.2.10) и (2.2.11) дают $\mathcal{E}_y = \mathcal{H}_x = 0$. Тем же способом, что и для ТЕ-мод, волновое уравнение для ТМ-мод может быть получено в виде

$$(\partial_x^2 \mathcal{H}_y / \partial x^2) + (\bar{n}^2 k_0^2 - \beta^2) \mathcal{H}_y = 0. \quad (2.4.55)$$

Решения для четных ТМ-мод

Внутри активного слоя $|x| < d/2$ четная мода описывается формулами

$$\mathcal{H}_y = B_e \cos(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.56)$$

$$\mathcal{E}_z = (j\kappa/\bar{n}_2^2 \omega \epsilon_0) B_e \sin(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.57)$$

где κ определяется соотношениями (2.4.37) и (2.2.34) при $\epsilon_2 = \bar{n}_2^2 \epsilon_0$, причем формула (2.4.57) следует из уравнения (2.2.11) и формулы (2.4.56). Снаружи активного слоя $|x| > d/2$ компоненты поля определяются выражениями

$$\mathcal{H}_y = B_e \cos(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (2.4.58)$$

$$\mathcal{E}_z = (x/|x|) (j\gamma/\bar{n}_1^2 \omega \epsilon_0) B_e \cos(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (2.4.59)$$

где γ определено соотношением (2.4.43). Из граничного условия (2.4.34), требующего непрерывности \mathcal{E}_z на границе раздела

диэлектриков $x = \pm d/2$, вытекает дисперсионное уравнение

$$\operatorname{tg}(\kappa d/2) = (\bar{n}_2^2 \gamma / \bar{n}_1^2 \kappa). \quad (2.4.60)$$

Амплитудный коэффициент B_e находится из выражения для вектора Пойнтинга:

$$B_e = \left\{ 2\omega \varepsilon_0 \bar{n}_2^2 P / \beta \left[\frac{d}{2} + \frac{(\bar{n}_1 \bar{n}_2)^2}{\gamma} \frac{\kappa^2 + \gamma^2}{\bar{n}_1^4 \kappa^2 + \bar{n}_2^4 \gamma^2} \right] \right\}^{1/2}. \quad (2.4.61)$$

Решения для нечетных ТМ-мод

Нечетные ТМ-моды при $|x| < d/2$ определяются выражениями

$$\mathcal{H}_y = B_o \sin(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.62)$$

$$\mathcal{E}_z = (-j\kappa / \bar{n}_2^2 \omega \varepsilon_0) B_o \cos(\kappa x) \exp[j(\omega t - \beta z)], \quad (2.4.63)$$

а при $|x| > d/2$ — выражениями

$$\mathcal{H}_y = (x/|x|) B_o \sin(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (2.4.64)$$

$$\mathcal{E}_z = (j\gamma / \bar{n}_1^2 \omega \varepsilon_0) B_o \sin(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)]. \quad (2.4.65)$$

Дисперсионное уравнение вытекает из условия непрерывности \mathcal{E}_z при $x = \pm d/2$ и имеет вид

$$\operatorname{tg}(\kappa d/2) = -\bar{n}_1^2 \kappa / \bar{n}_2^2 \gamma, \quad (2.4.66)$$

а коэффициент B_o определяется той же формулой (2.4.61), что и B_e . Как и ожидалось, выражения для ТМ- и ТЕ-мод имеют одинаковую структуру.

§ 5. РЕШЕНИЕ ДИСПЕРСИОННОГО УРАВНЕНИЯ В СЛУЧАЕ СИММЕТРИЧНОГО ТРЕХСЛОЙНОГО ПЛОСКОГО ВОЛНОВОДА

Показатели преломления

Для того чтобы решить дисперсионные уравнения, необходимо определить показатели преломления слоев \bar{N}_1 и \bar{N}_2 или их разность $\bar{N}_2 - \bar{N}_1$. Как будет показано ниже, в гетероструктурах вместо комплексного показателя преломления \bar{N} можно брать его действительную часть \bar{n} , поскольку изменение его мнимой части \bar{k} в гетероструктурах значительно меньше, чем характерное изменение \bar{n} . В этой главе мы будем рассматривать только системы GaAs — Al_xGa_{1-x}As. В табл. 5.2.3 приведены значения показателя преломления в других бинарных соединениях A^{III}B^V. Хотя \bar{n} зависит от длины волны, в материа-

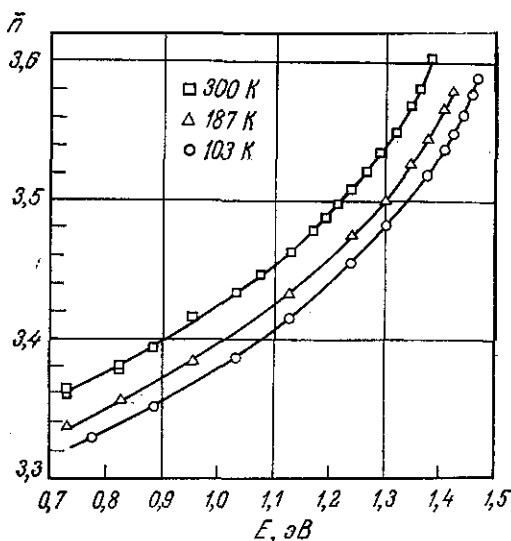


Рис. 2.5.1. Показатель преломления GaAs при указанных на рисунке температурах [15].

лах, для которых отсутствуют данные об этой зависимости, приближенное значение \bar{n} может быть получено из отношения ϵ/ϵ_0 , если воспользоваться соотношением (2.2.34).

Показатель преломления GaAs зависит от температуры, энергии фотона и концентрации примеси. Марпл [15] измерил значения \bar{n} при температурах 103, 187 и 300 К и при энергиях фотонов, меньших ширины запрещенной зоны. Образец имел концентрацию свободных электронов $\sim 6 \cdot 10^{16} \text{ см}^{-3}$. На рис. 2.5.1 представлены его результаты. Они показывают, что \bar{n} изменяется с температурой приблизительно так же, как это предсказывает формула (2.3.1). На рис. 2.5.2, а и б представлены значения показателя преломления [14], полученные при 297 К для сильно легированного GaAs *n*- и *p*-типов соответственно. Из результатов работы [14] может быть определено при фиксированной энергии фотона изменение \bar{n} при изменении равновесных концентраций электронов и дырок. На рис. 2.5.3 показана зависимость показателя преломления для фотонов с энергией 1,38 эВ от n_0 и p_0 . Зависимость \bar{n} от n_0 и p_0 влияет на волноводное распространение излучения в ОГС-лазерах. Спектральная зависимость показателя преломления $\text{Al}_x\text{Ga}_{1-x}\text{As}$ для $0 \leq x \leq 0,38$ представлена на рис. 2.5.4 [36]. Слои $\text{Al}_x\text{Ga}_{1-x}\text{As}$ в гетеролазерах обычно имеют концентрации примесей, достаточно низкие для того, чтобы можно было пренебречь зависимостью

от них показателя преломления. На рис. 2.5.5 показана зависимость \bar{n} от x при энергии фотона 1,38 эВ, являющейся характерной энергией квантов лазерного излучения ДГС-лазеров с активным слоем, легированным Si. Вставка на рис. 2.5.5 показывает, что \bar{n} слегка отклоняется вниз от линейного изменения (штриховая линия) в интервале от $\bar{n} = 3,590$ для GaAs [14, 36] до $\bar{n} = 2,971$ для AlAs [37]. Экспериментальные данные могут быть описаны формулой

$$\bar{n}(x) = 3,590 - 0,710x + 0,091x^2, \quad (2.5.1)$$

по которой можно оценить значения \bar{n} для $x > 0,38$. В табл. 2.5.1 представлены значения показателя преломления $\text{Al}_x\text{Ga}_{1-x}\text{As}$ при комнатной температуре и энергии фотона 1,38 эВ для x ,

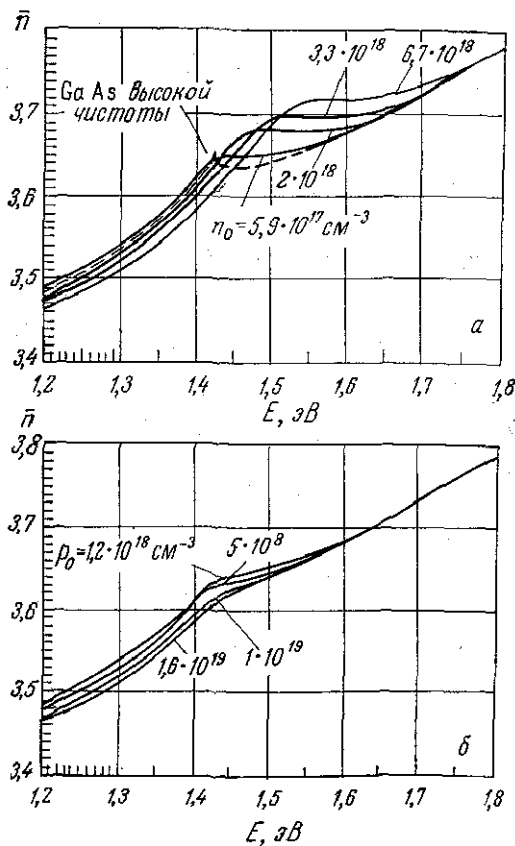


Рис. 2.5.2. Показатель преломления GaAs при температуре 297 К и указанных на рисунках равновесных концентрациях носителей. а — n-тип; б — p-тип [14].

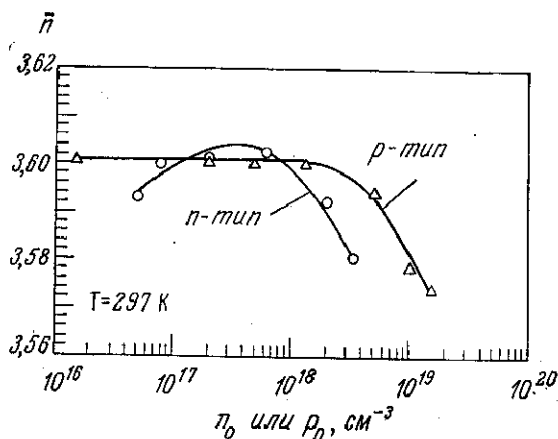


Рис. 2.5.3. Зависимость показателя преломления GaAs для фотонов с энергией 1,38 эВ от равновесных концентраций электронов и дырок. Экспериментальная ошибка определения \bar{n} составляла $\pm 0,005$ [14].

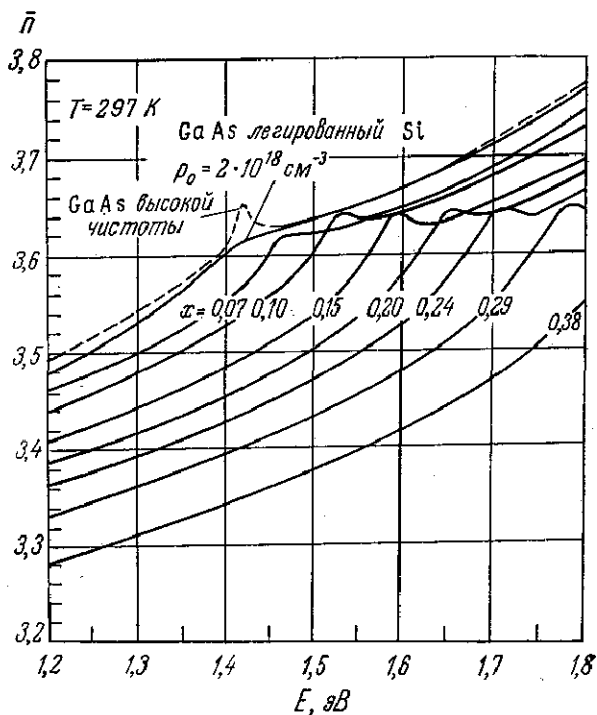


Рис. 2.5.4. Спектральная зависимость показателя преломления $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Для сравнения приведены данные для GaAs высокой чистоты [14] и GaAs, легированного Si [36].

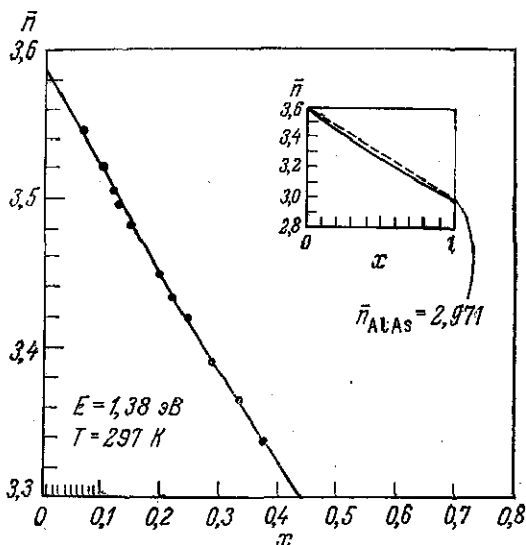


Рис. 2.5.5. Зависимость показателя преломления $\text{Al}_x\text{Ga}_{1-x}\text{As}$ от x при длине волны излучения гетеролазера, соответствующей энергии фотона 1,38 эВ. Штриховая прямая на вставке соответствует линейной интерполяции по значениям показателя преломления 3,590 для GaAs и 2,971 для AlAs [37]. Неопределенность в экспериментальных значениях показателя преломления составляет $\pm 0,005$, в то время как при используемых методах определения состава неопределенность в значениях x обычно считают равной $\pm 0,02$ [36].

изменяющегося с шагом 0,05 от 0 до 0,70. Эти значения \bar{n} будут использованы при численных расчетах ниже в этой главе.

Другие твердые растворы соединений $\text{A}^{\text{III}}\text{B}^{\text{V}}$ изучены значительно хуже, чем $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Так, для $\text{GaP}_x\text{As}_{1-x}$ имеются лишь ограниченные данные о показателе преломления [38, 39]. Афромовиц [40] предложил полуэмпирический метод вычисления показателя преломления для тройных твердых растворов соединений $\text{A}^{\text{III}}\text{B}^{\text{V}}$ и применил его к системам $\text{Al}_x\text{Ga}_{1-x}\text{As}$, $\text{GaP}_x\text{As}_{1-x}$, $\text{Ga}_x\text{In}_{1-x}\text{P}$.

Мнимая часть показателя преломления определяется коэффициентом экстинкции \bar{k} и связана с коэффициентом поглощения соотношением (2.2.61). На рис. 2.5.6 и 2.5.7 приведены значения коэффициента поглощения при 297 К для сильно легированного GaAs n - и p -типов соответственно [41]. В области

Таблица 2.5.1

Показатель преломления $\text{Al}_x\text{Ga}_{1-x}\text{As}$ при 297 К, $E = 1,38 \text{ эВ}$ [36]

| x | 0 | 0,05 | 0,10 | 0,15 | 0,20 | 0,25 | 0,30 | 0,35 | 0,40 | 0,45 | 0,50 | 0,55 | 0,60 | 0,65 | 0,70 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| \bar{n} | 3,590 | 3,555 | 3,520 | 3,486 | 3,452 | 3,418 | 3,385 | 3,353 | 3,321 | 3,289 | 3,258 | 3,227 | 3,197 | 3,167 | 3,138 |

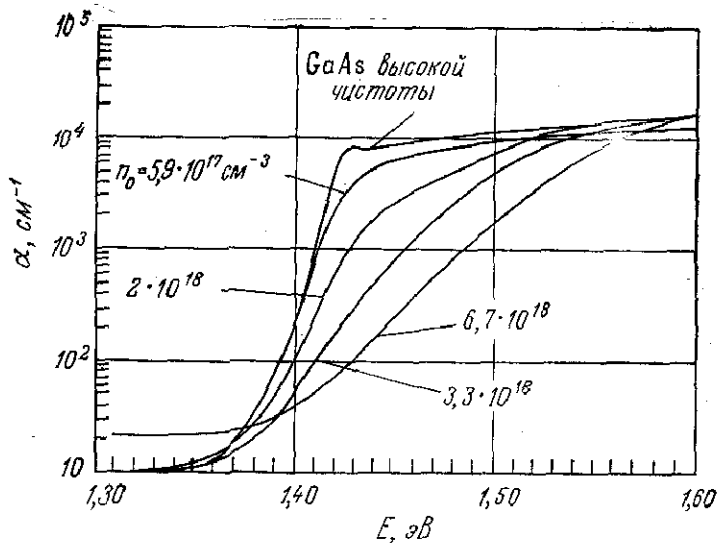


Рис. 2.5.6. Коэффициент поглощения сильно легированного n-GaAs при температуре 297 К [41].

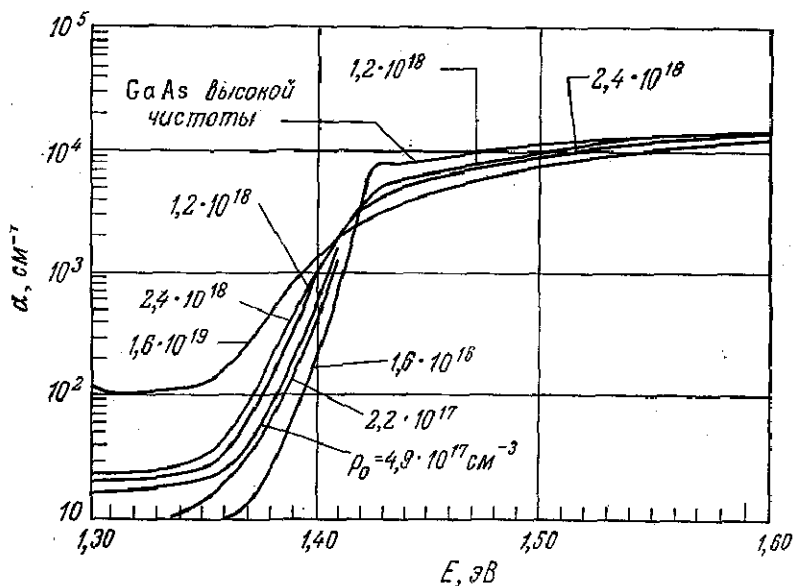


Рис. 2.5.7. Коэффициент поглощения сильно легированного p-GaAs при температуре 297 К. Для сравнения в области $\alpha < 1 \cdot 10^3 \text{ cm}^{-1}$ приведены данные для образцов с более низкой концентрацией дырок [41].

низких энергий или при уровне накачки, достаточно высоком, чтобы вызвать вынужденное излучение, поглощение происходит на свободных носителях. Влияние поглощения на свободных носителях на порог генерации лазера обсуждается в § 8, гл. 3. В GaAs *n*-типа поглощение на свободных носителях, равное 20 см^{-1} при $n = 6,7 \cdot 10^{18} \text{ см}^{-3}$, дает \bar{k} , примерно равное $1,4 \cdot 10^{-4}$. Сравнение рис. 2.5.6 и 2.5.7 показывает, что поглощение на свободных носителях больше в GaAs *p*-типа, но даже в этом случае \bar{k} все же очень мало. Из этих рисунков следует также, что коэффициент поглощения в указанном диапазоне энергий фотонов имеет максимальное значение $\sim 1 \cdot 10^4 \text{ см}^{-1}$, что дает максимальное значение $\bar{k} \approx 0,072$. Таким образом, в гетероструктурах, в которых на границах волновода происходит значительное изменение \bar{n} , мнимой частью \bar{N} можно пренебречь, за исключением области, где $\alpha \geq 10^3 \text{ см}^{-1}$. Однако когда изменение \bar{n} на границах мало, что имело место в ситуациях, рассмотренных в § 3 настоящей главы, важную роль может сыграть учет коэффициента экстинкции \bar{k} , что обусловлено усилением в активной области. Численные значения \bar{n} , приведенные на рис. 2.5.2—2.5.5 и в табл. 2.5.1, будут использованы при решении дисперсионных уравнений для типичных гетеролазеров.

Графическое решение

Графическое решение полезно использовать для иллюстрации таких понятий, как число мод и условия отсечки. Приведенное здесь решение было предложено Коллином [42]. Мы рассмотрим только ТЕ-моды: решения, соответствующие ТМ-модам, могут быть получены аналогичным путем.

Для четных ТЕ-мод дисперсионное уравнение (2.4.45) может быть записано в виде

$$(\chi d/2) \operatorname{tg}(\chi d/2) = \gamma d/2. \quad (2.5.2)$$

Складывая соотношения (2.4.37) и (2.4.43), исключим из них β^2 :

$$(\bar{n}_2^2 - \bar{n}_1^2)(k_0 d/2)^2 = (\chi d/2)^2 + (\gamma d/2)^2. \quad (2.5.3)$$

Это уравнение представляет собой уравнение окружности в плоскости $(\chi d/2, \gamma d/2)$ с радиусом $(\bar{n}_2^2 - \bar{n}_1^2)^{1/2} (k_0 d/2)$. Пересечения кривых, соответствующих уравнениям (2.5.2) и (2.5.3), в плоскости $(\chi d/2, \gamma d/2)$ определяют искомые решения. Для данной длины волны должны быть известны толщина d активного слоя и показатели преломления \bar{n}_2 и \bar{n}_1 (или их разность).

Строится зависимость $\gamma d/2 = f(\chi d/2)$, определяемая уравнением (2.5.2). На рис. 2.5.8 ей соответствуют штриховые ли-

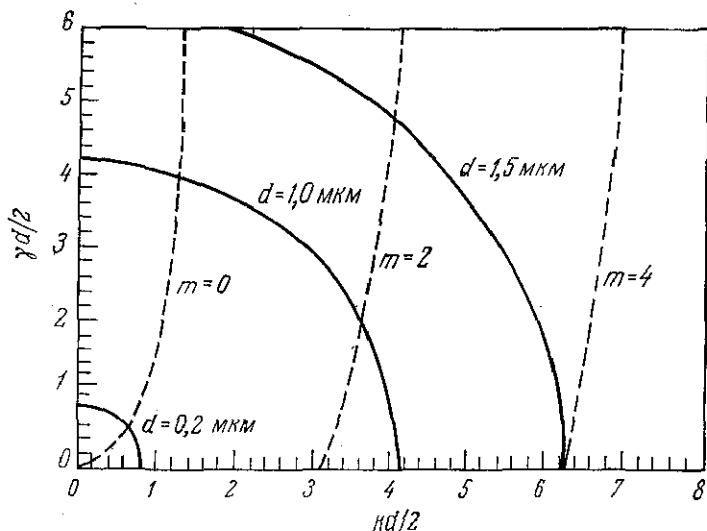


Рис. 2.5.8. Графическое решение дисперсионного уравнения (2.4.45) для четных TE-мод.

нии. Для того чтобы получить решения, рассмотрим ДГС-лазер на основе GaAs—Al_{0,3}Ga_{0,7}As с $\lambda_0 = 0,9000$ мкм (1,38 эВ). Для этой структуры из табл. 2.5.1 находим $\bar{n}_2 = 3,590$ и $\bar{n}_1 = 3,385$. На рис. 2.5.8 для $d = 0,2$ 1,0 и 1,5 мкм построены окружности (сплошные линии) с радиусами, определяемыми левой частью уравнения (2.5.3). Для $d = 0,2$ мкм окружность пересекается только с одной штриховой кривой, соответствующей $m = 0$. Координаты точки пересечения определяют решение для моды низшего порядка (основной моды). Каким бы малым ни было d , в симметричных структурах для этой моды не существует условия отсечки. Полученные таким путем значения γ и κ позволяют вычислить распределение полей. Увеличение d до 1,0 мкм приводит к появлению пересечений вблизи $\kappa d/2 \approx 3,6$ для моды, соответствующей $m = 2$, и вблизи $\kappa d/2 \approx 1,3$ для основной моды. Рассматриваются только положительные значения $\gamma d/2$, поскольку отрицательные значения этой величины дают решения, соответствующие волнам, затухающим в направлении распространения. Толщина активной области $d = 1,5$ мкм, при которой кривые, определяемые уравнениями (2.5.2) и (2.5.3), пересекаются при $\gamma d/2 = 0$, соответствует отсечке моды четвертого порядка ($m = 4$). При этом для моды второго порядка и основной моды точки пересечения дают значения $\kappa d/2$, равные 4,0 и 1,5 соответственно. Значения $\kappa d/2$, при которых происходит отсечка мод высшего порядка, определяются точками пересечения кривых $\gamma d/2 = (\kappa d/2) \text{tg}(\kappa d/2)$ с осью $\kappa d/2$.

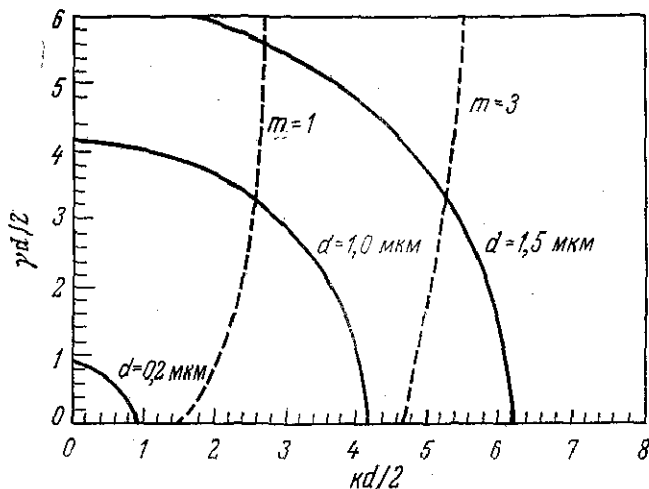


Рис. 2.5.9. Графическое решение дисперсионного уравнения (2.4.54) для нечетных ТЕ-мод.

Для четных m этим точкам соответствуют $kd/2 = m\pi/2$. Значение d при отсечке, как следует из уравнения (2.5.3), определяется выражением

$$d = m\lambda_0/2 (\bar{n}_2^2 - \bar{n}_1^2)^{1/2}, \quad (2.5.4)$$

где $m = 2, 4, \dots$

Для нечетных ТЕ-мод дисперсионное уравнение (2.4.54) может быть представлено в виде

$$-(kd/2) \operatorname{ctg}(kd/2) = \gamma d/2. \quad (2.5.5)$$

Уравнение (2.5.3) справедливо также и для нечетных мод. На рис. 2.5.9 построены кривые, соответствующие уравнениям (2.5.3) и (2.5.5), причем значения параметров брались те же, что и при построении кривых на рис. 2.5.8. Искомые решения определялись так же, как и в случае четных мод, а именно по пересечениям кривых, соответствующих уравнению (2.5.5), с окружностями радиусов $(\bar{n}_2^2 - \bar{n}_1^2)^{1/2} (k_0 d/2)$. Для нечетных ТЕ-мод отсечка также происходит при $kd/2 = m\pi/2$, и толщина активной области при этом определяется уравнением (2.5.4), но здесь уже $m = 1, 3, 5, \dots$. Таким образом, уравнение (2.5.4) с $m = 1, 2, 3, 4, 5, \dots$ дает условие отсечки как для четных, так и для нечетных мод и при заданном d определяет разрешенные моды высшего порядка. Уравнения (2.5.2), (2.5.3) и (2.5.5) легко могут быть решены с помощью микрокалькулятора методом проб и ошибок, если только рассматривается небольшое

Таблица 2.5.2

Значения величины β/k_0 для GaAs — Al_xGa_{1-x}As симметричного плоского трехслойного волновода с $\bar{n}_2 = 3,590$ и $\lambda_0 = 0,9000$ мкм

| d , мкм | $x=0,1, \bar{n}_1=3,520$ | $x=0,2, \bar{n}_2=3,448$ | $x=0,3, \bar{n}_1=3,385$ |
|-----------------------|--------------------------|--------------------------|--------------------------|
| β/k_0 для $m=0$ | | | |
| 0,1 | 3,524 | 3,463 | 3,415 |
| 0,15 | 3,528 | 3,477 | 3,440 |
| 0,2 | 3,533 | 3,492 | 3,463 |
| 0,3 | 3,543 | 3,516 | 3,499 |
| 0,4 | 3,551 | 3,534 | 3,524 |
| 0,5 | 3,558 | 3,546 | 3,539 |
| 0,75 | 3,570 | 3,564 | 3,561 |
| 1,0 | 3,576 | 3,573 | 3,572 |
| β/k_0 для $m=1$ | | | |
| 0,3 | — | — | — |
| 0,4 | — | — | 3,387 |
| 0,5 | — | 3,451 | 3,413 |
| 0,75 | 3,524 | 3,493 | 3,480 |
| 1,0 | 3,539 | 3,524 | 3,518 |
| β/k_0 для $m=2$ | | | |
| 0,5 | — | — | — |
| 0,75 | — | — | — |
| 1,0 | — | 3,456 | 3,434 |

число мод. Довольно грубые графики, подобные приведенным на рис. 2.5.8 и 2.5.9, дают приближенные решения. Например, приближенное значение $kd/2$ можно использовать для вычисления величины $\gamma d/2$ по уравнению (2.5.2) или (2.5.3). Значения $kd/2$ можно варьировать до тех пор, пока оба способа вычисления не дадут одно и то же значение $\gamma d/2$. Для значений λ_0 , \bar{n}_1 и \bar{n}_2 , использованных при построении кривых на рис. 2.5.8 и 2.5.9, имеем $k_0 = 69813$ см⁻¹ и для $d = 0,2$ мкм получаем $kd/2 = 0,6598$ и $\gamma d/2 = 0,5116$. Постоянная распространения β может быть найдена как из выражения (2.4.37), так и из выражения (2.4.43). Обычно она дается в виде величины β/k_0 , которую часто называют эффективным показателем преломления $\bar{n}_{эфф}$. Для приведенного выше примера $\beta/k_0 = 3,463$. Этот результат подчеркивает тот факт, что для волноводных мод

$$\bar{n}_2 > \beta/k_0 > \bar{n}_1. \quad (2.5.6)$$

В следующей части этого параграфа мы вычислим распределения интенсивности излучения с целью проиллюстрировать ряд свойств распространения волн в гетеролазерах.

Численные результаты

Свойства волноводных мод в гетеролазерах можно продемонстрировать на нескольких характерных примерах. В табл. 2.5.2 приведены значения величины β/k_0 для волновода, изображенного на рис. 2.3.2, при $x = 0,1, 0,2$ и $0,3$ и толщинах активной

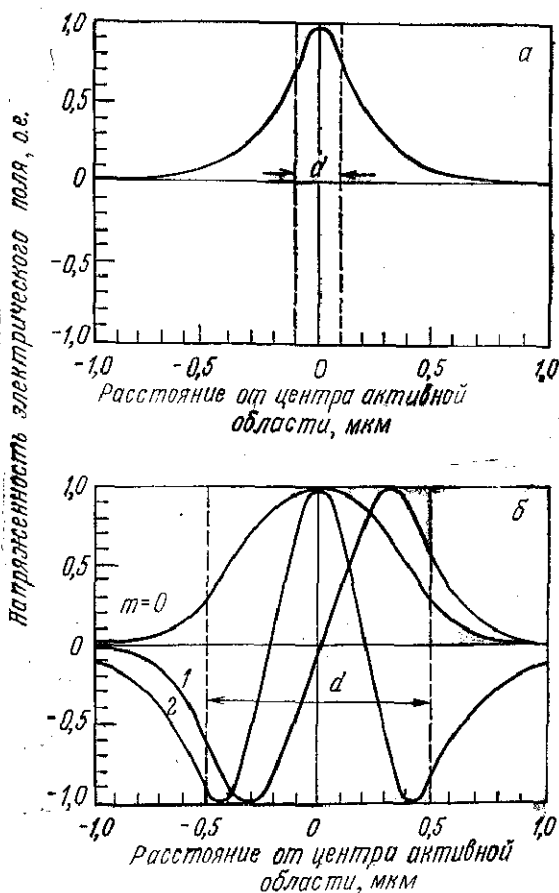


Рис. 2.5.10. Распределение электрического поля в симметричном трехслойном плоском волноводе при $t = 0$ и $z = 0$, определяемое выражениями (2.4.36) и (2.4.41). Значения параметров: $\lambda_0 = 0,9000$ мкм, $\bar{n}_1 = 3,590$, $\bar{n}_2 = 3,385$. а — основная ТЕ-мода, $d = 0,2$ мкм; б — ТЕ-моды: основная, первого и второго порядков, $d = 1,0$ мкм.

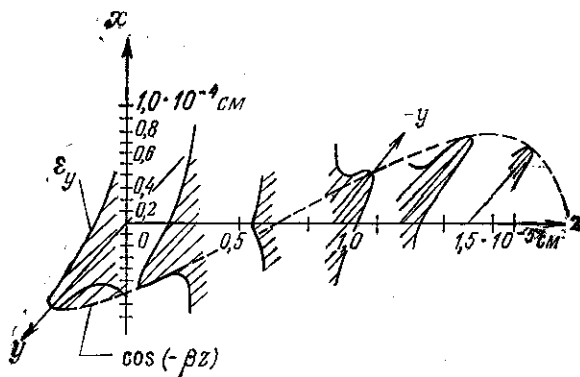


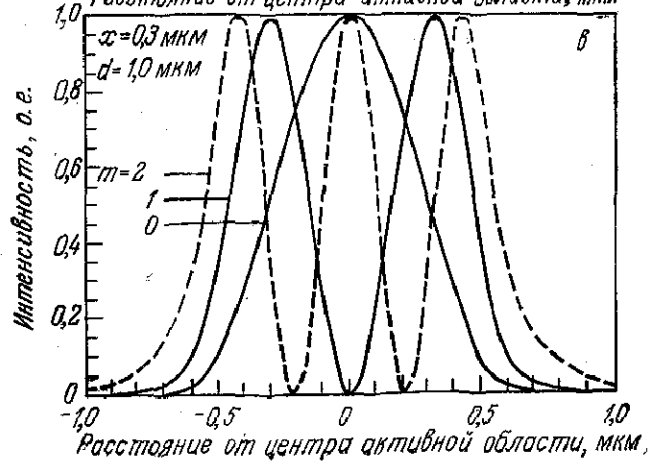
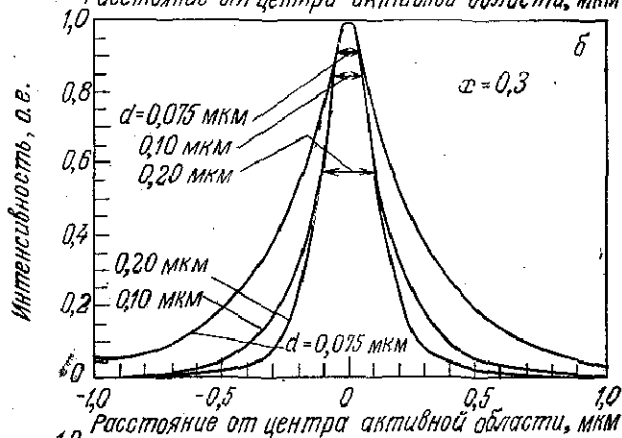
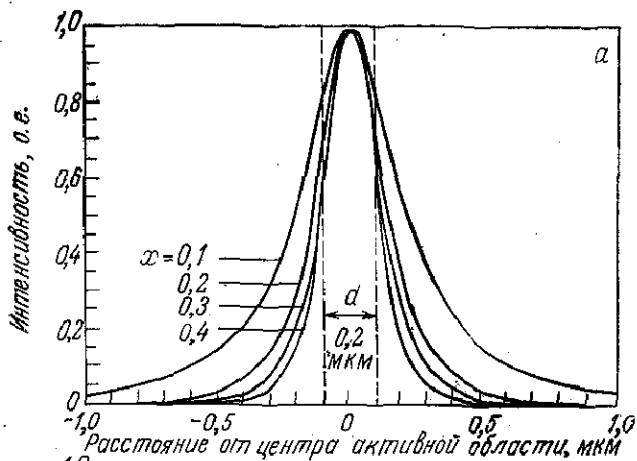
Рис. 2.5.11. Пространственное распределение в момент времени $t = 0$ электрического поля основной ТЕ-моды в симметричном трехслойном плоском волноводе при $d = 0,2$ мкм, $\lambda_0 = 0,9000$ мкм, $\bar{n}_1 = 3,590$, $\bar{n}_2 = 3,385$.

области от 0,1 до 1,0 мкм. Значения показателя преломления из табл. 2.5.1 использовались при численном решении дисперсионных уравнений (2.5.2) и (2.5.5). Данные табл. 2.5.2 показывают, что при уменьшении d величина β/k_0 приближается к \bar{n}_1 . Для мод высшего порядка вблизи отсечки $\beta/k_0 \approx \bar{n}_1$.

Конфигурации электрического поля определяются выражениями (2.4.36) и (2.4.41). На рис. 2.5.10, а приведено распределение электрического поля \mathcal{E}_y при $z = 0$ и $t = 0$ для ДГС с $d = 0,2$ мкм и $x = 0,3$. Многомодовый случай изображен на рис. 2.5.10, б для $d = 1,0$ мкм и $x = 0,3$. В центре волновода четные моды ($m = 0, 2, 4, \dots$) всегда имеют максимум, а нечетные моды ($m = 1, 3, 5, \dots$) равны нулю. На рис. 2.5.11 показано распределение электрического поля в волноводе в момент времени $t = 0$ для случая $d = 0,2$ мкм и $x = 0,3$. Амплитуда \mathcal{E}_y промодулирована множителем $\exp[j(\omega t - \beta z)]$. Эта волна аналогична плоской волне, картина распространения которой представлена на рис. 2.2.2. Однако в волноводном случае \mathcal{E}_y спадает до нуля в направлении, перпендикулярном направлению распространения, а расстояние между двумя последовательными максимумами \mathcal{E}_y не равно длине волны. Распространение мод в плоском волноводе можно пояснить, исходя из простых физических законов, если воспользоваться моделью зигзагообразных волн, которая рассматривается в § 6 этой главы.

Коэффициент оптического ограничения

Как видно из рис. 2.5.10, а, амплитуда \mathcal{E}_y может иметь значительную величину снаружи активного слоя трехслойного плоского диэлектрического волновода, изображенного на рис. 2.3.2.



Поэтому часть поля не взаимодействует с инжектированными носителями и, таким образом, не дает вклада в вынужденное излучение в ДГС-лазерах. Поскольку вынужденное излучение и усиление связаны с потоком фотонов, важно рассмотреть распределение поля с целью определения той части переносимой модой энергии, которая заключена в активном слое лазера. Оптический поток определяется через вектор Пойнтинга соотношением (2.2.57). Из уравнений (2.4.46) — (2.4.48) следует, что $\mathcal{E} \times \mathcal{H}^*$ пропорционально $|\mathcal{E}_y|^2$. Поэтому интенсивность излучения $I = |\mathbf{S}|$, пропорциональная $|\mathcal{E}_y|^2$, является более полезной физической величиной при исследовании лазеров, чем \mathcal{E}_y . Степень ограничения излучения активным слоем может быть представлена коэффициентом оптического ограничения Γ . Поскольку он определяется как отношение интенсивности света, приходящейся на активный слой, к суммарной интенсивности света, приходящейся на все слои, то для симметричного трехслойного плоского диэлектрического волновода выражение для Γ может быть получено из соотношений (2.4.36) и (2.4.41) в виде

$$\Gamma = \frac{\int_0^{d/2} \cos^2(\kappa x) dx}{\int_0^{d/2} \cos^2(\kappa x) dx + \int_{d/2}^{\infty} \cos^2(\kappa d/2) \exp[-2\gamma(x - d/2)] dx}, \quad (2.5.7)$$

что дает

$$\Gamma = \left(1 + \frac{\cos^2(\kappa d/2)}{\gamma [d/2 + (1/\kappa) \sin(\kappa d/2) \cos(\kappa d/2)]} \right)^{-1} \quad (2.5.8)$$

для четных ТЕ-мод. Аналогичные выражения получаются для нечетных ТЕ-мод, а также для ТМ-мод. Частое использование коэффициента оптического ограничения обусловлено тем, что он представляет собой часть переносимой модой энергии, распространяющуюся внутри активного слоя. Величины γ и κ примерно одинаковы для ТЕ- и ТМ-мод, поэтому различия в Γ для этих поляризаций малы. Вследствие этого малого различия,

Рис. 2.5.12. Зависимость квадрата модуля электрического вектора в ДГС-волноводе, изображенном на рис. 2.3.2, от расстояния, измеренного от центра активного слоя в направлении, перпендикулярном плоскости перехода. Длина волны излучения принята равной $\lambda_0 = 0.9000$ мкм (1,38 эВ). a — распределения интенсивности излучения при $d = 0,2$ мкм и молярных долях AlAs, указанных на рисунке; b — распределения интенсивности излучения при $x = 0,8$ и толщинах активного слоя, указанных на рисунке; c — распределение интенсивности для основной моды и мод первого и второго порядков при $d = 1,0$ мкм и $x = 0,3$.

а также из-за того, что в ДГС-лазерах излучение имеет преимущественно ТЕ-поляризацию, коэффициент Γ для ТМ-волн рассматриваться не будет.

Рис. 2.5.12, *а* показывает влияние молярной доли арсенида алюминия x на распределение интенсивности света $|\mathcal{E}_y|^2$ в направлении, перпендикулярном плоскости $p-n$ -перехода, для трехслойного плоского диэлектрического волновода, изображенного на рис. 2.3.2. Длина волны равна 0,90 мкм (1,38 эВ). Толщина активного слоя, равная 0,2 мкм, постоянна, а меняется состав пассивных областей. На рис. 2.5.12, *а* видно, что значительное ослабление ограничения происходит при уменьшении x от 0,2 до 0,1. На рис. 2.5.12, *б* представлена зависимость ограничения от d для $x = 0,3$. При уменьшении толщины активного слоя оптическое поле все больше распространяется в слое $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$, и все меньшая доля общей интенсивности излучения приходится на активный слой. На рис. 2.5.12, *в* показано, что в случае больших d , когда становятся разрешенными моды высшего порядка, при увеличении порядка моды уменьшается доля интенсивности света, приходящаяся на активный слой. Поэтому чем ниже порядок моды, тем больше коэффициент оптического ограничения.

На рис. 2.5.13, *а* и *б* представлена зависимость Γ для основной моды от молярной доли арсенида алюминия x и толщины активного слоя d . Показано быстрое уменьшение Γ при $d < < \lambda_0/\bar{n}_2$, т. е. когда толщина активного слоя становится меньше, чем длина волны излучения в полупроводнике (рис. 2.5.13, *а*). Участок, соответствующий толщинам активного слоя от 0,01 до 0,1 мкм, представлен на рис. 2.5.13, *б* в увеличенном масштабе. На рис. 2.5.14 приведена зависимость коэффициента оптического ограничения от толщины активного слоя для основной моды и мод высшего порядка при $x = 0,3$. Для мод высшего порядка Γ мало вблизи отсечки и при увеличении d приближается к значению коэффициента оптического ограничения для основной моды. В гетеролазере с толщиной активной области d достаточно большой, чтобы существовали моды высшего порядка, коэффициент оптического ограничения и коэффициент отражения на сколотых гранях определяют доминирующую моду. Коэффициент отражения на сколотых гранях обсуждается в § 8 этой главы. Зависимость порогового тока лазера от Γ рассматривается в § 8 гл. 3, а примеры использования этой зависимости даны в гл. 7. Представление части переносимой модой энергии излучения, распространяющейся в активном слое, через коэффициент оптического ограничения позволяет понять, какое влияние оказывает толщина активного слоя на коэффициент усиления, необходимый для достижения порога генерации. Коэффициент Γ важен также и для структур, не образующих симмет-

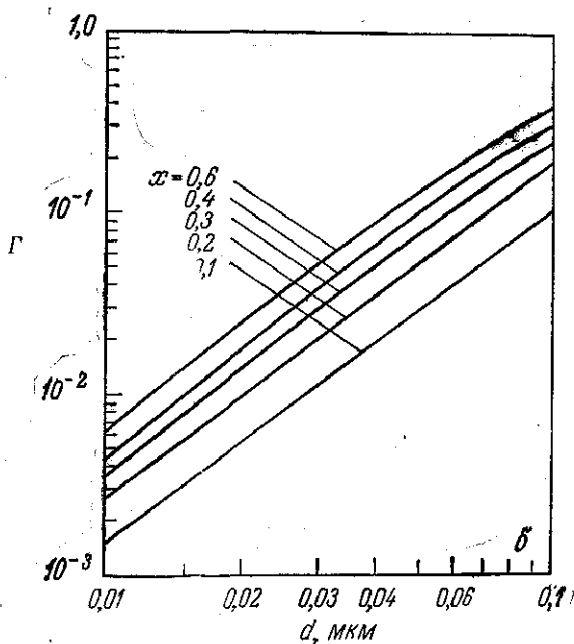
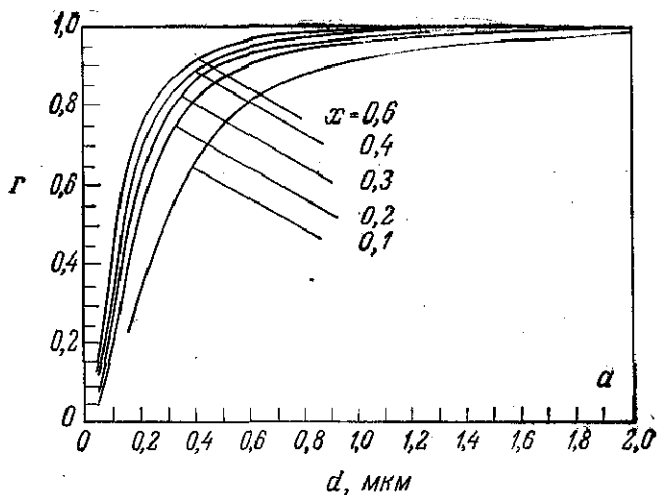


Рис. 2.5.13. Зависимость коэффициента оптического ограничения для основной моды от толщины активного слоя в GaAs—Al_xGa_{1-x}As симметричном трехслойном плоском диэлектрическом волноводе. а — зависимость Γ от d в области изменения d от 0,05 до 2,0 мкм; б — зависимость Γ от d в области изменения d от 0,01 до 0,1 мкм.

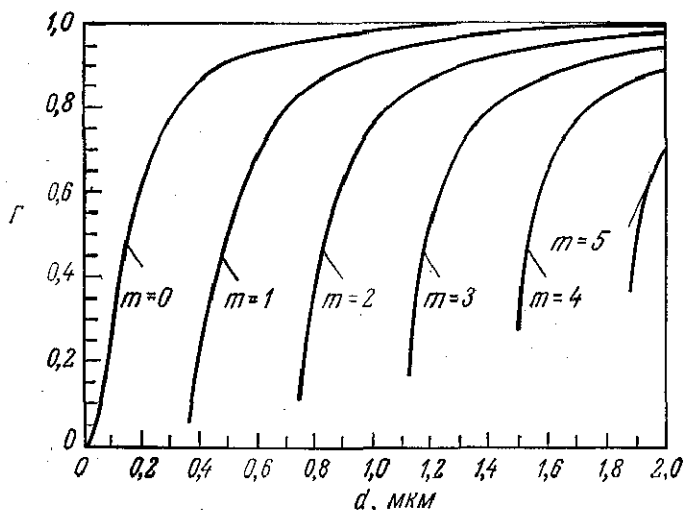


Рис. 2.5.14. Зависимость коэффициента оптического ограничения для основной моды и мод высшего порядка от толщины активного слоя в GaAs— $\text{Al}_{0,8}\text{Ga}_{0,2}\text{As}$ симметричном трехслойном плоском диэлектрическом волноводе.

ричный трехслойный плоский волновод. В этой главе приводятся результаты расчетов для асимметричных трехслойных плоских диэлектрических волноводов (§ 6) и для некоторых четырех- и пятислойных плоских волноводов (§ 9).

Приближенные решения

Думке [43] предложил приближенное решение дисперсионного уравнения для малых и больших d . При d , стремящемся к нулю, величина $\text{tg}(\kappa d/2)$ стремится к $\kappa d/2$, так что уравнение (2.4.45) может быть переписано в виде

$$\gamma \approx \kappa^2 d/2. \quad (2.5.9)$$

Из табл. 2.5.2 видно, что при уменьшении d значения β/k_0 приближаются к \bar{n}_1 . Тогда, подставляя γ^2 из выражения (2.4.43) в уравнение (2.5.3), получаем

$$\kappa \approx (\bar{n}_2^2 - \bar{n}_1^2)^{1/2} k_0 \quad (2.5.10)$$

и из уравнения (2.5.9) имеем

$$\gamma \approx (\bar{n}_2^2 - \bar{n}_1^2)^{1/2} k_0^2 d/2. \quad (2.5.11)$$

Это приближение основано на предположении, что

$$\beta/k_0 = \bar{n}_1. \quad (2.5.12)$$

Как видно из рис. 2.5.8, решение дисперсионного уравнения, соответствующее $m = 0$, при больших d дает $\kappa d/2 \approx \pi/2$, или

$$\kappa \approx \pi/d. \quad (2.5.13)$$

Тогда β/k_0 и γ могут быть получены из выражений (2.4.37) и (2.4.43) соответственно.

При малых d множитель $\cos(\kappa x)$ в выражении (2.4.36) для \mathcal{E}_y в области $|x| < d/2$ примерно равен единице. Тогда во всей области $|x| > 0$ для $\mathcal{E}_y(x)$ может быть использована экспоненциальная зависимость, выраженная соотношением (2.4.41) [43]:

$$\mathcal{E}_y(x) = E_0 \exp(-\gamma|x|), \quad (2.5.14)$$

где E_0 — амплитуда. Из этого экспоненциального приближения при малых d получается следующее выражение для коэффициента оптического ограничения:

$$\Gamma \approx \frac{\int_0^{d/2} E_0^2 \exp(-2\gamma x) dx}{\int_0^\infty E_0^2 \exp(-2\gamma x) dx}, \quad (2.5.15)$$

которое при малых γd с учетом выражения (2.5.11) для γ дает

$$\Gamma \approx \gamma d = 2\pi^2 (\bar{n}_2^2 - \bar{n}_1^2) d^2 / \lambda_0^2. \quad (2.5.16)$$

Для системы GaAs—Al_xGa_{1-x}As при $\lambda_0 = 0,90$ мкм из соотношения (2.5.1) имеем

$$\bar{n}_2^2 - \bar{n}_1^2 = 1,42\bar{n}_2 x, \quad (2.5.17)$$

где опущены члены более высокого порядка по x . Из выражений (2.5.16) и (2.5.17) при $\bar{n}_2 = 3,590$ получаем

$$\Gamma = 10^2 x d^2 / \lambda_0^2. \quad (2.5.18)$$

Как видно из рис. 2.5.13, б, при малых d коэффициент оптического ограничения Γ изменяется приблизительно как d^2 . Для $d = 0,1$ мкм, $x = 0,1$ и $\lambda_0 = 0,90$ мкм соотношение (2.5.18) дает $\Gamma = 0,12$, а из рис. 2.5.13, б получаем $\Gamma = 0,11$. Полученное при малых d приближенное соотношение (2.5.16) будет использовано нами в § 7 этой главы при выводе выражения, аппроксимирующего угол расходимости излучения в направлении, перпендикулярном плоскости $p-n$ -перехода.

Область применимости полученного приближения определяется условием $\text{tg}(\kappa d/2) \approx \kappa d/2$. Из выражений (2.5.10) и (2.5.17) следует

$$\kappa d/2 = \pi (1,42\bar{n}_2 x)^{1/2} d / \lambda_0. \quad (2.5.19)$$

Условие $\text{tg}(kd/2) \approx kd/2$ означает, что $kd/2$ должно быть меньше или равно 0,5 рад. При этом, полагая $\bar{n}_2 = 3,590$, из соотношения (2.5.19) имеем

$$d \leq 0,07\lambda_0/x^{1/2}. \quad (2.5.20)$$

Это соотношение показывает, что область применимости приближения, соответствующего малым d , зависит от состава слоев, окружающих активный слой.

Если данные по показателю преломления тройного твердого раствора отсутствуют, величина $\bar{n}_2^2 - \bar{n}_1^2$ может быть аппроксимирована выражением, аналогичным выражению (2.5.17). Для показателя преломления может быть взята линейная зависимость от состава:

$$\bar{n}_1(x) = \bar{n}_2 - bx. \quad (2.5.21)$$

Коэффициент b легко получить, зная $\bar{n}_1(1,0)$ — показатель преломления бинарного соединения, соответствующего $x = 1$. Показатели преломления бинарных соединений $A^{IV}B^V$ приведены в табл. 5.2.3. Из выражения (2.5.21) следует формула

$$\bar{n}_2^2 - \bar{n}_1^2 = 2b\bar{n}_2x, \quad (2.5.22)$$

в которой опущены члены порядка x^2 .

§ 6. МОДЕЛЬ ЗИГЗАГООБРАЗНЫХ ВОЛН И АСИММЕТРИЧНЫЕ ТРЕХСЛОЙНЫЕ ПЛОСКИЕ ВОЛНОВОДЫ

Модель зигзагообразных волн

Модель зигзагообразных волн [44—46] дает наиболее простое решение дисперсионного уравнения для определения распределения поля в асимметричном трехслойном плоском волноводе и помогает понять волноводный эффект. Эта модель позволяет определить величины κ , γ и β , которые используются в ранее выведенных выражениях для \mathcal{E}_y . Знание решений волнового уравнения, приведенных в § 4, 5 настоящей главы, необходимо для правильного понимания вывода и применения выражений, используемых в этой модели.

Для того чтобы вывести дисперсионное уравнение, необходимо сначала определить фазовый сдвиг при отражении плоской волны от границы раздела диэлектриков. Пространственный обзор вопросов, касающихся отражения и преломления на границе раздела диэлектриков, дан Лотшом [47]. Проникновение хвостов моды [см. выражение (2.4.41)] в прилегающий слой с меньшим показателем преломления приводит к горизонтальному сдвигу луча, который называется сдвигом Гуса — Хэнхена

[48, 49]. Этот сдвиг играет важную роль в случае тонких активных слоев, часто встречающихся в ДГС-лазерах, и будет рассмотрен в этом параграфе. Для пояснения модели зигзагообразных волн мы вычислили коэффициент оптического ограничения и распределение поля для некоторых асимметричных трехслойных плоских волноводов, параметры которых характерны для волноводов в ОГС-лазерах.

Отражение на границе раздела диэлектриков

Используя полученные выше решения, соответствующие четным ТЕ-модам, выражение (2.4.36) для \mathcal{E}_y можно переписать в виде

$$\mathcal{E}_y = (A_e/2) \{ \exp[j(\omega t + \kappa x - \beta z)] + \exp[j(\omega t - \kappa x - \beta z)] \}, \quad (2.6.1)$$

где \mathcal{E}_y представлено как суперпозиция двух плоских волн в области 2 на рис. 2.3.2. Рассмотрим второе слагаемое в выражении (2.6.1), описывающее распространение волны в положительном направлении осей x и z . На рис. 2.6.1, а показан процесс отражения этой плоской волны на границе раздела диэлектриков. Вектор электрического поля, соответствующий падающей волне, обозначен через \mathcal{E}_{yi} и направлен по нормали к плоскости рисунка. Для прошедшей и отраженной волн векторы

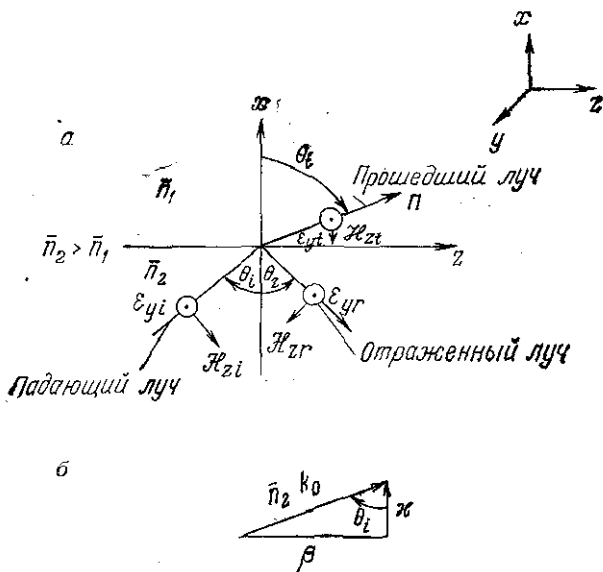


Рис. 2.6.1. Схематическое представление электрического и магнитного полей на границе раздела диэлектриков (а); падающее, отраженное и прошедшее поля обозначены индексами i , r и t соответственно. Постоянная распространения $\bar{n}_2 k_0$ и ее x - и z -составляющие (б).

электрического поля обозначены через \mathcal{E}_{yt} и \mathcal{E}_{yr} соответственно. Из формулы (2.4.37) имеем $\bar{n}_2^2 k_0^2 = \beta^2 + \kappa^2$, где x -компоненте постоянной распространения соответствует κ , а z -компоненте — β . Эта ситуация изображена на рис. 2.6.1, б. Мы можем написать

$$\beta = \bar{n}_2 k_0 \sin \theta_t, \quad (2.6.2)$$

$$\kappa = \bar{n}_2 k_0 \cos \theta_t \quad (2.6.3)$$

и, следовательно,

$$\operatorname{tg} \theta_t = \beta/\kappa. \quad (2.6.4)$$

Второе слагаемое в выражении (2.6.1) с учетом формул (2.6.2) и (2.6.3) принимает вид

$$\mathcal{E}_{yt} = \mathcal{E}_t \exp [j(\omega t - \bar{n}_2 k_0 x \cos \theta_t - \bar{n}_2 k_0 z \sin \theta_t)]. \quad (2.6.5)$$

Аналогичные формулы справедливы для электрических полей прошедшей и отраженной волн:

$$\mathcal{E}_{yt} = \mathcal{E}_t \exp [j(\omega t - \bar{n}_1 k_0 x \cos \theta_t - \bar{n}_1 k_0 z \sin \theta_t)], \quad (2.6.6)$$

$$\mathcal{E}_{yr} = \mathcal{E}_r \exp [j(\omega t + \bar{n}_2 k_0 x \cos \theta_r - \bar{n}_2 k_0 z \sin \theta_r)]. \quad (2.6.7)$$

Из граничного условия (2.4.34) следует непрерывность тангенциальных составляющих электрического поля при $x = 0$:

$$\begin{aligned} \mathcal{E}_t \exp [j(\omega t - \bar{n}_2 k_0 z \sin \theta_t)] + \mathcal{E}_r \exp [j(\omega t - \bar{n}_2 k_0 z \sin \theta_r)] = \\ = \mathcal{E}_t \exp [j(\omega t - \bar{n}_1 k_0 z \sin \theta_t)]. \end{aligned} \quad (2.6.8)$$

Это условие непрерывности справедливо при любых значениях z . Следовательно, падающая, отраженная и прошедшая волны должны иметь одну и ту же фазу при распространении в направлении оси z . Таким образом,

$$\bar{n}_2 \sin \theta_t = \bar{n}_2 \sin \theta_r = \bar{n}_1 \sin \theta_t, \quad (2.6.9)$$

что дает

$$\theta_t = \theta_r \quad (2.6.10)$$

и

$$\bar{n}_2 \sin \theta_t = \bar{n}_1 \sin \theta_t. \quad (2.6.11)$$

Уравнение (2.6.11) представляет собой закон Снеллиуса.

Магнитное поле \mathcal{H}_z связано с \mathcal{E}_y уравнением (2.4.38). Из этого уравнения и выражения (2.6.5) для электрического поля падающей волны следует уравнение

$$\partial \mathcal{H}_{zt} / \partial t = (j/\mu_0) \mathcal{E}_{yt} \bar{n}_2 k_0 \cos \theta_t, \quad (2.6.12)$$

интегрирование которого по t дает

$$\mathcal{H}_{zt} = (\bar{n}_2 k_0 / \omega \mu_0) \mathcal{E}_{yt} \cos \theta_t. \quad (2.6.13)$$

Множитель $\bar{n}_2 k_0 / \omega \mu_0$ при $\mu_0 = 1$ обычно записывают в виде \bar{n}_2 / c . Тогда

$$\mathcal{H}_{zt} = (\bar{n}_2 / c) \mathcal{E}_{yt} \cos \theta_t. \quad (2.6.14)$$

Аналогичным образом получаются выражения для отраженной и прошедшей волн:

$$\mathcal{H}_{zr} = -(\bar{n}_2 / c) \mathcal{E}_{yr} \cos \theta_r, \quad (2.6.15)$$

$$\mathcal{H}_{zt} = (\bar{n}_1 / c) \mathcal{E}_{yt} \cos \theta_t. \quad (2.6.16)$$

Из граничного условия (2.4.35) для тангенциальных составляющих магнитного поля следует непрерывность \mathcal{H}_z при $x = 0$:

$$\mathcal{H}_{zt} + \mathcal{H}_{zr} = \mathcal{H}_{zt}. \quad (2.6.17)$$

Из этого уравнения при учете выражений (2.6.14) — (2.6.16) следует

$$\bar{n}_2 \mathcal{E}_{yt} \cos \theta_t - \bar{n}_2 \mathcal{E}_{yr} \cos \theta_r = \bar{n}_1 \mathcal{E}_{yt} \cos \theta_t. \quad (2.6.18)$$

Используя выражения (2.6.8) и (2.6.10), заменим в уравнении (2.6.18) \mathcal{E}_{yt} на $\mathcal{E}_{yi} + \mathcal{E}_{yr}$ и θ_r на θ_t . Тогда получим

$$\mathcal{E}_{yr} = \frac{\bar{n}_2 \cos \theta_t - \bar{n}_1 \cos \theta_t}{\bar{n}_2 \cos \theta_t + \bar{n}_1 \cos \theta_t} \mathcal{E}_{yt}. \quad (2.6.19)$$

Учитывая тот факт, что мы вывели формулу (2.6.19) из граничных условий при $x = 0$, а также принимая во внимание соотношение (2.6.10), можно заменить в формуле (2.6.19) \mathcal{E}_{yr} и \mathcal{E}_{yt} на их амплитуды \mathcal{E}_r и \mathcal{E}_t . Эта формула, связывающая между собой амплитуды отраженной и падающей волн, называется формулой Френеля.

Прошедшая волна отсутствует при $\theta_t = 90^\circ$. Соответствующий угол падения можно определить из закона Снеллиуса (2.6.11):

$$\theta_c = \arcsin(\bar{n}_1 / \bar{n}_2). \quad (2.6.20)$$

Этот угол называется критическим углом. Волноводное распространение излучения возникает при выполнении неравенства $\theta_i \geq \theta_c$. Уравнение (2.6.11) может быть переписано в виде

$$\cos^2 \theta_t = 1 - (\bar{n}_2^2 / \bar{n}_1^2) \sin^2 \theta_i. \quad (2.6.21)$$

При $\theta_i \geq \arcsin(\bar{n}_1 / \bar{n}_2)$, когда отсутствует проходящая волна, член $(\bar{n}_2^2 / \bar{n}_1^2) \sin^2 \theta_i$ становится больше единицы и

$$\cos \theta_t = \pm j [(\bar{n}_2^2 / \bar{n}_1^2) \sin^2 \theta_i - 1]^{1/2}. \quad (2.6.22)$$

В этом выражении надо выбрать знак минус, чтобы определяемое формулой (2.6.6) \mathcal{E}_{yt} соответствовало затухающей волне.

Поэтому комплексный коэффициент отражения Γ , равный $\mathcal{E}_{yr}/\mathcal{E}_{yi}$, определяется выражением

$$\Gamma = \frac{\mathcal{E}_{yr}}{\mathcal{E}_{yi}} = \frac{\bar{n}_2 \cos \theta_i + j(\bar{n}_2^2 \sin^2 \theta_i - \bar{n}_1^2)^{1/2}}{\bar{n}_1 \cos \theta_i - j(\bar{n}_2^2 \sin^2 \theta_i - \bar{n}_1^2)^{1/2}}, \quad (2.6.23)$$

которое получается после подстановки выражения (2.6.22) в формулу (2.6.19). Выражение (2.6.23) представляет собой отношение двух комплексных чисел $(a + jb)/(a - jb)$, которое при записи этих чисел в экспоненциальной форме принимает вид

$$\Gamma = \exp(j\phi)/\exp(-j\phi) = \exp(2j\phi), \quad (2.6.24)$$

где

$$\operatorname{tg} \phi = (\bar{n}_2^2 \sin^2 \theta_i - \bar{n}_1^2)^{1/2} / \bar{n}_2 \cos \theta_i. \quad (2.6.25)$$

Из соотношений (2.6.2), (2.6.3) и (2.6.25) следует

$$\operatorname{tg} \phi = (\beta^2/k_0^2 - \bar{n}_1^2)^{1/2} / (\kappa/k_0), \quad (2.6.26)$$

и, учитывая выражение (2.4.43) для γ^2 ,

$$\operatorname{tg} \phi = \gamma/\kappa. \quad (2.6.27)$$

Таким образом, если мы знаем θ_i , то по формулам (2.6.3), (2.6.25) и (2.6.27) можем определить величины κ и γ . В рассматриваемой модели выводится дисперсионное уравнение для θ_i .

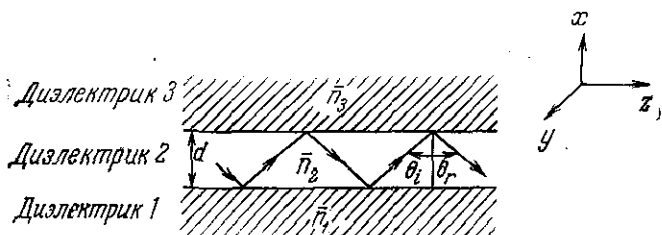
Дисперсионное уравнение

В этом разделе § 6 мы покажем, что волноводным модам соответствует набор дискретных значений угла θ_i . На рис. 2.6.2, а показана траектория луча, соответствующего волноводной моде в тонкой пленке из материала с показателем преломления \bar{n}_2 , окруженной слоями с показателями преломления \bar{n}_1 и \bar{n}_3 . На рис. 2.6.2, б изображены фазовые фронты плоских волн (штриховые линии), причем, чтобы не загромождать рисунок, фазовые фронты волн, отразившихся от верхней границы раздела сред, не показаны. Для вывода дисперсионного уравнения было предложено несколько эквивалентных методов [44, 50]. Здесь мы используем метод, описанный в работе Маркузе [50].

Важно обратить внимание на то обстоятельство, что на рис. 2.6.2, б все лучи, распространяющиеся в одном направлении, принадлежат одной и той же плоской волне и все точки фазового фронта должны соответствовать одинаковой фазе колебаний. Предположим, что исходящий из точки A в точку B луч AB не претерпевал отражения, так что соответствующий ему в точке B фазовый фронт направлен вдоль BD . Луч CD ,

исходящий из точки C в точку D , представляет собой отраженную волну, испытавшую полное внутреннее отражение в точках C и D . В точке D ему должен соответствовать фазовый фронт, находящийся в фазе с фазовым фронтом луча AB в точке B . Для этого необходимо, чтобы длина оптического пути (сдвиг фазы) луча AB отличалась от длины оптического пути луча CD на число, кратное 2π .

Соответствующая геометрия лучей представлена на рис. 2.6.2, в. Из него видно, что $CD = d/\cos \theta$, $DE = d \operatorname{tg} \theta$ и $EF = d/\operatorname{tg} \theta$, а также $BC = DE - EF$ и $AB = BC \sin \theta = d[\operatorname{tg} \theta - (1/\operatorname{tg} \theta)] \sin \theta$. При распространении луча в положительном направлении оптический путь равен произведению



а

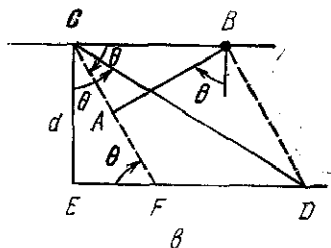
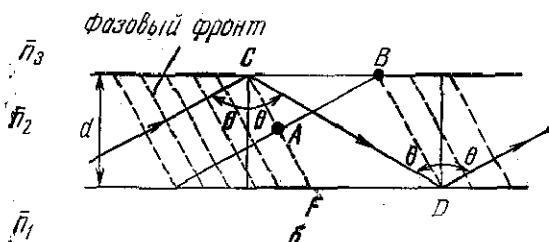


Рис. 2.6.2. Картина, поясняющая фазовое условие, из которого вытекает дисперсионное уравнение для трехслойного асимметричного плоского диэлектрического волновода; сдвиги Гуса — Хэнхена не учитываются. а — траектории лучей; б — взаимное расположение траекторий лучей и фазовых фронтов; в — геометрическая схема для определения оптических длин путей AB и CD .

постоянной распространения в направлении луча, взятой со знаком минус, $-\bar{n}_2 k_0$ на пройденное расстояние. Следовательно, разность оптических путей лучей CD и AB равна

$$-(CD - AB) \bar{n}_2 k_0 = -2\bar{n}_2 k_0 d \cos \theta. \quad (2.6.28)$$

Луч CD вследствие отражения в точках D и C приобретает дополнительные фазовые сдвиги Φ_1 и Φ_3 соответственно. Поэтому полный фазовый сдвиг определяется формулой

$$-2\bar{n}_2 k_0 d \cos \theta + \Phi_1 + \Phi_3 = -2m\pi, \quad (2.6.29)$$

где $m = 0, 1, 2, 3, \dots$ — номер моды. Из выражения (2.6.24) следует, что фазовый сдвиг при отражении равен 2ϕ , где ϕ определяется соотношением (2.6.25). Для ТЕ-волн получаются следующие формулы для фазовых сдвигов Φ_1 и Φ_3 :

$$\Phi_1 = 2 \operatorname{arctg} \left[\frac{(\bar{n}_2^2 \sin^2 \theta - \bar{n}_1^2)^{1/2}}{\bar{n}_2 \cos \theta} \right], \quad (2.6.30)$$

$$\Phi_3 = 2 \operatorname{arctg} \left[\frac{(\bar{n}_2^2 \sin^2 \theta - \bar{n}_3^2)^{1/2}}{\bar{n}_2 \cos \theta} \right]. \quad (2.6.31)$$

Подстановка выражений (2.6.30) и (2.6.31) в уравнение (2.6.29) дает дисперсионное уравнение для θ . Это трансцендентное уравнение содержит, к счастью, в качестве неизвестной только θ . Решение дисперсионного уравнения можно получить, если взять за начальное значение θ критический угол θ_c , определяемый выражением (2.6.20), и затем увеличивать угол θ до тех пор, пока он не будет удовлетворять уравнению (2.6.29). Тогда величина κ определяется при этом значении θ из соотношения (2.6.3), а значение γ на каждой границе — из соотношений (2.6.27) и (2.6.25). Заметим, что $\pi/2 > \theta_0 > \theta_m > \theta_c$, где θ_0 — решение дисперсионного уравнения, соответствующее основной моде, а θ_m — решения этого уравнения, соответствующие модам высшего порядка. Уравнение (2.6.29) справедливо и для ТМ-волн [44] при

$$\Phi_1 = 2 \operatorname{arctg} \left[\frac{\bar{n}_2^2 (\bar{n}_2^2 \sin^2 \theta - \bar{n}_1^2)^{1/2}}{\bar{n}_1 \bar{n}_2 \cos \theta} \right], \quad (2.6.32)$$

$$\Phi_3 = 2 \operatorname{arctg} \left[\frac{\bar{n}_2^2 (\bar{n}_2^2 \sin^2 \theta - \bar{n}_3^2)^{1/2}}{\bar{n}_3 \bar{n}_2 \cos \theta} \right]. \quad (2.6.33)$$

Важное отличие асимметричного трехслойного волновода от симметричного состоит в том, что для первого существует условие отсечки для основной моды, а для второго нет. Это условие зависит как от скачков показателя преломления на границах активного слоя, так и от толщины этого слоя. При приближении

значений \bar{n}_1 и \bar{n}_3 к значению \bar{n}_2 критический угол, определяемый соотношением (2.6.20), приближается к $\pi/2$, и член $-2\bar{n}_2 k_0 d \cos \theta$ становится слишком малым, чтобы в сумме с $\Phi_1 + \Phi_3$ давать нуль. Вопрос об отсечке мод в асимметричном трехслойном волноводе обсуждался Маркузе [51]. Распределение интенсивности света и коэффициент оптического ограничения в асимметричном волноводе, которым является ОГС-лазер, будут рассмотрены в конце параграфа. Там же приведены характерные значения параметров, при которых происходит отсечка основной моды.

Сдвиг Гуса — Хэнхена

Гус и Хэнхен [48] экспериментально установили, что луч, претерпевший полное внутреннее отражение на границе раздела диэлектриков, пространственно смещается вдоль линии пересечения этой границы с плоскостью падения. Рис. 2.6.3 поясняет этот эффект: луч как бы отражается от поверхности, параллельной границе раздела, но сдвинутой от последней на расстояние x_g . Вывод величины сдвига Гуса — Хэнхена для ТЕ-волн, который мы приводим здесь, предложен Когельником и др. [45].

Из формул (2.6.2) и (2.6.5) при $x = 0$, пренебрегая временной зависимостью, получаем следующее выражение для падающего электрического поля:

$$\mathcal{E}_{yi} = \mathcal{E}_i \exp(-i\beta z). \quad (2.6.34)$$

Для отраженного поля из соотношений (2.6.23) и (2.6.24) имеем

$$\mathcal{E}_{yr} = \mathcal{E}_{yi} \exp(j2\phi). \quad (2.6.35)$$

Фазовый сдвиг, обусловленный отражением, равен 2ϕ и, как следует из выражения (2.6.26), является функцией β . Для вычисления горизонтального сдвига Когельник и др. [45] рассмотрели пучок лучей, образованный двумя плоскими волнами с постоянными распространения $\beta \pm \Delta\beta$, падающими под мало различающимися углами. Соответствующее этому случаю поле при $x = 0$ определяется выражением

$$\mathcal{E}_{yi} = \mathcal{E}_i \{ \exp[-j(\beta + \Delta\beta)z] + \exp[-j(\beta - \Delta\beta)z] \}, \quad (2.6.36)$$

или

$$\mathcal{E}_{yi} = 2\mathcal{E}_i \cos(\Delta\beta z) \exp(-j\beta z). \quad (2.6.37)$$

При изменении β на малую величину $\Delta\beta$ фазовый сдвиг ϕ становится равным

$$\phi(\beta + \Delta\beta) = \phi(\beta) + (\partial\phi/\partial\beta)\Delta\beta = \phi + \Delta\phi. \quad (2.6.38)$$

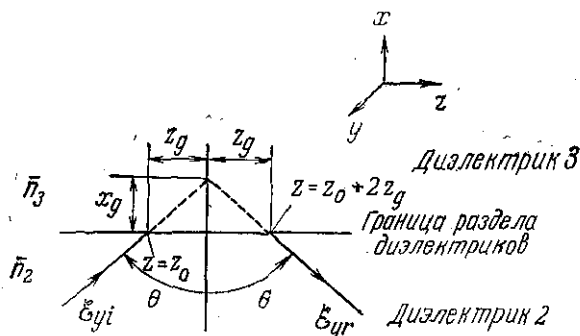


Рис. 2.6.3. Отражение луча на границе раздела диэлектриков с учетом сдвига Гуса — Хэнкена; $2z_g$ — горизонтальный сдвиг луча, x_g — кажущаяся глубина проникновения.

Преобразовав по формуле (2.6.35) каждый экспоненциальный член в выражении (2.6.36), описывающем падающую волну, получим для отраженной волны выражение, которое после замены $\phi(\beta \pm \Delta\beta)$ на $\phi \pm \Delta\phi$ принимает вид

$$\mathcal{E}_{yr} = \mathcal{E}_i \{ \exp[-j(\beta + \Delta\beta)z] \exp[j(2\phi + 2\Delta\phi)] + \exp[-j(\beta - \Delta\beta)z] \exp[j(2\phi - 2\Delta\phi)] \}. \quad (2.6.39)$$

Несложные преобразования этой формулы дают

$$\mathcal{E}_{yr} = \mathcal{E}_i \exp[-j(\beta z - 2\phi)] \{ \exp[j(\Delta\beta z - 2\Delta\phi)] + \exp[-j(\Delta\beta z - 2\Delta\phi)] \}, \quad (2.6.40)$$

или

$$\mathcal{E}_{yr} = 2\mathcal{E}_i \cos \Delta\beta (z - 2\Delta\phi/\Delta\beta) \exp[-j(\beta z - 2\phi)]. \quad (2.6.41)$$

Для того чтобы амплитуды в выражениях для \mathcal{E}_{yi} и \mathcal{E}_{yr} были одинаковы при $x=0$, аргументы у косинусов в выражениях (2.6.37) и (2.6.41) должны быть равны. Следовательно, если для падающей волны $z=z_0$, то для отраженной волны z должно равняться $z_0 + 2\Delta\phi/\Delta\beta$, что соответствует горизонтальному сдвигу, равному $2\Delta\phi/\Delta\beta$. Этот сдвиг обозначен на рис. 2.6.3 через $2z_g$, где

$$z_g = \partial\phi/\partial\beta. \quad (2.6.42)$$

Кажущаяся глубина проникновения, которой на рис. 2.6.3 соответствует величина x_g , связана с z_g соотношением

$$x_g = z_g/\text{tg } \theta. \quad (2.6.43)$$

Выражение для частной производной $\partial\phi/\partial\beta$ может быть получено из формулы (2.6.27) в виде [46]

$$z_g = \partial\phi/\partial\beta = (\kappa^2 + \gamma^2)^{-1} [\kappa(\partial\gamma/\partial\beta) - \gamma(\partial\kappa/\partial\beta)]. \quad (2.6.44)$$

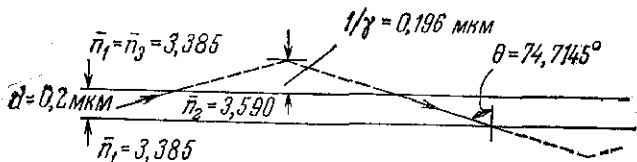


Рис. 2.6.4. Сдвиг Гуса — Хэнхена в симметричном GaAs—Al_{0,3}Ga_{0,7}As волноводе. Параметры волновода указаны на рисунке; $\lambda_0 = 0,90$ мкм.

Зависимость κ и γ от β определяется соотношениями (2.4.37) и (2.4.43), из которых следует

$$\kappa \partial \gamma / \partial \beta = \kappa \beta / \gamma, \quad (2.6.45)$$

$$\gamma \partial \kappa / \partial \beta = -\gamma \beta / \kappa. \quad (2.6.46)$$

Тогда горизонтальное смещение определяется формулой

$$z_g = \beta / \gamma \kappa = (1/\gamma) \operatorname{tg} \theta, \quad (2.6.47)$$

где использовано равенство (2.6.4). Для x_g из соотношения (2.6.43) получаем следующее выражение:

$$x_g = 1/\gamma. \quad (2.6.48)$$

Таким образом, кажущаяся глубина проникновения равна обратной величине постоянной затухания, определяемой из формулы (2.4.43) с учетом выражения (2.6.2) для β . Все полученные выражения справедливы для ТЕ-мод. Аналогичные выражения получаются и для ТМ-мод [44—46].

На рис. 2.5.11 представлена картина изменения электрического поля внутри волновода, образованного в симметричном ДГС-лазере, который рассматривался в § 5. Толщина активной области составляет 0,2 мкм, а ограничивающим слоям 1 и 3 соответствует состав с $x = 0,3$. На рис. 2.6.4 показана схематичная картина распространения волны в той же структуре с точки зрения модели зигзагообразных волн. Величина θ в этой модели для рассматриваемого примера может быть получена из соотношения (2.6.2) при значении β/k_0 , приведенном в § 5 настоящей главы (табл. 2.5.2). Как было показано выше, θ удовлетворяет также уравнению (2.6.29) с Φ_1 и Φ_3 , определяемыми выражениями (2.6.30) и (2.6.31). Сдвиг Гуса — Хэнхена, равный $2z_g$, составляет 1,434 мкм, т. е. горизонтальный сдвиг полностью отраженного луча в несколько раз больше толщины активной области.

Распределение интенсивности и коэффициент оптического ограничения в асимметричном трехслойном плоском волноводе

Выражения, описывающие электрическое поле в асимметричном волноводе, имеют такой же вид, как и полученные ранее для случая симметричного волновода. Однако максимум электрического поля, который всегда должен лежать внутри волновода, т. е. в диэлектрике 2 на рис. 2.6.2, *a*, не располагается в центре этого слоя при $\bar{n}_1 \neq \bar{n}_3$. Поэтому положение максимума поля является дополнительным параметром, который должен быть определен при рассмотрении асимметричного волновода. Постоянные затухания γ_1 и γ_3 также становятся различными. Ниже мы рассмотрим только основную моду ($m = 0$). Анализ мод высшего порядка можно найти у Тьена [44].

Уравнение (2.6.29) для $m = 0$ и $\kappa = \bar{n}_2 k_0 \cos \theta$ принимает вид

$$\kappa d = \phi_1 + \phi_3, \quad (2.6.49)$$

где $\Phi_1 = 2\phi_1$ и $\Phi_3 = 2\phi_3$ в соответствии с данным ранее определением. На рис. 2.3.2 за координату $x = 0$ брался центр волновода. Пусть в асимметричном волноводе, изображенном на рис. 2.6.5, x_2 и x_3 обозначают расстояния от границ раздела диэлектриков соответственно 1—2 и 2—3 до линии, соответствующей максимуму поля. При таких обозначениях имеем $d = x_2 + x_3$, и уравнение (2.6.49) переписывается в виде

$$\kappa(x_2 + x_3) = \phi_1 + \phi_3. \quad (2.6.50)$$

Поле должно описываться выражениями $A_e \cos(\kappa x_2)$ на границе раздела диэлектриков 1—2 и $A_e \cos(\kappa x_3)$ на границе раздела

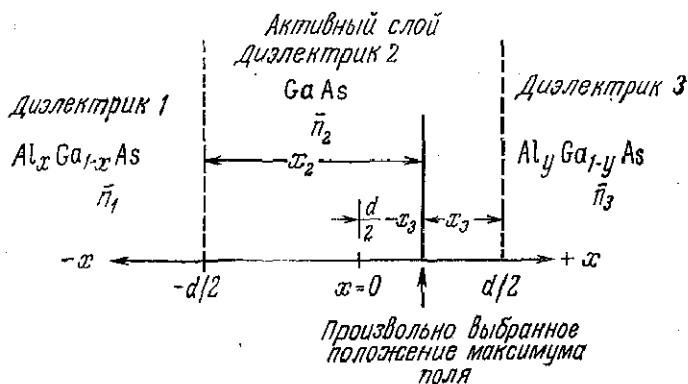


Рис. 2.6.5. Схематическое изображение асимметричного плоского трехслойного диэлектрического волновода и пояснение смысла некоторых величин, входящих в выражения, определяющие поле \mathcal{E}_y в волноводе. Максимуму поля соответствует координата $(d/2) - x_3$, где x_3 — расстояние от максимума поля до границы раздела диэлектриков 2—3 и $d = x_2 + x_3$.

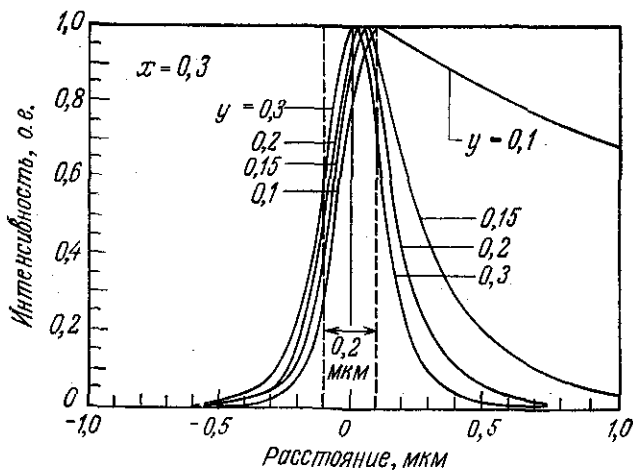


Рис. 2.6.6. Интенсивность электрического поля как функция расстояния от центра активного слоя в $\text{Al}_x\text{Ga}_{1-x}\text{As}-\text{GaAs}-\text{Al}_y\text{Ga}_{1-y}\text{As}$ асимметричном волноводе с $d = 0,2$ мкм и $\lambda_0 = 0,90$ мкм (1,38 эВ). Молярные доли AlAs: x — постоянная величина, y — переменная.

диэлектриков 2—3. Величины κx_2 и κx_3 выбираются так, чтобы

$$\kappa x_2 = \phi_1, \quad (2.6.51)$$

$$\kappa x_3 = \phi_3. \quad (2.6.52)$$

Из рис. 2.6.5 видно, что если центру волновода соответствует координата $x = 0$, то для достижения максимального значения поля \mathcal{E}_y в диэлектрике 2 в точке $x = (d/2) - x_3$ необходимо, чтобы аргумент y косинуса в выражении для \mathcal{E}_y равнялся нулю в этой точке. Поэтому в выражении (2.4.36) при $|x| < d/2$ аргумент x заменяется на $[x - (d/2 - x_3)]$, что дает

$$\mathcal{E}_y = A_e \cos \{ \kappa [x - (d/2 - x_3)] \}. \quad (2.6.53)$$

Из формулы (2.4.41) следует

$$\mathcal{E}_y = A_e \cos(\kappa x_2) \exp[-\gamma_1(-x + d/2)] \quad (2.6.54)$$

для $x < -d/2$ и

$$\mathcal{E}_y = A_e \cos(\kappa x_3) \exp[-\gamma_3(x - d/2)] \quad (2.6.55)$$

для $x > d/2$. Полученные выражения удовлетворяют граничным условиям, требующим, чтобы \mathcal{E}_y было непрерывно при $x = \pm d/2$. В формулах (2.6.53)–(2.6.55) опущен множитель $\exp[j(\omega t - \beta z)]$, который остается неизменным, как и в полученных ранее выражениях.

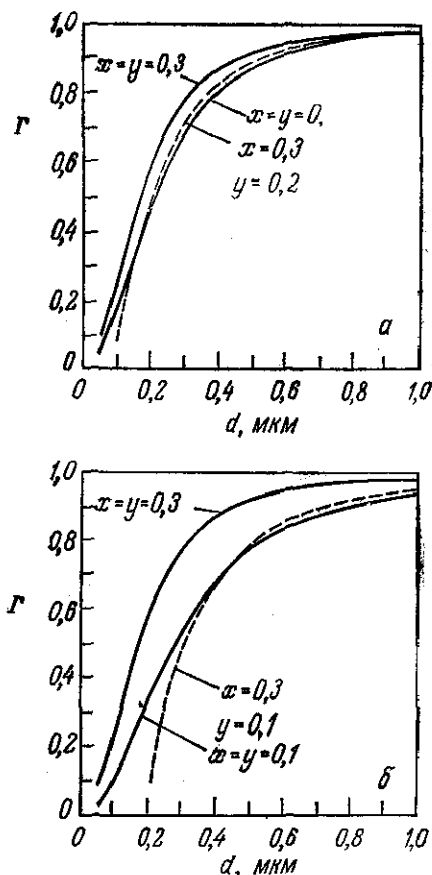


Рис. 2.6.7. Сравнение коэффициентов оптического ограничения в симметричном и асимметричном трехслойных плоских диэлектрических волноводах при $\lambda_0 = 0,90 \mu\text{м}$ (1,38 эВ). *а* — симметричные структуры с $x = y = 0,3$ и $x = y = 0,2$ и асимметричная с $x = 0,3, y = 0,2$; *б* — симметричные структуры с $x = y = 0,3$ и $x = y = 0,1$ и асимметричная с $x = 0,3, y = 0,1$.

Интенсивность электрического поля, равная $|\mathcal{E}_y|^2$, была вычислена по формулам (2.6.53)–(2.6.55) для $\text{Al}_x\text{Ga}_{1-x}\text{As}-\text{GaAs}-\text{Al}_y\text{Ga}_{1-y}\text{As}$ асимметричного трехслойного плоского диэлектрического волновода с $d = 0,2 \mu\text{м}$. Значения показателей преломления взяты из табл. 2.5.1 для $E = 1,38 \text{ эВ}$. Молярные доли AlAs брались следующие: $x = 0,3$; $y = 0,3$ для симметричного случая и $y = 0,2, 0,15, 0,1$ для демонстрации влияния асимметрии. На рис. 2.6.6 показаны полученные распределения интенсивности. Заметим, что максимум $|\mathcal{E}_y|^2$ сдвигается в направлении к слою с большим показателем преломления. Кроме

того, при изменении y от 0,15 до 0,1 распределение интенсивности резко меняется.

Рис. 2.6.7 показывает влияние асимметрии на коэффициент оптического ограничения. На рис. 2.6.7, а приведены кривые для $x = 0,3$ и $y = 0,2$, а на рис. 2.6.7, б — для $x = 0,3$ и $y = 0,1$. Кривая коэффициента оптического ограничения для асимметричного волновода практически совпадает с кривой, полученной для симметричного волновода, пассивные области которого содержат наименьшую молярную долю AlAs. Это совпадение наблюдается до тех пор, пока толщина d не становится достаточно малой, после чего значение коэффициента оптического ограничения асимметричного волновода резко снижается. Этот результат в применении к ДГС-лазерам показывает, что симметричные структуры имеют коэффициент оптического ограничения, максимально возможный для состава с данной молярной долей AlAs, и что влияние асимметрии наибольшее при малых толщинах активных слоев.

Распределение интенсивности и коэффициент оптического ограничения в ОГС-лазерах

Лазер на односторонней гетероструктуре представляет собой $n-p-P$ -структуру с молярной долей x арсенида алюминия в P -слое, обычно равной 0,2—0,5. Скачок показателя преломления

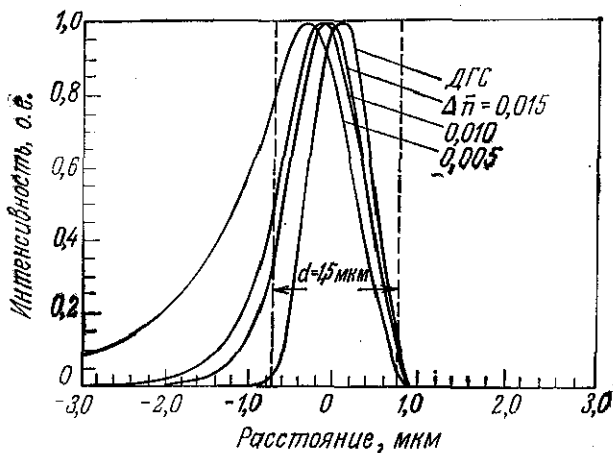


Рис. 2.6.8. Зависимость интенсивности электрического поля от расстояния в $n-p-P$ GaAs—Al_xGa_{1-x}As асимметричном волноводе при $\lambda_0 = 0,90$ мкм (1,38 эВ). Для сравнения приведено распределение интенсивности электрического поля в симметричной GaAs—Al_{0,3}Ga_{0,7}As ДГС. Скачок показателя преломления на $n-p$ -переходе обозначен через $\Delta\bar{n}$.

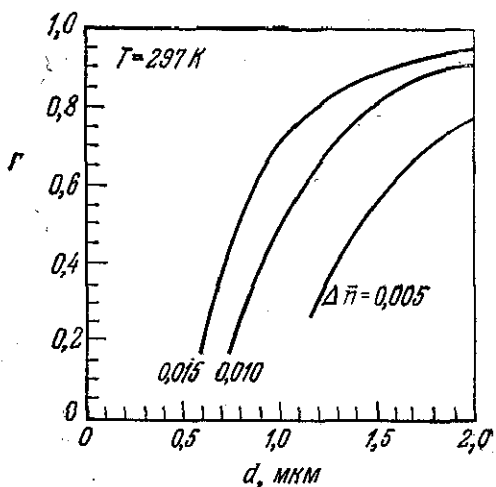


Рис. 2.6.9. Зависимость коэффициента оптического ограничения от толщины активного слоя асимметричного волновода в $n-p-P$ GaAs—Al_xGa_{1-x}As ОГС при $\lambda_0 = 0,90$ мкм (1,38 эВ). Скачок показателя преломления на $n-p$ -переходе обозначен через $\Delta\bar{n}$; для P -слоя $x = 0,3$. Кривые обрываются при значениях d , удовлетворяющих условию отсечки для основной моды.

ломления $\Delta\bar{n} = \bar{n}_2 - \bar{n}_3$ составляет примерно 0,14—0,21 на $p-P$ -гетерогранице. Величина скачка показателя преломления $\Delta\bar{n} = \bar{n}_2 - \bar{n}_1$ на $n-p$ -переходе на порядок меньше и зависит от ряда параметров, как это уже отмечалось в § 3 настоящей главы при рассмотрении гомолазеров. На рис. 2.6.8 показано, какое влияние на распределение интенсивности излучения оказывает скачок показателя преломления $\Delta\bar{n}$ на $n-p$ -переходе при $d = 1,5$ мкм и молярной доле x арсенида алюминия в P -слое, равной 0,3. Для сравнения на этом же рисунке представлено распределение интенсивности в GaAs—Al_{0,3}Ga_{0,7}As симметричном волноводе.

На рис. 2.6.9 приведены кривые, характеризующие зависимость коэффициента оптического ограничения от толщины активного слоя при значениях $\Delta\bar{n}$ на $p-n$ -переходе, равных 0,015, 0,010 и 0,005. Эти кривые обрываются при значениях d , при которых выполняется условие отсечки для основной моды. Выполнение этого условия при малых $\Delta\bar{n}$ означает, что вынужденное излучение в ОГС-лазерах может возникать только при толщинах активного слоя $d \geq 0,6-1,0$ мкм. Приведенные здесь расчетные значения коэффициента оптического ограничения будут использованы в § 3 гл. 7 при анализе экспериментальных данных по плотности порогового тока в ОГС-лазерах.

§ 7. РАСХОДИМОСТЬ ЛУЧА

Расчет дальней зоны

Детальное знание картины излучения в дальней зоне полупроводниковых лазеров весьма важно для их правильного применения. Например, от картины дальней зоны гетеролазера зависит способ ввода излучения в оптическое волокно, а также получаемая эффективность ввода [52]. Кроме того, распределение излучения в дальней зоне лазера характеризует распределение поля внутри волновода. Как будет показано в этом параграфе, картина излучения в дальней зоне представляет собой по существу интеграл Фурье от поля, распространяющегося в волноводе. Поэтому сравнение экспериментальной картины дальней зоны с расчетной, основанной на вычислении распределения поля в волноводе, показывает, насколько хорошо выражения для \mathcal{E}_y и численные значения показателей преломления, приведенные выше, описывают действительную картину распространения излучения в волноводе.

На рис. 2.7.1 показана структура, которую мы будем рассматривать в этом параграфе. Она представляет собой трехслойный плоский диэлектрический волновод с границей раздела полупроводник — воздух, расположенной при $z = 0$. Волновод бесконечен в направлении оси y ; координата $x = 0$ соответствует центру области с показателем преломления \bar{n}_2 . Для того чтобы получить картину излучения в дальней зоне в направлении, перпендикулярном плоскости перехода, мы определим

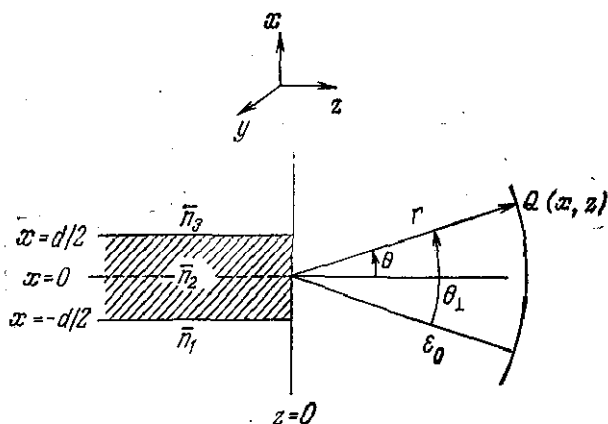


Рис. 2.7.1. К вычислению распределения интенсивности излучения в дальней зоне ДГС-лазера. Направляемая волна распространяется вдоль оси z в центральном слое с показателем преломления \bar{n}_2 ; \bar{n}_1 и \bar{n}_3 — показатели преломления прилегающих слоев. Размеры лазера в x - и y -направлениях велики по сравнению с толщиной d активной области.

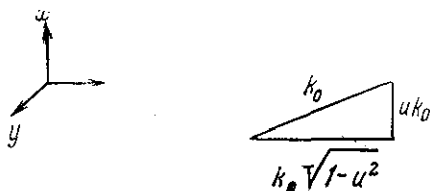


Рис. 2.7.2. Компоненты постоянной распространения k_0 , выраженные через безразмерный параметр u .

электрическое поле в точке $Q(x, z)$ в свободном пространстве. Как и в случае волновода, рассмотренном выше в § 4, для ТЕ-волн в свободном пространстве при $z > 0$ будем иметь $\mathcal{E}_x = \mathcal{E}_z = \mathcal{H}_y = 0$ и $\partial/\partial y = 0$. Тогда волновое уравнение принимает вид (2.4.13)

$$\frac{\partial^2 \mathcal{E}_y}{\partial x^2} + \frac{\partial^2 \mathcal{E}_y}{\partial z^2} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathcal{E}_y}{\partial t^2},$$

где $\varepsilon = \varepsilon_0$.

При использовании метода разделения переменных, как было показано в § 4, решения ищутся в виде

$$\mathcal{E}_y(x, z, t) = XZT. \quad (2.7.1)$$

В этом случае волновое уравнение для электрического поля, имеющего временную зависимость вида $\exp(j\omega t)$, запишется так:

$$(X''/X) + (Z''/Z) = \mu_0 \varepsilon_0 (T''/T) = -\omega^2 \mu_0 \varepsilon_0 = -k_0^2. \quad (2.7.2)$$

Рис. 2.7.2 показывает связь x - и z -компонент волнового вектора с k_0 . Из (2.7.2) следует

$$X''/X = -Z''/Z - k_0^2 = -u^2 k_0^2, \quad (2.7.3)$$

где $-u^2 k_0^2$ — постоянная разделения. Вследствие однородности среды используется всего одна постоянная разделения (которая обычно берется в виде произведения k_0 на безразмерную постоянную u). Решение для X имеет вид

$$X = A \exp(juk_0 x) + B \exp(-juk_0 x). \quad (2.7.4)$$

Из этого выражения видно, что uk_0 представляет собой x -компоненту волнового вектора, изображенного на рис. 2.7.2.

Из уравнения (2.7.3) вытекает уравнение для Z в следующей форме:

$$Z''/Z = -k_0^2 (1 - u^2), \quad (2.7.5)$$

откуда

$$Z = C \exp[j(1 - u^2)^{1/2} k_0 z] + D \exp[-j(1 - u^2)^{1/2} k_0 z]. \quad (2.7.6)$$

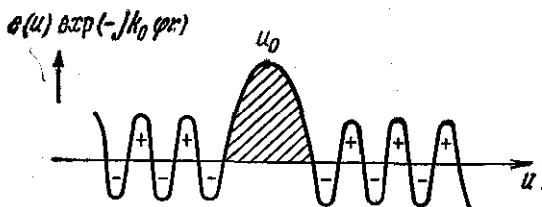


Рис. 2.7.3. График подынтегральной функции в выражении (2.7.11). Видно, что вклад в интеграл от области, где подынтегральная функция быстро осциллирует, примерно равен нулю.

В формулах (2.7.4) и (2.7.6) в дальнейшем учитываем только одно из двух слагаемых, поскольку другое соответствует противоположному знаку перед постоянной распространения. Тогда частное решение при заданном значении u будет иметь вид

$$\mathcal{E}_y(x, z) = e(u) \exp(-juk_0 x) \exp[-j(1-u^2)^{1/2} k_0 z], \quad (2.7.7)$$

где произведение коэффициентов B и D обозначено через $e(u)$ и является функцией u . Непосредственной проверкой легко убедиться, что выражение (2.7.7) является решением уравнения (2.4.13). Общее решение получается путем интегрирования частного решения по всем значениям u :

$$\mathcal{E}_y(x, z) = \int_{-\infty}^{\infty} e(u) \exp(-juk_0 x) \exp[-j(1-u^2)^{1/2} k_0 z] du. \quad (2.7.8)$$

Для того чтобы оценить интеграл (2.7.8), необходимо использовать метод стационарной фазы [53, 54]. В дальней зоне (рис. 2.7.1) $x = r \sin \Theta$ и $z = r \cos \Theta$. Тогда из формулы (2.7.8) имеем

$$\mathcal{E}_y(x, z) = \int_{-\infty}^{\infty} e(u) \exp\{-jk_0 r [u \sin \Theta + (1-u^2)^{1/2} \cos \Theta]\} du. \quad (2.7.9)$$

Вводя новое обозначение

$$\varphi = u \sin \Theta + (1-u^2)^{1/2} \cos \Theta, \quad (2.7.10)$$

общее решение (2.7.9) перепишем в виде

$$\mathcal{E}_y(x, z) = \int_{-\infty}^{\infty} e(u) \exp(-jk_0 \varphi r) du. \quad (2.7.11)$$

В дальней зоне r велико, и, следовательно, аргумент в скобках быстро меняется с изменением φ . Наличие быстро осциллирующей функции приводит к уменьшению вклада в интеграл от

функции $e(u)$. Как показано на рис. 2.7.3, площади участков под кривой, помеченные знаками $+$ и $-$, примерно равны друг другу. Однако может существовать область (на рис. 2.7.3 ей соответствует заштрихованный участок), в которой подынтегральное выражение дает ненулевой вклад в интеграл (2.7.11). В этой области в точке максимума u_0 имеем [53]

$$\left. \frac{\partial \varphi}{\partial u} \right|_{u=u_0} = 0. \quad (2.7.12)$$

(Дальнейшее обсуждение метода стационарной фазы можно найти в работе [53].)

Вблизи точки максимума u_0 обычно полагают

$$u = u_0 + v \quad (2.7.13)$$

и разлагают $\varphi(u)$ в ряд Тейлора:

$$\varphi(u) = \varphi(u_0) + \varphi'(u_0)v + [\varphi''(u_0)/2!]^2 v^2 + \dots \quad (2.7.14)$$

Если пренебречь в выражении (2.7.14) членами высшего порядка, то в нем останутся только два слагаемых, поскольку $\varphi'(u_0)$, согласно (2.7.12), равно нулю. Из (2.7.10) и (2.7.12) вытекает следующее уравнение:

$$\left. \frac{\partial \varphi}{\partial u} \right|_{u=u_0} = \sin \Theta - \frac{u_0}{(1-u_0^2)^{1/2}} \cos \Theta = 0. \quad (2.7.15)$$

Преобразовав это уравнение к виду $(1-u_0^2) \sin^2 \Theta = u_0^2 \cos^2 \Theta$, легко найти его решение

$$u_0 = \sin \Theta. \quad (2.7.16)$$

Тогда из выражений (2.7.10) и (2.7.16) получаем

$$\varphi(u_0) = 1 \quad (2.7.17)$$

и

$$\varphi''(u_0) = \left. \frac{d}{du} \left[\sin \Theta - \frac{u}{(1-u^2)^{1/2}} \cos \Theta \right] \right|_{u=u_0} = -\frac{1}{\cos^2 \Theta}. \quad (2.7.18)$$

Соотношения (2.7.17) и (2.7.18) позволяют получить из формулы (2.7.14) приближенное выражение для $\varphi(u)$. Используя это выражение, интеграл (2.7.11) можно переписать в виде

$$\mathcal{E}_y(x, z) = \int_{-\infty}^{\infty} e(u) \exp \left[-jk_0 r \left(1 - \frac{v^2}{2 \cos^2 \Theta} \right) \right] dv. \quad (2.7.19)$$

Разложим $e(u)$ в ряд

$$e(u) = e(u_0 + v) = e(u_0) + e'(u_0)v + \dots \quad (2.7.20)$$

Пренебрегая членами высшего порядка и учитывая тот факт, что при интегрировании нечетный член $ve'(u_0)$ дает нуль, получим

$$e(u) \approx e(u_0). \quad (2.7.21)$$

В этом случае интеграл примет вид

$$\mathcal{E}_y(x, z) = e(u_0) \exp(-jk_0 r) \int_{-\infty}^{\infty} \exp(jv^2 k_0 r / 2 \cos^2 \Theta) dv. \quad (2.7.22)$$

Полагая $a = k_0 r / 2 \cos^2 \Theta$ и учитывая равенство [53]

$$\int_{-\infty}^{\infty} \exp(jav^2) dv = (\pi/a)^{1/2} \exp(j\pi/4), \quad (2.7.23)$$

получаем

$$\mathcal{E}_y(x, z) = (2\pi/k_0 r)^{1/2} \exp(j\pi/4) \cos \Theta e(u_0) \exp(-jk_0 r). \quad (2.7.24)$$

Для того чтобы выразить множитель $e(u_0) \exp(-jk_0 r)$ через значения поля в волноводе, положим в выражении (2.7.8) $z=0$. Полю внутри волновода будет соответствовать $\mathcal{E}_y(x, 0)$, поскольку $\mathcal{E}_y(x, z)$ должно быть непрерывно при $z=0$. Выражение (2.7.8) при $z=0$ имеет вид

$$\mathcal{E}_y(x, 0) = \int_{-\infty}^{\infty} e(u) \exp(-juk_0 x) du. \quad (2.7.25)$$

Из этого выражения видно, что $\mathcal{E}_y(x, 0)$ является фурье-образом величины

$$e(u) = (1/2\pi) \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) \exp(juk_0 x) dk_0 x. \quad (2.7.26)$$

Учитывая равенство (2.7.16) и подставляя выражение (2.7.26) в (2.7.24), получим формулу, определяющую поле в свободном пространстве:

$$\begin{aligned} \mathcal{E}(\Theta) &= (k_0/2\pi r)^{1/2} \exp(j\pi/4) \exp(-jk_0 r) \cos \Theta \times \\ &\times \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) \exp(jk_0 x \sin \Theta) dx. \end{aligned} \quad (2.7.27)$$

Это выражение и является искомым результатом. Оно представляет картину излучения в дальней зоне в виде произведения множителя, содержащего $\cos \Theta$, на интеграл Фурье от распределения электрического поля моды внутри волновода. В дальней зоне отношение интенсивности излучения, распространяющегося в направлении, определяемом углом Θ , к интенсивности

излучения, распространяющегося в направлении $\Theta = 0$, равно

$$\frac{I(\Theta)}{I(0)} = \frac{|\mathcal{E}(\Theta)|^2}{|\mathcal{E}(0)|^2} = \frac{\cos^2 \Theta \left| \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) \exp(j \sin \Theta k_0 x) dx \right|^2}{\left| \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) dx \right|^2}. \quad (2.7.28)$$

В большинстве расчетов картины излучения в дальней зоне полупроводникового лазера использовались выражения, аналогичные выражению (2.7.27), за тем исключением, что в них не учитывался множитель $\cos \Theta$. Наличие этого множителя, называемого угловым фактором Гюйгенса, приводит к значительному уменьшению вычисляемого значения интенсивности излучения, распространяющегося под большими углами. Киркби и Томпсон [55] использовали множитель $(\cos \Theta)^{1/2}$ на том основании, что он «не противоречит каким-либо экспериментальным результатам». Хокэм [56] путем довольно сложных вычислений показал, что множитель $g(\Theta)$ в выражении для $\mathcal{E}_y(x, z)$ следует брать в виде

$$g(\Theta) = \frac{2 \cos \Theta \{(\bar{n}^2 - \sin^2 \Theta)^{1/2} + (\beta/k_0)\}}{\cos \Theta + (\bar{n}^2 - \sin^2 \Theta)^{1/2}}, \quad (2.7.29)$$

где \bar{n} — показатель преломления в лазере. Левин [57, 58], проводя дальнейшее уточнение выражения для $g(\Theta)$, заменил последнее слагаемое в числителе формулы (2.7.29) на величину $\beta/k_0 \bar{n}$, однако это не оказало заметного влияния на численное значение $g(\Theta)$. Определяемый формулой (2.7.29) множитель $g(\Theta)$ ведет себя почти как $\cos \Theta$. Выражения, описывающие картину излучения в дальней зоне, были предложены также Батлером и Зоруччи [59].

Сравнение экспериментальной и расчетной картин излучения в дальней зоне симметричного ДГС-лазера

Экспериментальная картина излучения в дальней зоне обычно характеризуется расходимостью луча, определяемой как полный угол, измеренный на уровне половины (3 дБ) максимальной интенсивности. В плоскости $p-n$ -перехода расходимость луча $\Theta_{||}$, определенная указанным выше образом, составляет $\sim 10^\circ$. Мы не будем здесь останавливаться на рассмотрении расходимости луча в плоскости $p-n$ -перехода, поскольку до сих пор предметом нашего исследования были волны, не ограниченные в направлении оси y . Обсуждение этого вопроса проведено в гл. 7, посвященной лазерам с полосковой геометрией,

в которых формируется волновод также и по оси y . Расходимость луча в плоскости, перпендикулярной плоскости $p-n$ -перехода, зависит как от разности показателей преломления на границах раздела между слоями, так и от толщины активного слоя.

В случае симметричного ДГС-лазера подстановка выражений (2.4.36) и (2.4.41), определяющих \mathcal{E}_y , в интегралы, входящие в соотношение (2.7.28), позволяет переписать это соотношение в виде [60]

$$\frac{I(\Theta)}{I(0)} = \frac{|g(\Theta)|^2}{|g(0)|^2} \left[\gamma \kappa^2 \left(\frac{\gamma \cos[k_0(d/2) \sin \Theta] - k_0 \sin \Theta \sin[k_0(d/2) \sin \Theta]}{(\kappa^2 - k_0^2 \sin^2 \Theta)(\gamma^2 + k_0^2 \sin^2 \Theta)} \right) \right]^2, \quad (2.7.30)$$

где вместо $\cos \Theta$ использован множитель $g(\Theta)$, определяемый формулой (2.7.29). В других случаях интегрирование в соотношении (2.7.28) может быть проведено численно. На рис. 2.7.4 приведена экспериментальная картина дальней зоны симметричного GaAs—Al_{0,3}Ga_{0,7}As ДГС-лазера с $d = 0,18$ мкм. Экспериментальные значения $I(\Theta)/I(0)$ для распределения излучения в плоскости, перпендикулярной плоскости $p-n$ -перехода, представлены на рис. 2.7.4 светлыми кружками и сравниваются со значениями $I(\Theta)/I(0)$, вычисленными по формуле (2.7.30). Вычисленное и экспериментальное значение расходимости пучка по уровню половины максимальной интенсивности Θ_{\perp} отличаются

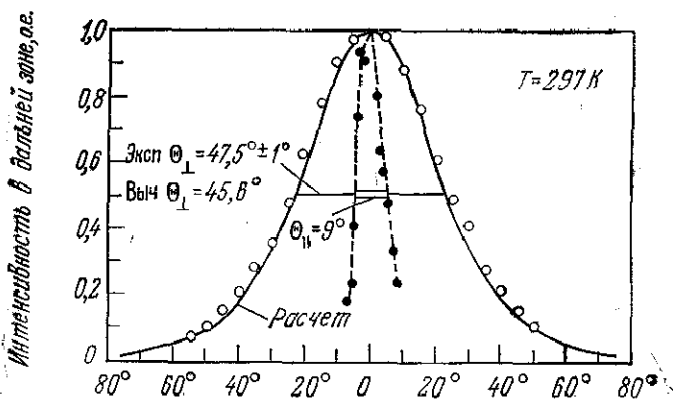


Рис. 2.7.4. Картина дальней зоны симметричного ДГС-лазера с $d = 0,18$ мкм и $x = 0,3$. Сплошная линия — отношение $I(\Theta)/I(0)$, рассчитанное по формуле (2.7.30). Экспериментальное распределение интенсивности в дальней зоне: светлые кружки — в плоскости, перпендикулярной плоскости $p-n$ -перехода, темные кружки — в плоскости $p-n$ -перехода; Θ_{\perp} и Θ_{\parallel} — соответствующие расходимости луча, определенные по уровню половины максимальной интенсивности [61].

между собой только на $\sim 1^\circ$, а вычисленная и экспериментальная зависимости $I(\Theta)/I(0)$ имеют одинаковую форму даже при больших углах. Если вместо коэффициента $[g(\Theta)]^2$ использовать $\cos^2 \Theta$, то вычисленное значение Θ_1 будет меньше на $1,5^\circ$. Экспериментальная картина излучения в дальней зоне в плоскости $p-n$ -перехода представлена на рис. 2.7.4 темными кружками.

На рис. 2.7.5 сравниваются с экспериментальными значениями [61] вычисленные по формуле (2.7.30) значения расходимости пучка по уровню половины максимальной интенсивности Θ_1 , представленные в виде функции толщины активной области симметричного GaAs—Al_xGa_{1-x}As ДГС-лазера для ряда значений молярной доли AlAs. Расчетная расходимость луча определяется как удвоенный угол Θ , при котором формула (2.7.30) дает значение $I(\Theta)/I(0) = 0,5$. Единственное существенное расхождение экспериментальных и теоретических результатов возникает при $x = 0,30$. Однако изменение значений показателей преломления, которые используются в выражениях для \mathcal{E}_y при получении соотношения (2.7.30), менее чем на 1% привело бы

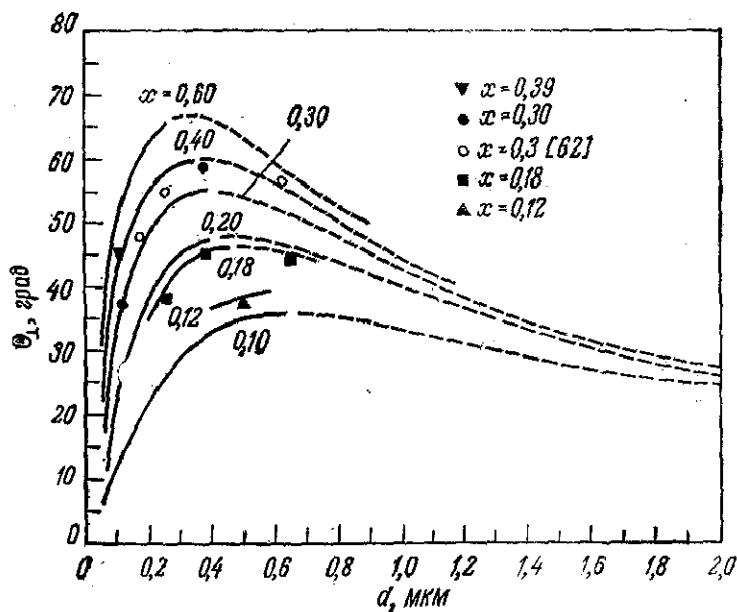


Рис. 2.7.5. Расходимость луча, измеренная на уровне половины максимальной интенсивности (3 дБ), как функция толщины активного слоя d и состава x . Сплошные и штриховые кривые — расходимость, вычисленная по формуле (2.7.30) для основной TE-моды. Штриховые части кривых соответствуют толщинам активных слоев, при которых могут возникать моды высшего порядка. Экспериментальные данные взяты из работы [61].

к полному согласию эксперимента с теорией и при этом значении x . Следует отметить, что наибольшая неопределенность возникает при определении для каждого частного случая конкретного значения x . Из рис. 2.7.4 и 2.7.5 можно сделать вывод, что численные значения \bar{n} , выражения, определяющие \mathcal{E}_y , а также полученная зависимость $I(\Theta)$ правильно описывают процесс распространения излучения в GaAs—Al_xGa_{1-x}As ДГС-лазерах.

Приближенное выражение для расходимости луча

Приведенные в § 5 настоящей главы формулы, аппроксимирующие \mathcal{E}_y , были использованы Думке [43] с целью получения приближенного выражения для Θ_{\perp} . Формула (2.7.27), описывающая поле в свободном пространстве, при использовании выражения (2.5.14) для $\mathcal{E}_y(x, 0)$ принимает вид

$$\mathcal{E}(\Theta) = (k_0/2\pi r)^{1/2} \exp(j\pi/4) \exp(-jk_0 r) \cos \Theta \times \\ \times \left(E_0 \int_{-\infty}^0 \exp[(\gamma + jk_0 \sin \Theta) x] dx + E_0 \int_0^{\infty} \exp[-(\gamma - jk_0 \sin \Theta) x] dx \right), \quad (2.7.31)$$

или

$$\mathcal{E}(\Theta) = (k_0/2\pi r)^{1/2} \exp(j\pi/4) \exp(-jk_0 r) \cos \Theta \times \\ \times [2E_0 \gamma / (\gamma^2 + k_0^2 \sin^2 \Theta)]. \quad (2.7.32)$$

При малых углах Θ , таких, что $\cos \Theta \approx 1$, интенсивность в дальней зоне проявляет следующую зависимость от Θ :

$$I(\Theta) \propto |\mathcal{E}(\Theta)|^2 \propto \gamma^2 / (\gamma^2 + k_0^2 \sin^2 \Theta)^2. \quad (2.7.33)$$

Угол $\Theta_{1/2}$, соответствующий направлению, в котором интенсивность излучения равна половине максимальной, определяется из равенства

$$I(\Theta_{1/2})/I(0) = \gamma^4 / (\gamma^2 + k_0^2 \sin^2 \Theta_{1/2})^2 = 1/2, \quad (2.7.34)$$

и, следовательно,

$$\sin \Theta_{1/2} = (\sqrt{2} - 1)^{1/2} \gamma / k_0. \quad (2.7.35)$$

Если угол $\Theta_{1/2}$ мал настолько, что $\sin \Theta_{1/2} \approx \Theta_{1/2}$, то расходимость луча Θ_{\perp} определяется формулой [43]

$$\Theta_{\perp} (\text{рад}) = 2\Theta_{1/2} \approx 2(\sqrt{2} - 1)^{1/2} \gamma / k_0 = 0,20\gamma\lambda_0, \quad (2.7.36)$$

справедливой для малых d и выражающей Θ_{\perp} в радианах.

Постоянная затухания γ описывается приближенно выражением (2.5.11)

$$\gamma \approx (\bar{n}_2^2 - \bar{n}_1^2) k_0^2 d/2,$$

использование которого позволяет представить формулу (2.7.36) в виде

$$\Theta_{\perp} \text{ (рад)} \approx 4,0 (\bar{n}_2^2 - \bar{n}_1^2) d/\lambda_0. \quad (2.7.37)$$

Согласно выражению (2.5.17), величина $\bar{n}_2^2 - \bar{n}_1^2$ примерно равна $1,42\bar{n}_2x$. Тогда из формулы (2.7.37) для значения $\bar{n}_2 = 3,590$ получаем окончательное выражение, определяющее Θ_{\perp} :

$$\Theta_{\perp} \text{ (рад)} = 20xd/\lambda_0 \quad (2.7.38)$$

или

$$\Theta_{\perp} \text{ (град)} = 1,1 \cdot 10^3 xd/\lambda_0. \quad (2.7.39)$$

Для симметричного GaAs—Al_{0,3}Ga_{0,7}As ДГС-лазера при $d = 0,1$ мкм и $\lambda_0 = 0,90$ мкм приближенная формула (2.7.39) дает $\Theta_{\perp} = 36,7^\circ$, в то время как из точного выражения (2.7.30) получается значение Θ_{\perp} , равное $39,5^\circ$. Согласие между приближенным и точным выражениями улучшается при уменьшении Θ_{\perp} . Думке [43] вывел также формулу, определяющую Θ_{\perp} при больших значениях d , однако она применима лишь при таких больших d , при которых в волноводе возбуждаются моды высшего порядка, а не основная мода.

§ 8. КОЭФФИЦИЕНТ ОТРАЖЕНИЯ НА ТОРЦЕВЫХ ГРЯНЯХ ЛАЗЕРА

Наблюдавшееся излучение ДГС-лазеров имело в большинстве случаев электрическую поляризацию, а не магнитную. Кроме того, при толщинах активного слоя, при которых возможно возбуждение мод высшего порядка, часто наблюдалась только основная мода. Как будет показано в § 8 гл. 3, необходимое для достижения порога генерации усиление зависит как от коэффициента оптического ограничения для моды, так и от коэффициента отражения на торцевых гранях для этой моды. Коэффициент отражения на зеркале лазера R_m для конкретной моды входит в пороговое условие для усиления в виде величины $(1/L) \ln(1/R_m)$ (см. § 8 гл. 3 и § 4 гл. 7). Коэффициент оптического ограничения не отличается существенно для ТЕ- и ТМ-поляризаций. Из всего сказанного следует, что отбор преимущественной поляризации излучения обусловлен главным образом различием коэффициентов отражения на торцевых гранях для ТЕ- и ТМ-волн, а момент, когда при увеличении толщины активного слоя начинают возбуждаться моды высшего порядка, определяется как коэффициентом оптического ограничения, так и коэффициентом отражения. Поэтому исследование коэффициента отражения на торцевых гранях необходимо для понимания происходящего в лазерах отбора мод и поляризации излучения.

Если поле волноводной моды находится в основном внутри активного слоя, то модель зигзагообразных волн позволяет определить угол падения соответствующей плоской волны на зеркальную грань лазера, и коэффициент отражения для нее может быть определен по формулам Френеля. К сожалению, в большинстве случаев распространяющиеся в лазерах волны проникают в прилегающие к активному слою диэлектрические области, как это показано на рисунках в § 5 настоящей главы. Поэтому электрическое поле у зеркальной грани не может быть представлено в виде одной плоской волны. Для получения численных значений R_m были использованы два близких подхода. Райнхарт и др. [63], Гордон [64] и Крупка [65] использовали метод аппроксимации, предложенный Мак-Кенной [66]. Икегами [67] определял коэффициент отражения, исходя из граничных условий на границе раздела полупроводник — воздух. Поскольку исследования по определению коэффициента отражения на торцевых гранях лазера не привели к получению решений в замкнутой форме, мы рассмотрим этот вопрос только качественно. Поведение коэффициента отражения будет проиллюстрировано численными результатами, полученными Икегами [67].

В методе, предложенном Мак-Кенной [66], электрическое поле при $z = 0$ (рис. 2.7.1) представляется в виде разложения по плоским волнам. Далее по формулам Френеля определяется коэффициент отражения для каждой из этих плоских волн, падающих под разными углами на торцевую грань лазера, а затем эти коэффициенты отражения суммируются с целью получения коэффициента отражения для моды. Поскольку часть поля приходится на активный слой с коэффициентом преломления \bar{n}_2 , а остальная часть — на прилегающие диэлектрические слои с коэффициентами преломления \bar{n}_1 , в формулах Френеля в качестве показателя преломления полупроводника используется эффективный показатель преломления β/k_0 [см. формулу (2.6.19)]. Этот способ расчета описан в приложении к работе [65]. Икегами разложил электрическое и магнитное поля на плоские волны и коэффициент отражения при $z = 0$ получил из требования непрерывности на этой границе ТЕ- и ТМ-полей. В обоих подходах необходим большой объем вычислений на ЭВМ.

В работе [67] содержится много рисунков, иллюстрирующих полученные результаты. Два важнейших результата отображены на рис. 2.8.1 и 2.8.2. Приведенные на них кривые рассчитаны для длины волны излучения 0,86 мкм и $\bar{n}_2 = 3,60$. Показатель преломления \bar{n}_0 среды, окружающей лазер, полагался равным единице. Исследовался коэффициент отражения по интенсивности R_m , равный квадрату абсолютной величины ко-

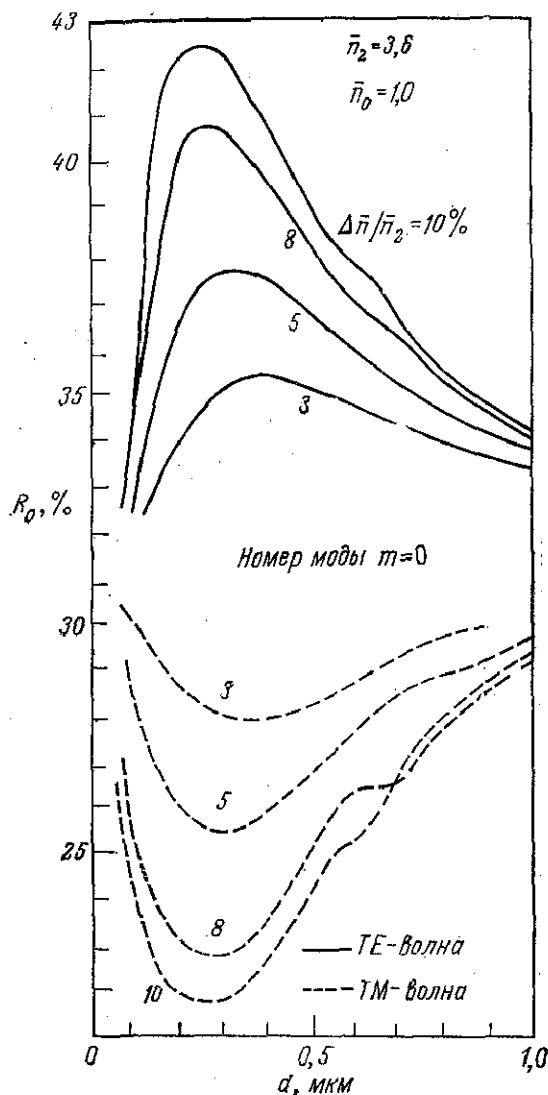


Рис. 2.8.1. Коэффициент отражения R_0 для основной моды как функция толщины активного слоя в симметричных ДГС-лазерах. Параметром является отношение $\Delta \bar{n} / \bar{n}_2 = [(\bar{n}_2 - \bar{n}_1) / \bar{n}_2] \cdot 100\%$ [67].

эфициента отражения по амплитуде поля. На рис. 2.8.1 коэффициент отражения R_0 для основной моды представлен как функция толщины активного слоя. Параметром семейства кривых служит относительная разность показателей преломления

$$\Delta \bar{n} / \bar{n}_2 = (\bar{n}_2 - \bar{n}_1) / \bar{n}_2, \quad (2.8.1)$$

выраженная в процентах. Сплошные кривые соответствуют ТЕ-волнам, штриховые кривые — ТМ-волнам. Результаты расчетов представленные на этом рисунке, показывают, что более высокие значения коэффициента отражения для ТЕ-моды определяют выбор этой поляризации для электромагнитных колебаний в ДГС-лазерах. По мере того как d возрастает, внутри активного слоя локализуется все большая часть моды. При этом падение зигзагообразной волны на торцевую грань приближается к нормальному, и коэффициент отражения для нее приближается к своему значению при нормальном падении. При уменьшении d величина R_0 для ТЕ-волн увеличивается, поскольку в этом случае падение плоских волн на торцевую грань

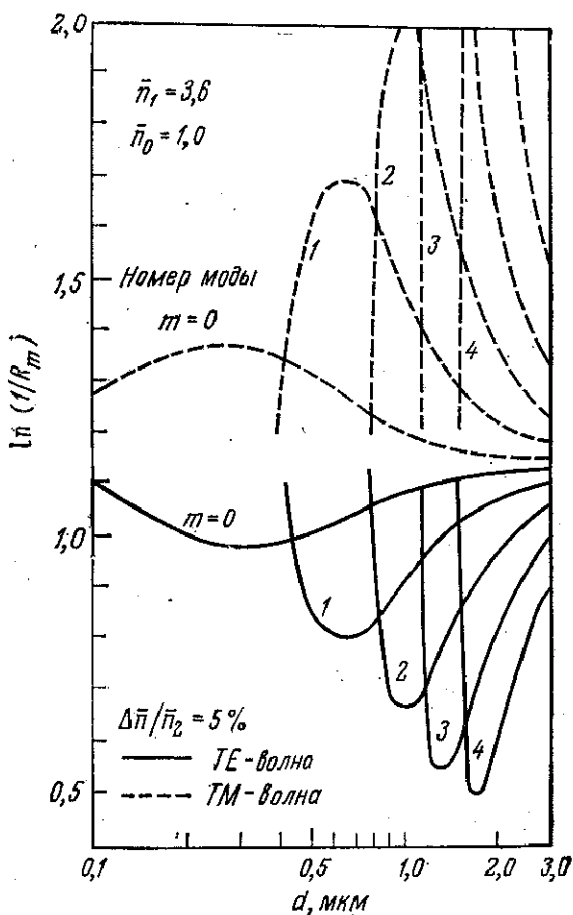


Рис. 2.8.2. Потери $\ln(1/R_m)$ для m -й моды как функция толщины активного слоя в ДГС-лазерах. Значение параметра $\Delta\bar{n}/\bar{n}_2$ равно 5% [67].

отклоняется от нормального. Угол полного внутреннего отражения составляет $\sim 16^\circ$, и при приближении к нему угла падения величина R_0 должна стремиться к единице. Однако когда d становится меньше, чем длина волны в полупроводнике (0,86 мкм/3,6), мода частично проникает в прилегающие диэлектрические слои, и R_0 уменьшается. Максимум R_0 достигается при толщине активного слоя, близкой к λ_0/\bar{n}_2 . Коэффициент отражения для ТМ-моды уменьшается, когда d становится меньше 1 мкм, поскольку угол падения соответствующей плоской волны все больше отличается от нормального и приближается к углу Брюстера, при котором отражение для ТМ-поляризации минимально. Однако при приближении d к величине λ_0/\bar{n}_2 ТМ-мода начинает частично переходить из активного слоя в прилегающие диэлектрические слои, и значение коэффициента отражения возрастает.

Поскольку потери для моды определяются выражением $(1/L)\ln(1/R_m)$, удобно исследовать величину $\ln(1/R_m)$, а не R_m . На рис. 2.8.2 приведены расчетные кривые потерь для мод с $m = 0-4$ при $\Delta\bar{n}/\bar{n}_2 = 5\%$, чему соответствует молярная доля AlAs, примерно равная 0,25. Как показано на этом рисунке, мода с $m = 1$ может возбуждаться уже начиная с $d \approx 0,4$ мкм. Тем не менее экспериментально эта мода не наблюдается вплоть до значений $d \approx 0,6$ мкм [61]. Потери для мод с $m = 0$ и $m = 1$ при $d = 0,45$ мкм равны, однако, как следует из рис. 2.5.14, при этом значении d величина Γ в два раза больше для $m = 0$, чем для $m = 1$. По мере приближения d к значению, равному 0,6 мкм, член, соответствующий потерям, становится меньше для моды с $m = 1$, а значения Γ для $m = 0$ и $m = 1$ для $d = 0,6$ мкм не отличаются между собой так сильно, как при меньших значениях d . Мода с $m = 2$ наблюдалась, начиная примерно с $d = 1,0$ мкм. Из рис. 2.8.2 видно, что вблизи этого значения d потери для моды с $m = 2$ имеют минимальное значение. Поэтому ясно, что существование мод высшего порядка требует, чтобы для них потери при отражении, определяемые величиной $\ln(1/R_m)$, имели приблизительно минимальные значения. Очевидно также, что для отбора мод важны как коэффициент оптического ограничения, так и коэффициент отражения на торцевых гранях лазера.

§ 9. ВОЛНОВОДЫ В ЧЕТЫРЕХ- И ПЯТИСЛОЙНЫХ ГЕТЕРОСТРУКТУРАХ

Гетеролазеры с отдельным ограничением

Дальнейшее раскрытие свойств гетероструктур связано с использованием многослойных структур. Простейшей из них с точки зрения расчета оптических полей является симметричный пя-

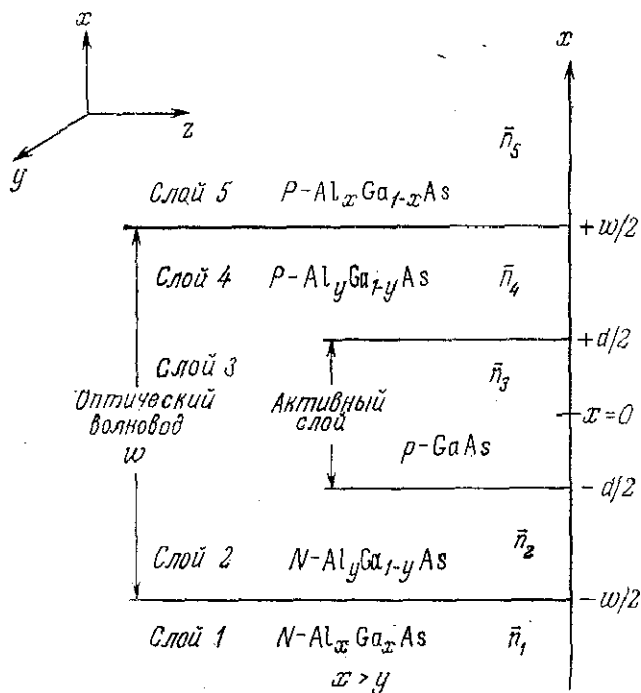


Рис. 2.9.1. Представление ДГС—РО-лазера как симметричного пятислойного диэлектрического плоского волновода.

тислойный волновод, обеспечивающий раздельное электронное и оптическое ограничение [62, 68—70]. Эта структура была названа двусторонней гетероструктурой с раздельным ограничением (ДГС—РО) [68—70]. На рис. 2.9.1 дано схематическое изображение такой пятислойной структуры с толщиной активной области d и толщиной оптического резонатора w . Тип проводимости и молярная доля AlAs указаны для каждого слоя. В ДГС—РО-лазере для состава $\text{Al}_y\text{Ga}_{1-y}\text{As}$ -слоя выбирается $y \approx 0,1 - 0,12$. Это делает скачок ширины запрещенной зоны на границе между GaAs и $\text{Al}_y\text{Ga}_{1-y}\text{As}$ достаточно большим для того, чтобы ограничивать большинство инжектированных носителей областью внутри GaAs-слоя (см. § 6 гл. 4). В то же время скачок показателя преломления на этой границе при малых значениях d не обеспечивает оптического ограничения. Большая величина скачка показателя преломления между $\text{Al}_y\text{Ga}_{1-y}\text{As}$ и $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ($x > y$) служит для создания оптического ограничения и тем самым формирует оптический волновод. Оптические поля в симметричном ДГС—РО-лазере рас-

смагиваются в этом разделе § 9, а рабочие характеристики обсуждаются в § 5 гл. 7.

Для симметричного плоского диэлектрического ДГС — РО-волновода, изображенного на рис. 2.9.1, границе раздела полупроводник — воздух соответствует $z = 0$. Волновод бесконечен в направлении оси y , и центр активной области из GaAs лежит в плоскости $x = 0$. В этой структуре $x > y$ и $\bar{n}_3 > \bar{n}_2 = \bar{n}_4 > \bar{n}_1 = \bar{n}_5$. Будет рассмотрена лишь основная ТЕ-мода, поскольку только она наблюдалась экспериментально в таких структурах с малыми толщинами активного слоя ($d < 0,5$ мкм). Как было указано в § 4 настоящей главы, для ТЕ-волн $\mathcal{E}_x = \mathcal{E}_z = 0$, а \mathcal{E}_y является решением уравнения (2.4.28). В активном слое GaAs, где $|x| < d/2$, решение для электрического поля получается такое же, как и для симметричного трехслойного плоского волновода (2.4.36) [70]:

$$\mathcal{E}_y = A \cos(\kappa_3 x) \exp[j(\omega t - \beta z)],$$

где

$$\kappa_3^2 = \bar{n}_3^2 k_0^2 - \beta^2. \quad (2.9.1)$$

Магнитное поле \mathcal{H}_z определяется формулой (2.4.40). В области с показателем преломления $\bar{n}_2 = \bar{n}_4$, которой соответствует $d/2 < |x| < \omega/2$, имеем

$$\mathcal{E}_y = [B \cos(\kappa_2 x) + C \sin(\kappa_2 x)] \exp[j(\omega t - \beta z)], \quad (2.9.2)$$

где

$$\kappa_2^2 = \bar{n}_2^2 k_0^2 - \beta^2. \quad (2.9.3)$$

Наконец, для области $\text{Al}_x\text{Ga}_{1-x}\text{As}$, где $|x| > \omega/2$, получаем

$$\mathcal{E}_y = D \exp(-\gamma x) \exp[j(\omega t - \beta z)], \quad (2.9.4)$$

где

$$\gamma^2 = \beta^2 - \bar{n}_1^2 k_0^2. \quad (2.9.5)$$

Дисперсионное уравнение получается при использовании соотношений (2.4.34) и (2.4.35), выражающих требование непрерывности \mathcal{E}_y и \mathcal{H}_z на каждой границе раздела. Поскольку структура симметрична, можно рассматривать только границы раздела между слоями 1 и 2, 2 и 3 или слоями 3 и 4, 4 и 5. Граничные условия приводят к однородной системе линейных уравнений для коэффициентов:

$$A \cos(\kappa_3 d/2) - B \cos(\kappa_2 d/2) - C \sin(\kappa_2 d/2) = 0, \quad (2.9.6)$$

$$-A \kappa_3 \sin(\kappa_3 d/2) + B \kappa_2 \sin(\kappa_2 d/2) - C \kappa_2 \cos(\kappa_2 d/2) = 0, \quad (2.9.7)$$

$$B \cos(\kappa_2 \omega/2) + C \sin(\kappa_2 \omega/2) - D \exp(-\gamma \omega/2) = 0, \quad (2.9.8)$$

$$-B \kappa_2 \sin(\kappa_2 \omega/2) + C \kappa_2 \cos(\kappa_2 \omega/2) + D \gamma \exp(-\gamma \omega/2) = 0. \quad (2.9.9)$$

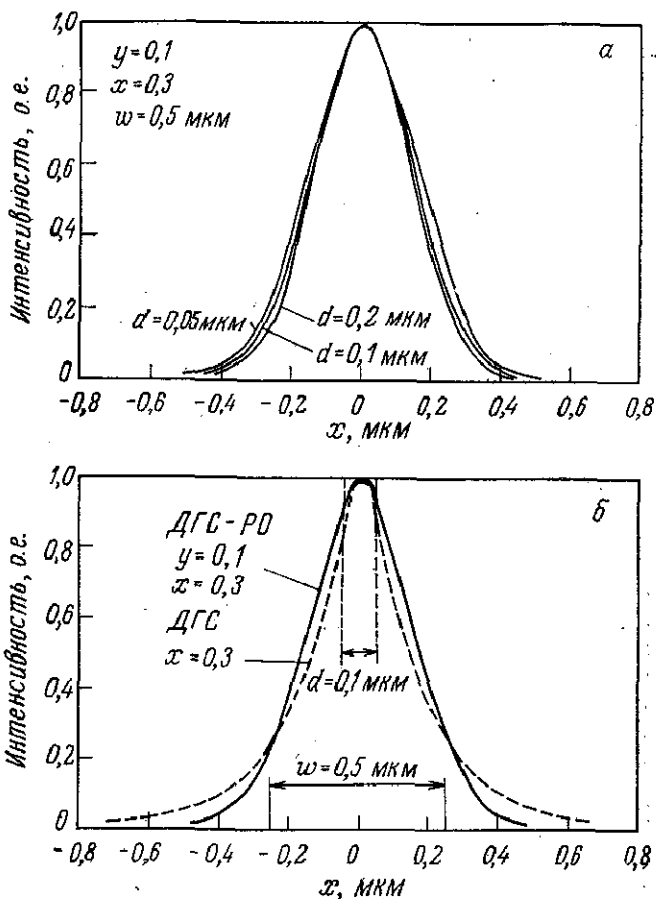


Рис. 2.9.2. Квадрат вектора электрического поля в гетеролазерах как функция расстояния от центра активного слоя в направлении, перпендикулярном плоскости $p-n$ -перехода. Энергия фотонов лазерного излучения принята равной 1,380 эВ. a — распределение интенсивности при разных толщинах активного слоя d и фиксированном значении w ; b — сравнение распределений интенсивности в ДГС — РО- и ДГС-лазерах при одинаковой толщине активного слоя и при молярных долях AlAs, указанных на рисунке [70].

Решение существует только в том случае, если детерминант системы равен нулю. Из этого условия получаем следующее дисперсионное уравнение:

$$\begin{aligned} & \kappa_2 \cos \left\{ \kappa_2 \left[\frac{1}{2} (\omega - d) \right] \right\} \{ \kappa_3 \sin (\kappa_3 d / 2) - \gamma \cos (\kappa_3 d / 2) \} + \\ & + \sin \left\{ \kappa_2 \left[\frac{1}{2} (\omega - d) \right] \right\} \{ \kappa_2^2 \cos (\kappa_3 d / 2) + \kappa_3 \gamma \sin (\kappa_3 d / 2) \} = 0. \quad (2.9.10) \end{aligned}$$

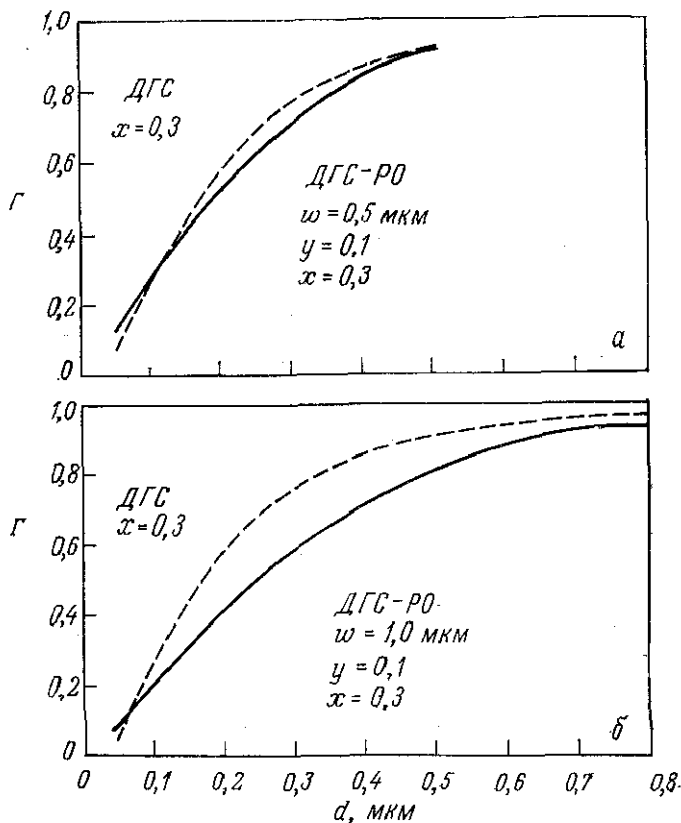


Рис. 2.9.3. Доля переносимой модой энергии Γ , приходящаяся на активный слой, как функция толщины активного слоя [70]. а — для ДГС — РО-лазера, $\omega = 0,5$ мкм; б — для ДГС — РО-лазера, $\omega = 1,0$ мкм.

В этом уравнении три параметра κ_3 , κ_2 и γ при заданных значениях показателей преломления и фиксированной длине волны зависят только от постоянной распространения β , как это следует из соотношений (2.9.1), (2.9.3) и (2.9.5). Поэтому уравнение (2.9.10) служит для определения β .

Рис. 2.9.2 иллюстрирует влияние различных параметров на рассчитанные кривые распределения интенсивности $|\mathcal{E}_y|^2$. Рис. 2.9.2, а демонстрирует воздействие толщины активного слоя на распределение оптической интенсивности. Из этого рисунка видно, что при заданных x и y распределение интенсивности определяется главным образом толщиной оптического резонатора ω . На рис. 2.9.2, б сравниваются распределения интенсивности для ДГС — РО-лазера с $y = 0,1$ и $x = 0,3$ и для ДГС-ла-

зера с $x = 0,3$. Обе структуры имеют одинаковую толщину активной области, равную $0,1$ мкм; для ДГС — РО-лазера $w = 0,5$ мкм. Как видно из этого рисунка, кривая распределения интенсивности для ДГС — РО-лазера при малых расстояниях x от центра активной области характеризуется большей шириной, чем кривая для ДГС-лазера, однако при возрастании x она быстрее стремится к нулю.

На рис. 2.9.3 проводится сравнение коэффициентов оптического ограничения для ДГС — РО- и ДГС-лазеров при двух значениях w , равных $0,5$ и $1,0$ мкм. При $w = 0,5$ мкм величина Γ для ДГС — РО-лазера при $d < 0,12$ мкм больше, а при $d > 0,12$ мкм немного меньше, чем для ДГС-лазера (рис. 2.9.3, а). При возрастании w до $1,0$ мкм точка пересечения кривых, характеризующих Γ в ДГС — РО- и ДГС-лазерах, сдвигается в сторону меньших d (рис. 2.9.3, б). Такое поведение согласуется с ожидаемым, и следует считать, что оно характерно для этих кривых. Точное значение d , при котором коэффициент оптического ограничения в ДГС — РО- и ДГС-лазерах имеет одинаковую величину, зависит также от таких факторов, как фактическое содержание Al в слое и отклонение от симметрии в толщинах слоев в реальном приборе.

Структура с раздельным ограничением была использована в качестве волновода в лазерах с распределенной обратной связью путем создания в ней гофра на границе раздела между слоями 4 и 5 на рис. 2.9.1. Эта структура используется потому, что при комнатной температуре создание гофра на границе активного слоя ведет к чрезмерно сильной безызлучательной рекомбинации. Распределенная обратная связь в гетеролазерах обсуждается в следующем параграфе. Использование ДГС — РО в приборах полезно также с точки зрения достижения низкой плотности порогового тока и малой расходимости луча при больших мощностях излучения. Эти свойства симметричных и асимметричных ДГС — РО-лазеров обсуждаются в § 5 гл. 7.

Лазер на четырехслойной гетероструктуре

Следующим более сложным типом диэлектрического плоского волновода является четырехслойный волновод. Лазеры на четырехслойной гетероструктуре готовились путем создания $p-n$ -перехода в слое GaAs, заключенном между слоями $P-Al_xGa_{1-x}As$ и $N-Al_xGa_{1-x}As$ [71]. Такая структура называется гетероструктурой с расширенным волноводом (РВ) [71—73]; она показана на рис. 2.9.4. Как видно на этом рисунке, GaAs-волновод разделяется на тонкую активную область p -типа и широкую пассивную область n -типа. Преимущество этой структуры состоит в том, что в ней возрастание толщины вол-

новода не сопровождается таким большим возрастанием плотности порогового тока, как в ДГС-лазере. Катастрофическая деградация (§ 2 гл. 8) при больших импульсных мощностях в структурах с РВ уменьшается благодаря тому, что полная мощность распределяется по большой площади. Паоли и др. [74] описали аналогичную четырехслойную гетероструктуру, работавшую на основной поперечной моде при толщинах оптических волноводов 2 мкм. Эта структура была названа $PpnN$ -гетероструктурой, где N и P соответствуют слоям $Al_xGa_{1-x}As$ N - и P -типов, а n и p — слоям $GaAs$ n - и p -типов.

Был проведен ряд исследований четырехслойных волноводов [27, 65, 75—77]. Дисперсионное уравнение получается при сшивании полей на границах раздела так же, как и в случае ДГС — РО-лазера. В граничных слоях $Al_xGa_{1-x}As$ поля экспоненциально затухают, в то время как во внутренних двух слоях они представляются в виде суммы синусоидального и косинусоидального членов [27, 75]. Из граничных условий получают однородную систему линейных уравнений для коэффициентов, подобно тому как была получена система уравнений (2.9.6) — (2.9.9) для симметричного ДГС — РО-лазера. Однако теперь в эту систему входит большее число уравнений (шесть), содержащих больше членов, а получающееся в результате дисперсионное уравнение значительно более сложно. Дисперсионное

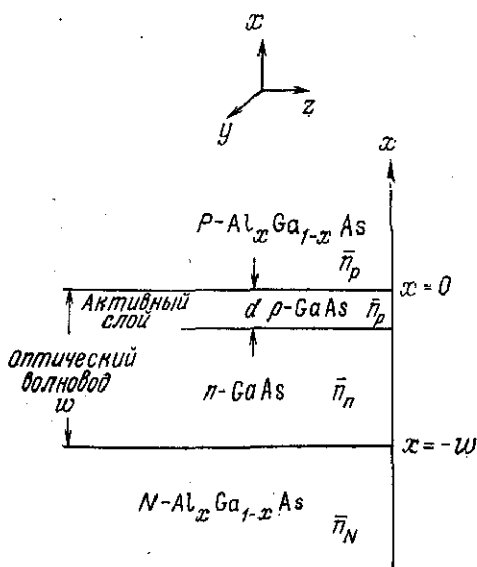


Рис. 2.9.4. Представление ДГС-лазера с РВ как четырехслойного плоского диэлектрического волновода.

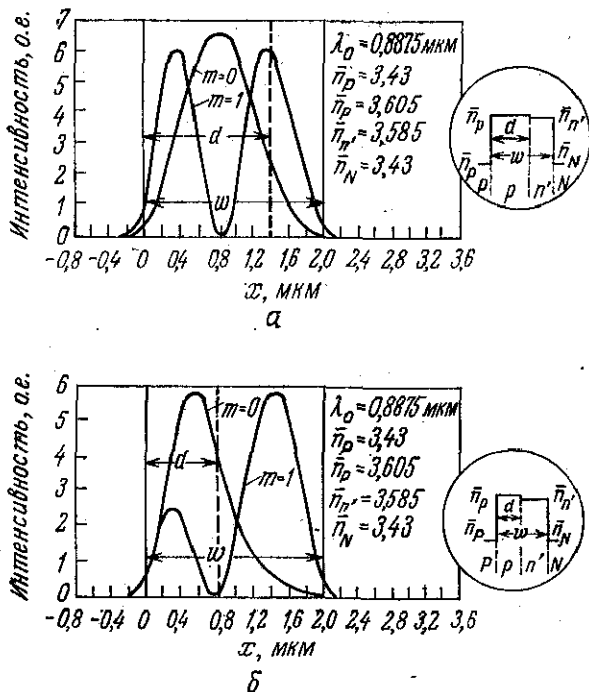


Рис. 2.9.5. Влияние толщины активного слоя на форму кривой распределения интенсивности в основной моде и в моде первого порядка для $Ppn'N$ -структуры при $\Delta\bar{n} = 0,02$. а — $w = 2,0$ мкм и $d = 1,4$ мкм; б — $w = 2,0$ мкм и $d = 0,8$ мкм [65].

уравнение для четырехслойного волновода приведено в работе [27].

Обсуждение характерных решений для полей в $PpnN$ -структуре или структуре с РВ затруднено вследствие неопределенности скачка показателя преломления на $n-p$ -границе. Анализ более определенной структуры в которой достигается лучшая селекция мод, был дан Крупкой [65]. В этой структуре небольшое количество Al вводится в n -слой ($x \approx 0,03$), обозначаемый теперь как n' , для того чтобы получить определенное значение $\Delta\bar{n} = \bar{n}_p - \bar{n}_{n'}$. (Индекс y показателя преломления соответствует слою, к которому этот показатель относится.) На рис. 2.9.5 приведены два примера распределения интенсивности в основной моде и в моде первого порядка в такой структуре для случая $\Delta\bar{n} = 0,02$. Эта структура обозначается как $Ppn'N$ -гетеролазер; на рис. 2.9.5 она показана в кружке. Толщина активного слоя p -GaAs равна d , а толщина оптического резонатора, являющаяся суммой толщин p - и n' -областей, равна w . Рис. 2.9.5, а

соответствует значениям $d = 1,4$ мкм и $w = 2,0$ мкм. В этом случае два пика моды первого порядка имеют одинаковую интенсивность. Однако когда d уменьшается до 0,8 мкм, интенсивность одного пика моды первого порядка, приходящегося на активный слой, составляет только половину интенсивности другого (рис. 2.9.5, б). Этот рисунок показывает, как выбор отношения d/w можно использовать для селекции мод.

Изменение распределения интенсивности излучения по модам может быть рассмотрено также с точки зрения коэффициента оптического ограничения. На рис. 2.9.6 для $\Delta\bar{n} = 0,02$ приведены кривые, определяющие Γ как функцию d при $w = 1, 2$ и 3 мкм для основной моды и для моды первого порядка [78]. На рис. 2.9.7 представлены аналогичные кривые для $\Delta\bar{n} = 0,04$ [78]. Эти рисунки показывают, что существуют обширные области значений d , в которых коэффициент оптического ограничения для основной моды больше, чем для моды первого порядка. Зависимость порогового тока от Γ обсуждается в § 8

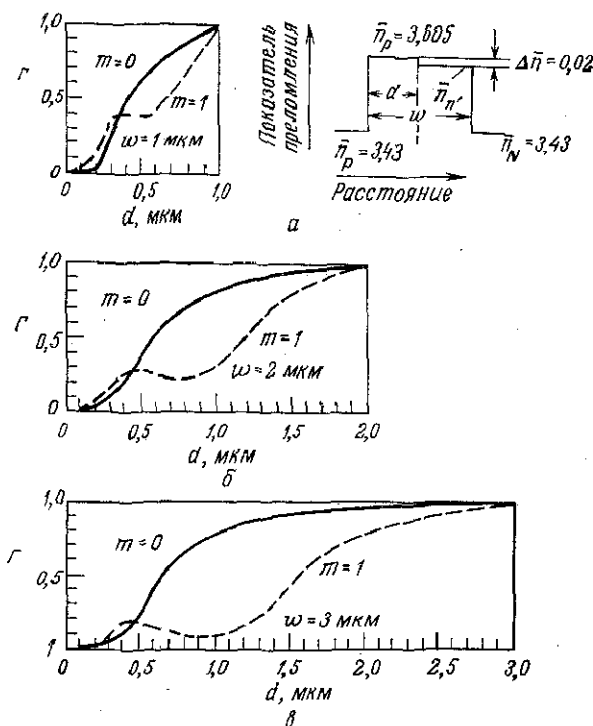


Рис. 2.9.6. Коэффициент оптического ограничения для основной моды и моды первого порядка как функция толщины активного слоя для Ppr'N-структуры с $\Delta\bar{n} = 0,02$ и $\lambda_0 = 0,8875$ мкм. Значения показателей преломления приведены справа на рис. а. а — $w = 1$ мкм; б — $w = 2$ мкм; в — $w = 3$ мкм [78].

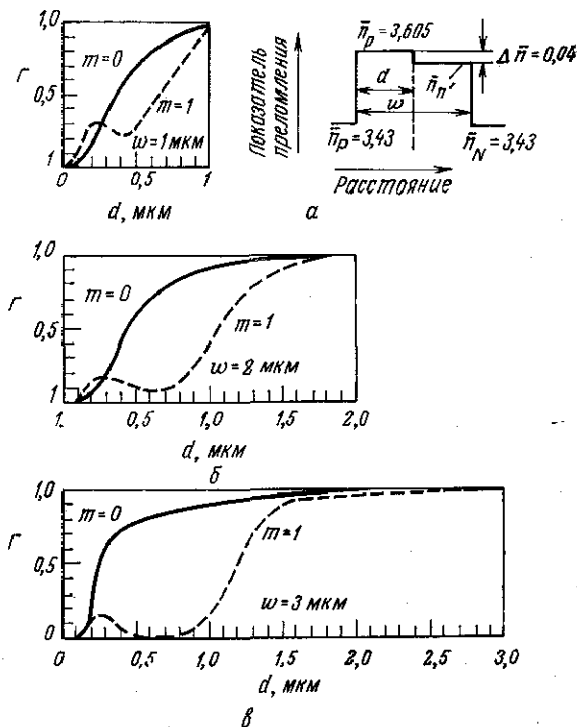


Рис. 2.9.7. Коэффициент оптического ограничения для основной моды и моды первого порядка как функция толщины активного слоя для $PpN'N$ -структуры с $\Delta n = 0,04$ и $\lambda_0 = 0,8875$ мкм. Значения показателей преломления приведены справа на рис. a . a — $w = 1$ мкм; $б$ — $w = 2$ мкм; $в$ — $w = 3$ мкм [78].

гл. 3. Следует ожидать, что мода с наибольшим значением Γ будет доминирующей лазерной модой. Коэффициенты отражения на торцевых гранях для основной моды и для моды первого порядка в $PpN'N$ -структурах, изображенных на рис. 2.9.6 и 2.9.7, имеют примерно одинаковые значения и не обеспечивают селекции мод [65]. Дальнейшее обсуждение лазеров с четырехслойными волноводами дано в § 5 гл. 7.

§ 10. ЛАЗЕРЫ С РАСПРЕДЕЛЕННОЙ ОБРАТНОЙ СВЯЗЬЮ

Вводные замечания

В § 8 настоящей главы рассматривался коэффициент отражения на образованных скалыванием или полировкой торцевых гранях резонатора. За счет этого отражения создается обратная

связь, необходимая для генерации лазерного излучения. Обратная связь может быть также получена посредством периодического изменения показателя преломления внутри оптического волновода, которое обычно создается гофрированием границы раздела между двумя диэлектрическими слоями. Лазеры, в которых обратная связь образуется за счет такой гофрированной структуры, называются лазерами с распределенной обратной связью (РОС) или с распределенным брэгговским отражателем (РБО). Эти гетеролазеры могут с успехом применяться в интегральной оптике в качестве источников излучения. Основные принципы их работы и конструирования были установлены еще до того, как технология выращивания гетероструктур была развита до такого высокого уровня, который необходим для практического создания этих приборов [79—81]. В первом лазере с РОС использовался краситель в матрице желатина на стеклянной подложке [79]. После большой предварительной работы, проведенной при низких температурах, в 1975 г. была получена генерация лазерного излучения в гетеролазерах с РОС [82, 83] и РБО [84] при комнатной температуре.

При анализе работы лазера с периодической структурой были использованы три подхода. Когельник и Шэнк [80] использовали метод связанных волн, в котором рассматривается распространение плоских волн в периодической структуре и получают простые выражения, определяющие резонансные длины волн и пороговые условия. Ярив [85] также использовал формализм связанных волн для анализа распространения волны в периодическом волноводе и получил выражение для коэффициента связи, весьма полезное при рассмотрении гетеролазеров. Де Вэймс и Холл [86], Кодиро и Вэнг [87], а также Вэнг [88] дали анализ распространения волн в периодических структурах, используя блоховские волны и предполагая, что важную роль играют только два луча. В этом параграфе мы будем следовать теории связанных волн, развитой Когельником и Шэнком [80], поскольку этот подход проще всего обосновать, исходя из полученных в предыдущем параграфе уравнений. Тем не менее все три указанных выше способа рассмотрения эквивалентны. Различие между методом связанных волн и подходом, в котором используются блоховские волны, можно понять, если проанализировать выражения, которые выводятся в следующем разделе этого параграфа.

Уравнения для связанных волн

В этом разделе мы рассмотрим селективность по длинам волн в структурах с РОС, определим резонансные длины волн и усиление на этих длинах волн на пороге генерации. Следуя

Когельнику и Шэнку [80], мы рассмотрим распространение плоской волны в периодической структуре, а применение развитой теории к гетероструктурам отложим до конца параграфа. Эта линейная теория для модели связанных волн применима только вблизи порога. Волновое уравнение в среде, которая может обладать потерями или усилением, записывалось нами в виде (2.2.42)

$$\partial^2 \mathcal{E}_x(z) / \partial z^2 - \Gamma^2 \mathcal{E}_x(z) = 0,$$

где Γ^2 определяется формулой, вытекающей из соотношений (2.2.38), (2.2.49) и (2.2.50):

$$-\Gamma^2 = (\bar{n}^2 - \bar{k}^2 - 2j\bar{n}\bar{k}) (2\pi/\lambda_0)^2. \quad (2.10.1)$$

Предположим, что показатель преломления имеет синусоидальную зависимость от z (рис. 2.10.1, *a*) и описывается выражением

$$\bar{n}(z) = \bar{n} + \bar{n}_a \cos Kz. \quad (2.10.2)$$

Здесь

$$K = 2\pi/\Lambda, \quad (2.10.3)$$

где Λ — период пространственных колебаний показателя преломления, а \bar{n}_a — амплитуда колебаний. Из условия Брэгга [89] следует

$$\Lambda = \lambda_b / 2\bar{n}, \quad (2.10.4)$$

где λ_b — длина волны в свободном пространстве, удовлетворяющая условию Брэгга. Можно предположить, что существует также пространственное изменение усиления или поглощения [80]. Однако в рассматриваемых здесь случаях усиление берется постоянным и рассматривается только пространственное изменение показателя преломления. В публикациях, касающихся структур с РОС, рассматривалось усиление амплитуды поля, а не усиление интенсивности. Поскольку $P \propto \mathcal{E}\mathcal{E}^*$, или $P_0 \exp(-\alpha z) \propto \mathcal{E}_0^2 \exp(-2\alpha_F z)$, для коэффициента поглощения или коэффициента усиления по амплитуде поля будет использоваться символ α_F в отличие от введенного в § 2 настоящей главы коэффициента α (т. е. $\alpha = 2\alpha_F$). Коэффициент экстинкции, согласно соотношению (2.2.61), теперь определяется формулой

$$\bar{k}_F = \alpha_F \lambda_0 / 2\pi. \quad (2.10.5)$$

Величина α_F считается положительной для поглощения и отрицательной для усиления.

Подстановка выражений (2.10.2) и (2.10.5) в формулу (2.10.1) при α_F , соответствующем усилению, дает

$$-\Gamma^2 = (2\pi/\lambda_0)^2 \{ [\bar{n} + \bar{n}_a \cos(Kz)]^2 - (\alpha_F \lambda_0 / 2\pi)^2 - 2j[\bar{n} + \bar{n}_a \cos(Kz)](\alpha_F \lambda_0 / 2\pi) \}. \quad (2.10.6)$$

Для плоской волны из соотношения (2.2.39) следует $\beta = \bar{n}2\pi/\lambda_0$, так что имеет место неравенство

$$\beta \gg \alpha_F. \quad (2.10.7)$$

Рассматривая малые изменения показателя преломления,

$$\bar{n} \gg \bar{n}_a, \quad (2.10.8)$$

и учитывая неравенство (2.10.7), формулу (2.10.6) можно представить в следующем виде¹⁾:

$$-\Gamma^2 = \beta^2 - 2j\beta\alpha_F + (4\pi/\lambda_0)\beta\bar{n}_a \cos(Kz). \quad (2.10.9)$$

Тогда уравнение (2.2.42) может быть переписано в виде

$$[\partial^2 \mathcal{E}_x(z)/\partial z^2] + (\beta^2 - 2j\beta\alpha_F) \mathcal{E}_x(z) = -(4\pi/\lambda_0)\beta\bar{n}_a \cos(2\beta_b z) \mathcal{E}_x(z), \quad (2.10.10)$$

где β_b определяется из формул (2.10.3) и (2.10.4) следующим образом:

$$K = 4\pi\bar{n}/\lambda_b = 2\beta_b. \quad (2.10.11)$$

Периодическое возмущение \bar{n} может давать бесконечное число дифракционных порядков. Однако вблизи длин волн, удовлетворяющих условию Брэгга, только два порядка имеют значительную амплитуду и удовлетворяют условию фазового синхронизма [80]. В рассматриваемой модели связанных волн пренебрегается всеми порядками дифракции, кроме этих двух, представляющих собой распространяющиеся навстречу друг другу волны $R(z)\exp(-j\beta_b z)$ и $S(z)\exp(j\beta_b z)$ [80]. Из-за наличия в среде усиления эти волны при распространении по диэлектрическому слою усиливаются, и благодаря брэгговскому рассеянию между ними происходит обмен энергией. Таким образом, полное электрическое поле внутри диэлектрического слоя с периодической структурой представляется в виде суммы двух волн:

$$\mathcal{E}_x(z) = R(z)\exp(-j\beta_b z) + S(z)\exp(j\beta_b z). \quad (2.10.12)$$

где $R(z)$ и $S(z)$ — комплексные амплитуды. Определение поля $\mathcal{E}_x(z)$ по формуле (2.10.12) отличает формализм связанных мод, развитый Когельником и Шэнком [80], от метода блоховских волн [86—88]. Различие этих методов мы обсудим в конце параграфа. Согласно неравенствам (2.10.7) и (2.10.8), обмен энергией между встречными волнами происходит медленно, и

¹⁾ В работе [80] используется временная зависимость вида $\exp(-j\omega t)$, так что $-\Gamma^2 = (\bar{n}^2 - k^2 + 2j\bar{n}k)k_0^2$ и формула (2.10.9) имеет вид $-\Gamma^2 = \beta^2 + 2j\beta\alpha_F + (4\pi/\lambda_0)\beta\bar{n}_a \cos Kz$. Это изменение знака сохраняется в последующих выражениях.

вторыми производными $\partial^2 R/\partial z^2$ и $\partial^2 S/\partial z^2$ можно пренебречь [80]. В этом случае подстановка выражения (2.10.12) в волновое уравнение (2.10.10) и группировка членов, содержащих множители $\exp(-j\beta_b z)$ или $\exp(j\beta_b z)$, дают два уравнения для связанных волн

$$-\frac{\partial R}{\partial z} - \left[\alpha_F + j \left(\frac{\beta^2 - \beta_b^2}{2\beta_b} \right) \right] R = j(\pi \bar{n}_a / \lambda_0) S, \quad (2.10.13)$$

$$\frac{\partial S}{\partial z} - \left[\alpha_F + j \left(\frac{\beta^2 - \beta_b^2}{2\beta_b} \right) \right] S = j(\pi \bar{n}_a / \lambda_0) R. \quad (2.10.14)$$

При выводе этих уравнений отношение β/β_b заменялось на единицу, поскольку рассматриваются длины волн, мало отличающихся от λ_b . Величина $(\beta^2 - \beta_b^2)/2\beta_b$ в уравнениях для связанных волн обычно записывается в виде

$$\delta = (\beta^2 - \beta_b^2)/2\beta_b = (\beta - \beta_b)(\beta + \beta_b)/2\beta_b \approx \beta - \beta_b. \quad (2.10.15)$$

Множитель в правой части уравнений (2.10.13) и (2.10.14) был назван постоянной (коэффициентом) связи κ_c :

$$\kappa_c = \pi \bar{n}_a / \lambda_0. \quad (2.10.16)$$

Это выражение для постоянной связи справедливо для приближения плоских волн и синусоидального изменения показателя преломления. Для гетероструктур с гофром в волноводе коэффициенты связи могут быть определены из выражений, полученных Яривом [85]. Они будут рассмотрены в конце параграфа. С учетом формул (2.10.15) и (2.10.16) получаем следующую форму уравнений для связанных волн [80]:

$$-R' - (\alpha_F + j\delta)R = j\kappa_c S, \quad (2.10.17)$$

$$S' - (\alpha_F + j\delta)S = j\kappa_c R. \quad (2.10.18)$$

Решение уравнений для связанных волн

Уравнения для связанных волн можно решить, используя подстановку $\exp(\gamma_c z)$ для неизвестных $R(z)$ и $S(z)$. Эта подстановка приводит к системе вспомогательных уравнений

$$-\gamma_c R - (\alpha_F + j\delta)R = j\kappa_c S, \quad (2.10.19)$$

$$\gamma_c S - (\alpha_F + j\delta)S = j\kappa_c R. \quad (2.10.20)$$

Выражая S из уравнения (2.10.20) и подставляя полученное выражение в уравнение (2.10.19), получим формулу, определяющую комплексную постоянную распространения γ_c :

$$\gamma_c^2 = \kappa_c^2 + (\alpha_F + j\delta)^2. \quad (2.10.21)$$

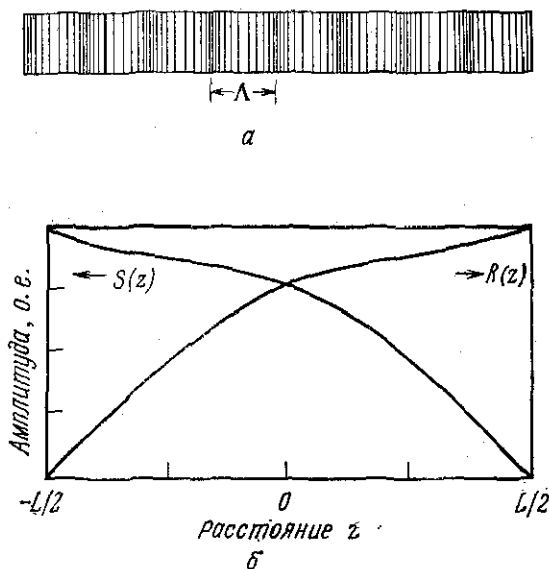


Рис. 2.10.1. Пространственное изменение показателя преломления, описываемое формулой $\bar{n}(z) = \bar{n}_0 + \bar{n}_a \cos [(2\pi/\Lambda)z]$ [79] (а). Зависимость амплитуд распространяющихся влево (S) и вправо (R) волн от расстояния (б).

Таким образом, общие решения могут быть представлены в виде

$$R(z) = r_1 \exp(\gamma_c z) + r_2 \exp(-\gamma_c z), \quad (2.10.22)$$

$$S(z) = s_1 \exp(\gamma_c z) + s_2 \exp(-\gamma_c z). \quad (2.10.23)$$

В изображенной на рис. 2.10.1, а структуре волны начинают распространяться с нулевыми амплитудами в конечных точках $z = \pm L/2$ и обмениваются энергией друг с другом вследствие периодического изменения $\bar{n}(z)$. Граничными условиями являются

$$R(-L/2) = S(L/2) = 0. \quad (2.10.24)$$

Случай ненулевой отражательной способности зеркал был рассмотрен Чинном [90] и Стрейфером и др. [91]. Граничные условия (2.10.24) дают

$$r_1/r_2 = s_2/s_1 = -\exp(\gamma_c L). \quad (2.10.25)$$

Используя эти равенства, получаем

$$R(z) = 2r_1 \exp(-\gamma_c L/2) \operatorname{sh}[\gamma_c(z + L/2)], \quad (2.10.26)$$

$$S(z) = -2s_2 \exp(-\gamma_c L/2) \operatorname{sh}[\gamma_c(z - L/2)]. \quad (2.10.27)$$

Симметричные [$\mathcal{E}(-z) = \mathcal{E}(z)$] и антисимметричные [$\mathcal{E}(-z) = -\mathcal{E}(z)$] решения вследствие симметрии структуры подразумевают выполнение дополнительных соотношений [80]

$$r_1 = \pm s_2 \quad (2.10.28)$$

и, учитывая равенство (2.10.25),

$$r_2 = \pm s_1. \quad (2.10.29)$$

Множители перед гиперболическими синусами в выражениях (2.10.26) и (2.10.27) можно положить равными единице и записать продольное распределение поля мод в структуре с РОС в следующей форме [80]:

$$R(z) = \text{sh}[\gamma_c(z + L/2)], \quad (2.10.30)$$

$$S(z) = \pm \text{sh}[\gamma_c(z - L/2)]. \quad (2.10.31)$$

Эти решения для волн $R(z)$ и $S(z)$, распространяющихся в противоположных направлениях, иллюстрируются на рис. 2.10.1, б.

Распределения поля являются модами, соответствующими дискретным собственным значениям γ_c . Эти собственные значения зависят от длины структуры L и постоянной связи κ_c . Для того чтобы получить дисперсионное уравнение, надо подставить выражения (2.10.30) и (2.10.31) в уравнения (2.10.17) и (2.10.18), а затем взять сумму и разность полученных уравнений. В результате этих операций получим

$$-\gamma_c \text{sh}(\gamma_c L/2) - (\alpha_F + j\delta) \text{ch}(\gamma_c L/2) = \pm j\kappa_c \text{ch}(\gamma_c L/2), \quad (2.10.32)$$

$$-\gamma_c \text{ch}(\gamma_c L/2) - (\alpha_F + j\delta) \text{sh}(\gamma_c L/2) = \mp j\kappa_c \text{sh}(\gamma_c L/2). \quad (2.10.33)$$

При выводе уравнений (2.10.32) и (2.10.33) было сделано сокращение на общие множители $\text{sh}(\gamma_c z)$ и $\text{ch}(\gamma_c z)$ соответственно. Сложение и вычитание уравнений (2.10.32) и (2.10.33) дает

$$\gamma_c + (\alpha_F + j\delta) = \mp j\kappa_c \exp(-\gamma_c L), \quad (2.10.34)$$

$$\gamma_c - (\alpha_F + j\delta) = \pm j\kappa_c \exp(\gamma_c L). \quad (2.10.35)$$

Дисперсионное уравнение, выраженное через структурные параметры κ_c и L , получается в результате сложения уравнений (2.10.34) и (2.10.35) и представляет собой трансцендентное уравнение на собственные значения величины γ_c :

$$\kappa_c = \pm j\gamma_c / \text{sh}(\gamma_c L). \quad (2.10.36)$$

Каждому значению γ_c соответствуют определенные длина волны и усиление на пороге генерации. Эти величины можно получить, вычитая уравнение (2.10.34) из уравнения (2.10.35). В результате получим

$$(\alpha_F + j\delta) = \mp j\kappa_c \text{ch}(\gamma_c L) = \gamma_c \text{cth}(\gamma_c L). \quad (2.10.37)$$

Таким образом, для каждого значения γ_c , определяемого из уравнения (2.10.36), существуют соответствующие значения коэффициента усиления α_F и резонансной длины волны, представленной через величину δ . Кроме того, каждое значение γ_c определяет продольное распределение поля моды в соответствии с формулами (2.10.30) и (2.10.31). Приближенные решения полученных уравнений, характеризующие пороговое усиление и спектр мод, будут даны ниже.

Спектр мод и усиление на пороге

Для случая $\alpha_F \gg \kappa_c$ из соотношения (2.10.21) получаем приближенное выражение для комплексной постоянной распространения

$$\gamma_c \approx (\alpha_F + j\delta). \quad (2.10.38)$$

Тогда дисперсионное уравнение (2.10.34) может быть переписано в виде

$$2(\alpha_F + j\delta) \approx \mp j\kappa_c \exp[-(\alpha_F + j\delta)L]. \quad (2.10.39)$$

Уравнение (2.10.39) непосредственно связывает коэффициент усиления на пороге генерации $-\alpha_F = g_F$ со структурными параметрами κ_c и L лазера с РОС. Это условие генерации в лазере с РОС может быть записано в более привычной форме

$$1 \approx [\mp j\kappa_c/2(\alpha_F + j\delta)] \exp[-(\alpha_F + j\delta)L], \quad (2.10.40)$$

которая аналогична приведенному в следующей главе условию (3.8.5) для системы с параллельными отражающими поверхностями. Для лазера с РОС пороговое условие может быть получено из соотношения (2.10.40), если приравнять абсолютную величину и фазу правой части единице и нулю соответственно:

$$[\kappa_c^2/4(\alpha_F^2 + \delta^2)] \exp(-2\alpha_F L) \approx 1, \quad (2.10.41)$$

$$\mp (q + 1/2)\pi - \arctg(\delta/\alpha_F) - \delta L = 0, \quad (2.10.42)$$

где $q = 0, \pm 1, \pm 2, \dots$. Сравнение формул (2.10.41) и (3.8.6) показывает, что $-2\alpha_F$ соответствует величине $(g - \alpha_i)$ в формуле (3.8.6), а $[\kappa_c^2/4(\alpha_F^2 + \delta^2)]$ — произведению коэффициентов отражения зеркал.

Вблизи длины волны, удовлетворяющей условию Брэгга, $\delta = 0$ и поэтому $-\alpha_F \gg \delta$. Тогда с учетом равенства (2.10.16) выражения (2.10.41) и (2.10.42) принимают вид

$$(\pi\bar{n}_a/\lambda_0)^2 \exp(-2\alpha_F L) \approx 4\alpha_F^2, \quad (2.10.43)$$

$$\delta = \pm (q + 1/2)\pi/L. \quad (2.10.44)$$

Используя формулу (2.10.15), определяющую δ , и полагая $\lambda_0/\lambda_b \approx 1$, из выражения (2.10.44) получаем формулы, определяющие резонансные длины волн:

$$\delta L = 2\pi\bar{n} [(\lambda_b - \lambda_0)/\lambda_b^2] L = \pm (q + 1/2)\pi, \quad (2.10.45)$$

или

$$\lambda_0 = \lambda_b \pm [(q + 1/2)\lambda_b^2/2\bar{n}L]. \quad (2.10.46)$$

Полученный результат показывает, что резонансные длины волн располагаются с интервалом $\lambda_b^2/2\bar{n}L$. Аналогичный результат получается из выражения (3.8.11) для случая параллельных зеркал. Из соотношения (2.10.46) следует также, что λ_0 никогда не равно λ_b при любых целых q . Поэтому в такой структуре с РОС резонанс никогда не наблюдается на длине волны, удовлетворяющей условию Брэгга. Используя выражение (2.10.46) для λ_0 , из (2.10.43) можно получить выражение для коэффициента усиления $g_F = -\alpha_F$.

Полученные выше выражения демонстрируют также свойства спектральной селективности структур с РОС. Из соотношения (2.10.41) следует, что при таком отклонении длины волны от брэгговской, что $\delta = \alpha_F$, величина κ_c^2 должна быть удвоена, чтобы значение коэффициента усиления на пороге осталось тем же. Из соотношения (2.10.45) при $\delta = -\alpha_F = g_F$ имеем

$$\delta = -\alpha_F \approx (2\pi\bar{n}/\lambda_b^2)(\lambda_b - \lambda_0), \quad (2.10.47)$$

и превышение над порогом будет существовать в спектральной полосе шириной, приблизительно определяемой формулой

$$(\Delta\lambda/\lambda_b) \approx (\lambda_b/4\pi\bar{n}L) \ln G, \quad (2.10.48)$$

где введено обозначение

$$G = \exp(-2\alpha_F L). \quad (2.10.49)$$

Таким образом, при фиксированном значении G спектральная селективность обратно пропорциональна длине гофра. Такое поведение естественно, поскольку при увеличении длины увеличивается число ребер гофра.

Когельник и Шэнк [80] получили ряд полезных результатов при численном решении на ЭВМ комплексного трансцендентного уравнения (2.10.36). На рис. 2.10.2 приведен рассчитанный спектр излучения. Расстояние между штриховыми прямыми соответствует интервалу $\lambda_b^2/2\bar{n}L$ между длинами волн, получающемуся из выражения (2.10.46) при предположении $-\alpha_F \gg \kappa_c$. Спектр излучения симметричен относительно брэгговской длины волны, так что появлению в спектре сразу двух линий соответствует одно и то же значение коэффициента усиления. Порог генерации возрастает при удалении длины волны излучения

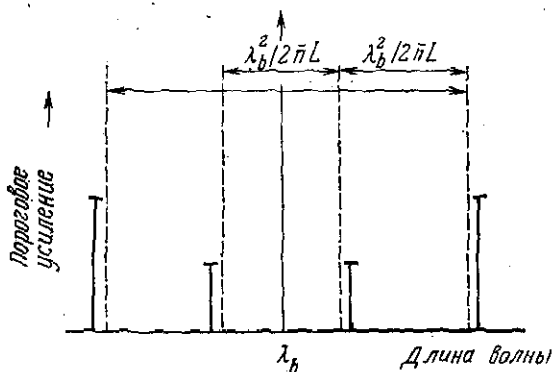


Рис. 2.10.2. Спектр мод и необходимое пороговое усиление для структуры с периодическим показателем преломления [80].

от значения λ_b , и тем самым обеспечивается спектральная селективность. Проведенные вычисления показали также, что, когда κ_c мало, резонансные длины волн находятся вблизи длин волн, определяемых формулой (2.10.46), и при увеличении κ_c усиление становится меньше, а длина волны первого резонанса удаляется от соответствующего значения, определяемого указанной формулой.

Несмотря на то что в проведенном выше анализе лазеров с РОС рассматривались бегущие волны, а не волноводные моды, основные идеи этого анализа справедливы также и для волноводных мод в гетеролазерах [92]. Примером может являться использование уравнений для связанных волн, описывающих перераспределение энергии между распространяющимися в противоположных направлениях волнами как результат периодического изменения показателя преломления. Было показано, что влияние такого изменения показателя преломления можно описать через постоянную связи. Выражение, определяющее постоянную связи для гетеролазера с РОС, отличается от полученного выше при рассмотрении бегущей волны. Различны в этих случаях и постоянные распространения. Разделение мод и спектральная селекция являются важными свойствами также и для гетеролазера с РОС.

Блоховские волны и волноводные моды

При использовании метода блоховских волн для анализа распространения волн в волноводах с периодической структурой решение волнового уравнения (2.2.42) записывается в виде [86—88]

$$\mathcal{E}_x(z) = A\phi_1(z) \exp(\Gamma z) + B\phi_2(z) \exp(-\Gamma z). \quad (2.10.50)$$

В этом выражении $\phi_1(z)$ и $\phi_2(z)$ — блоховские волны, распространяющиеся соответственно вперед и назад, каждая из них является периодической функцией с периодом Λ . Комплексная эффективная постоянная распространения обеих волн обозначена через Γ . Вблизи длин волн, удовлетворяющих условию Брэгга, каждая из этих двух блоховских волн может быть аппроксимирована суммой первых двух членов (основного и первой гармоники) в их разложениях в ряд Фурье. В результате такой процедуры $\mathcal{E}_x(z)$ будет представляться в виде суммы четырех слагаемых: распространяющихся вперед и назад первичных лучей, а также лучей, получающихся при обратном рассеянии первичных. Амплитуды рассеянных обратных лучей связываются с амплитудами соответствующих первичных лучей подстановкой $\mathcal{E}_x(z)$ в уравнение (2.10.10). В этом уравнении β и α_F полагаются равными соответственно постоянной распространения и коэффициенту усиления для волноводной моды в эквивалентном невозмущенном волноводе. Амплитуды двух первичных лучей связываются между собой через граничные условия на концах периодической структуры. Из граничных условий вытекает также дисперсионное уравнение для определения усиления на пороге генерации и резонансной длины волны. Рассмотрение данного вопроса с помощью метода блоховских волн отличается от анализа, проведенного Когельником и Шэнком, формой записи приближенного выражения для $\mathcal{E}_x(z)$. Поскольку оба подхода основаны на использовании уравнения (2.2.42) и дают для структуры с РОС одинаковые результаты, они различаются скорее по форме, чем по содержанию.

Теория связанных мод, развитая Яривом [85] применительно к исследованию волноводных мод, основывается на волновом уравнении (2.4.10) для невозмущенных мод. В этом случае периодическое изменение показателя преломления представляется как распределенный источник поляризации $P_{\text{возм}}(x, t)$. Вектор электрического смещения \mathcal{D} связан с вектором напряженности электрического поля соотношением (2.2.3). Это соотношение часто записывают в виде [93]

$$\mathcal{D} \equiv \varepsilon \mathcal{E} = \varepsilon_0 \mathcal{E} + \mathbf{p}, \quad (2.10.51)$$

где \mathbf{p} — вектор поляризации. Волновое уравнение при наличии возмущения записывается в виде [85]

$$\nabla^2 \mathcal{E}_y(\mathbf{r}, t) = \mu \varepsilon (\partial^2 \mathcal{E}_y / \partial t^2) + \mu (\partial^2 / \partial t^2) [\mathbf{p}_{\text{возм}}(\mathbf{r}, t)]. \quad (2.10.52)$$

В этом уравнении величины μ и ε соответствуют невозмущенному волноводу а возмущение, вызванное периодическим изменением показателя преломления, содержится в $\mathbf{p}_{\text{возм}}(\mathbf{r}, t)$. Таким образом, исходные уравнения теории связанных мод, описывающей распространение направляемых волн, и теории, раз-

витой на основе бегущих волн [80] или волн Блоха [86—88], различны. Для получения решения \mathcal{E}_y в волноводе с периодической структурой разлагают в ряд по ортогональным нормированным плоским волнам $\mathcal{E}_y^m(x)$, соответствующим случаю однородного волновода. Такое представление \mathcal{E}_y используется в волновом уравнении для того, чтобы получить решение в общем виде, не зависящем от конкретной формы возмущения. Этот метод позволяет представить постоянную связи для ТЕ-волн в виде [85]

$$\kappa_c = \frac{\omega \bar{n}_{\text{эфф}}^2 (-1)^l}{4\pi l} \int_{-\infty}^{\infty} \Delta \bar{n}^2(x) [\mathcal{E}_y^m(x)]^2 dx, \quad (2.10.53)$$

где $\bar{n}_{\text{эфф}} = \beta/k_0$ — эффективный показатель преломления волновода, $\Delta \bar{n}$ — амплитуда изменения показателя преломления, l — номер того члена в разложении рассматриваемой периодической структуры в ряд Фурье, который создает обратную связь, m — порядок моды однородного волновода. Номер l , обеспечивающий выполнение условия Брэгга, равен

$$l = 2\Lambda \bar{n}_{\text{эфф}}/\lambda_b. \quad (2.10.54)$$

Выражение (2.10.53) позволяет вычислить постоянную связи для заданного периодического профиля показателя преломления по невозмущенным полям $\mathcal{E}_y^m(x)$, описывающим соответствующие моды волновода.

Как уже обсуждалось в этой части параграфа, анализ блоховских волн и теория связанных мод дают решения волнового уравнения при помощи различных приближенных методов. Однако Ярив и Говер [94] сравнили эти две теории и показали, что формализм связанных мод и блоховский формализм эквивалентны. Имеются случаи, когда один подход более удобен, чем другой, поэтому полезно знать оба этих подхода.

Периодическое изменение показателя преломления в гетероструктурах

Вместо синусоидального возмущения показателя преломления, представленного на рис. 2.10.1, а, гетероструктуры могут изготавливаться с гофром (решеткой) на гетерогранице типа изображенного на рис. 2.10.3, а. Такое гофрирование дает периодическое изменение $\bar{n}(x, z)$ и приводит к тем же эффектам, что и рассмотренное выше синусоидальное возмущение. Решетки на поверхности с малыми периодами $\sim 1000 \text{ \AA}$ [95] создавались в пленке фоторезиста при помощи оптической интерференции лазерных пучков [95, 96]. Эти решетки могут быть перенесены

на диэлектрик ионным травлением [95—97] или химическим травлением [98—100].

Как было показано выше, волны, распространяющиеся в противоположных направлениях, возникают при выполнении условия Брэгга $2\Lambda = \lambda_0 / \bar{n}_{\text{эфф}}$, где $\bar{n}_{\text{эфф}} = \beta / k_0$. Полезно рассмотреть решетки высшего порядка, имеющие $\Lambda = l\lambda_0 / 2\bar{n}_{\text{эфф}}$, где l — порядок решетки. Граница раздела между материалами с \bar{n}_1 и \bar{n}_2 представлена вновь на рис. 2.10.3, б, где показаны лучи, рассеиваемые из волновода гофром с периодом Λ [101]. Для волны, распространяющейся направо, лучи, рассеиваемые последовательными выступами, находятся в фазе, если луч l

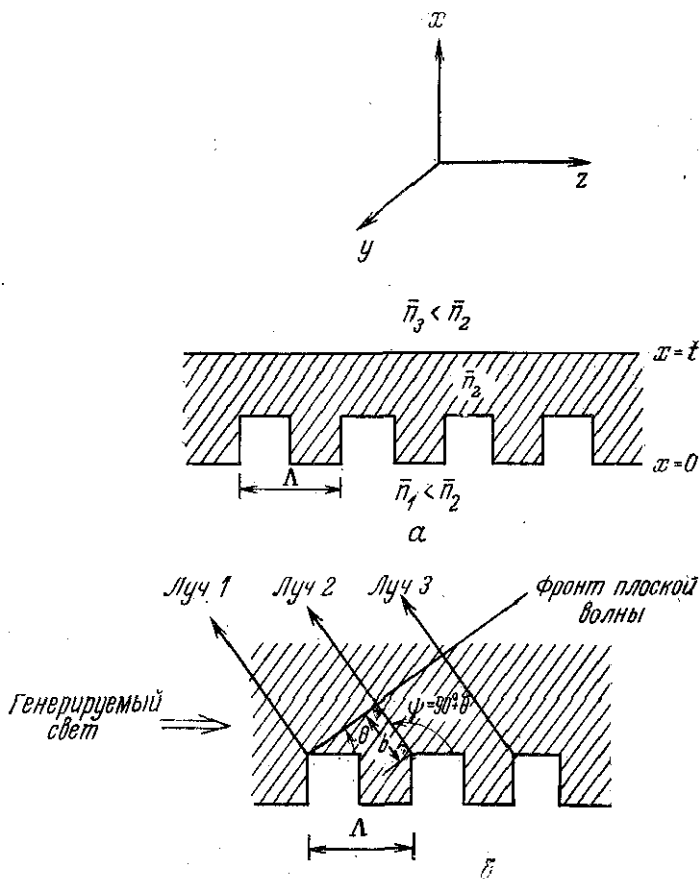


Рис. 2.10.3. Представление гофра с периодом Λ в волноводе, образованном в гетероструктуре (а). Лучи, рассеиваемые из волновода гофром (б); θ — угол между волновым фронтом и плоскостью p — n -перехода; ψ — угол между падающим и рассеянным лучами [101].

находится в фазе с лучом 2 и т. д. Это условие выполняется, если разность хода лучей $\Lambda + b$ равна целому числу, умноженному на длину волны в диэлектрике $\lambda_0/\bar{n}_{эфф}$, т. е. [101]

$$b + \Lambda = i(\lambda_0/\bar{n}_{эфф}), \quad (2.10.55)$$

где i — целое число, а b — расстояние, показанное на рис. 2.10.3, b и равно

$$b = \Lambda \sin \theta. \quad (2.10.56)$$

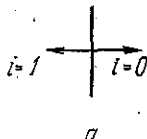
Уравнение (2.10.55) может быть переписано в виде

$$\sin \theta = i(\lambda_0/\Lambda \bar{n}_{эфф}) - 1 = (2i/l) - 1. \quad (2.10.57)$$

Выражение в правой части этого уравнения не может по модулю превосходить единицу, поскольку θ должно быть действительным.

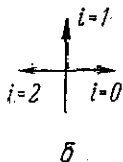
1-й порядок

$$\Lambda = \lambda_0 / 2 \bar{n}_{эфф}$$



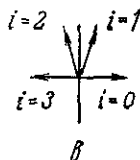
2-й порядок

$$\Lambda = \lambda_0 / \bar{n}_{эфф}$$



3-й порядок

$$\Lambda = 3\lambda_0 / 2 \bar{n}_{эфф}$$



4-й порядок

$$\Lambda = 2\lambda_0 / \bar{n}_{эфф}$$

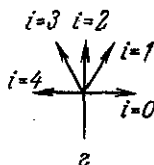


Рис. 2.10.4. Направления рассеяния волны в зависимости от порядка рассеивающей решетки.

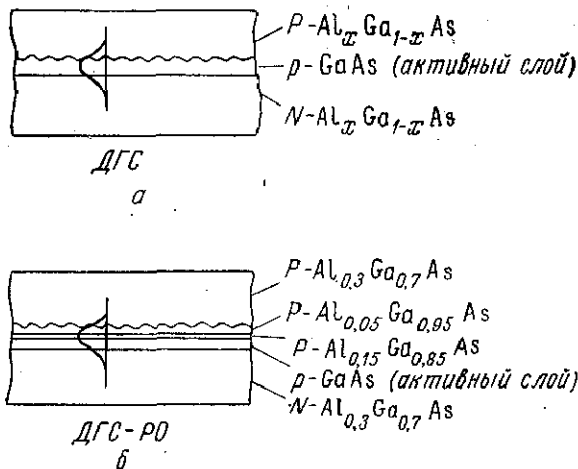


Рис. 2.10.5. Схематическое представление лазерного диода с РОС. Показано расположение гофра и распределение оптического поля. *a* — ДГС-лазер; *б* — ДГС — РО-лазер с $x = 0,3$, $y = 0,15$ и $z = 0,05$ [103].

Условие Брэгга для рассеяния в первом порядке дает $\Lambda = \lambda_0 / 2\bar{n}_{эфф}$, и из уравнения (2.10.57) следует

$$\sin \theta = 2i - 1, \quad (2.10.58)$$

где i должно быть равно 0 или 1, чтобы $2i - 1 \leq 1$. Решение с $i = 0$ дает $\theta = -90^\circ$, и для луча 2 получаем $\psi = 90^\circ + \theta = 0$, что соответствует рассеянию вперед. Для $i = 1$ угол θ равен

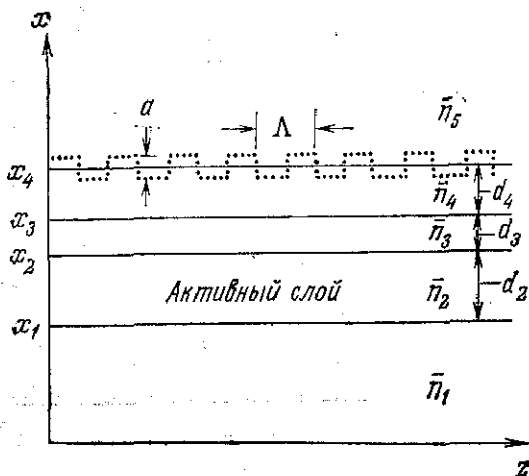


Рис. 2.10.6. Схематическое представление ДГС — РО-волновода, используемое при расчете $\bar{n}_{эфф}$ и κ_0 . Прямоугольный гофр имеет глубину a и период Λ [103].

90° , а $\psi = 180^\circ$, т. е. происходит рассеяние назад, чем и создается распределенная обратная связь. На рис. 2.10.4, а схематически изображены эти рассеянные лучи. Аналогичные результаты могут быть получены и для волны, распространяющейся в обратном направлении. Для решетки второго порядка ($l=2$) имеем $\Lambda = \lambda_0 / \bar{n}_{эфф}$, и в уравнении (2.10.57) $i-1 \leq 1$ при $i=0, 1, 2$. Решение с $i=0$ вновь определяет рассеянную вперед волну с $\theta = -90^\circ$ и $\psi = 0$. Однако решение с $i=1$ теперь приводит к $\theta = 0$ и $\psi = 90^\circ$, что определяет луч, рассеянный перпендикулярно направлению распространения. Решение с $i=2$ дает $\psi = 180^\circ$, что соответствует условию создания РОС. Эти рассеянные лучи, а также лучи, получающиеся при рассеянии на решетках третьего и четвертого порядков, изображены схематично на рис. 2.10.4, б, в и г соответственно. Рассеянные лучи

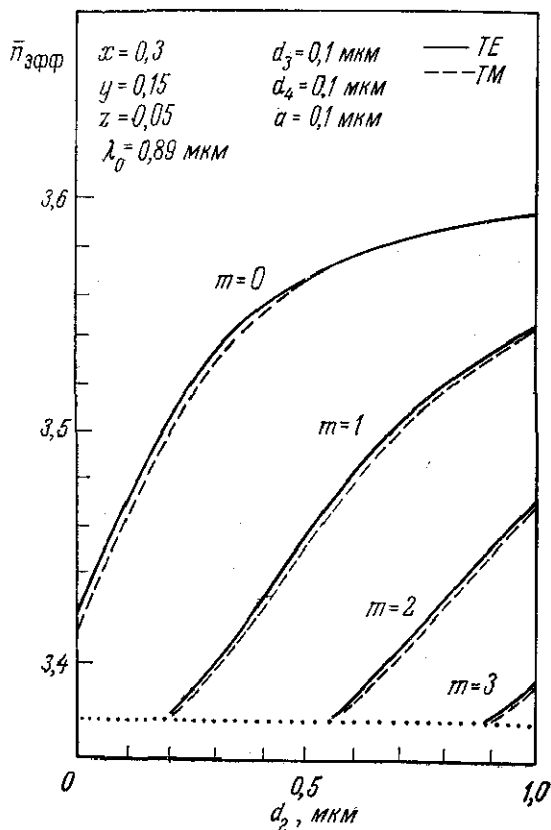


Рис. 2.10.7. Эффективный показатель преломления β/k_0 для поперечных мод с $m = 0, 1, 2, 3$ как функция толщины активного слоя [103].

могут выйти из полупроводника через его поверхность, параллельную направлению распространения первичных лучей, только в том случае, если их угол падения на эту поверхность меньше критического. Для GaAs с $\bar{n} \approx 3,6$ критический угол примерно равен 16° . Поэтому в направлении, перпендикулярном направлению волноводного распространения излучения, могут наблюдаться лишь лучи, соответствующие рассеянию на решетках только четных порядков, как, например, лучи на рис. 2.10.4, рассеянные решетками второго и четвертого порядков.

Примеры возможных значений постоянной связи для ОГС- и ДГС-лазеров с РОС даны Стрейфером и др. [102]. На рис. 2.10.5, а изображена гофрированная граница раздела в ДГС-лазере. Было показано, однако, что безызлучательная рекомбинация инжектированных носителей на гофрированной границе между активной областью и прилегающим к ней слоем $\text{Al}_x\text{Ga}_{1-x}\text{As}$ делает недостижимыми в таких лазерах низкие по-

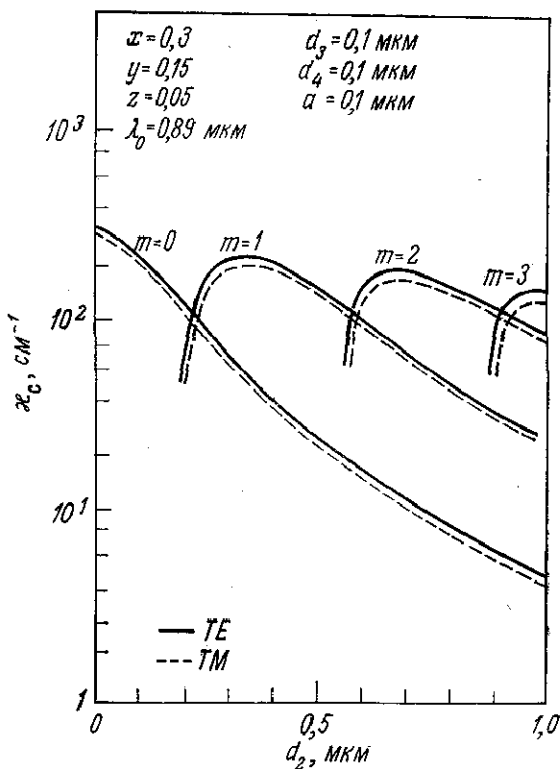


Рис. 2.10.8. Постоянная связи для изображенного на рис. 2.10.6 волновода как функция толщины активного слоя [103].

роги при комнатной температуре. Эта трудность была преодолена в ДГС — РО-лазерах путем удаления гофрированной поверхности от активного слоя [82, 83]. На рис. 2.10.5, б показан ДГС — РО-лазер с РОС, созданный Аики и др. [103]. Слой $P\text{-Al}_{0,05}\text{Ga}_{0,95}\text{As}$ был выращен для того, чтобы облегчить выращивание слоя $P\text{-Al}_{0,3}\text{Ga}_{0,7}\text{As}$ методом ЖФЭ. На рис. 2.10.5 показаны также поперечные распределения электрических полей в ДГС- и ДГС — РО-лазерах; видно, что в области с периодическим изменением показателя преломления электрическое поле имеет значительную амплитуду. Постоянная распространения и распределение электрического поля в волноводе, изображенном на рис. 2.10.6, были вычислены без учета гофра. На рис. 2.10.7 представлена зависимость эффективного показателя преломления β/k_0 для ТЕ- и ТМ-волн от толщины активного слоя [103]. Значения параметров, которые использовались при вычислениях, приведены на рисунке. При заданной длине волны излучения для рассчитанного значения $\bar{n}_{\text{эфф}}$ период гофра определяется по формуле (2.10.54). На рис. 2.10.8 приведены вычисленные по формуле (2.10.53) значения постоянной связи для ТЕ-мод при глубине гофра a и $l=3$ ($\Lambda \approx 0,37$ мкм). Было найдено, что постоянные связи для ТЕ- и ТМ-волн почти одинаковы. Выражение для постоянной связи ТМ-моды, полученное в работе [103], несколько отличается от используемого в работе [104]. Однако рассчитанные численные значения, видимо, не должны существенно отличаться от результатов, представленных на рис. 2.10.8. Экспериментальные характеристики лазеров с РОС обсуждаются в § 12 гл. 7.

§ 11. ЗАКЛЮЧЕНИЕ

В этой главе мы изложили основы теории и привели расчетные данные по волноводному распространению излучения в плоских диэлектрических волноводах. Сначала мы ввели уравнения Максвелла и, исходя из этих фундаментальных уравнений, получили формулы для конкретных случаев. Уравнения Максвелла используются также при исследовании взаимодействия свободных носителей с электромагнитным излучением (гл. 3) и при построении энергетических зонных диаграмм гетеропереходов (гл. 4).

Описано несколько различных подходов к решению волнового уравнения. Наиболее удобная форма дисперсионного уравнения для определения постоянной распространения в симметричных и асимметричных трехслойных плоских диэлектрических волноводах была получена на основе модели зигзагообразных волн. Коэффициент оптического ограничения, представляющий собой долю энергии моды, заключенную внутри активного слоя,

характеризует влияние распространения электромагнитной волны в гетеролазере на порог генерации в нем. Приведены типичные численные значения этого коэффициента в $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ диэлектрических плоских волноводах. Было найдено, что отражательная способность лазерных зеркал для ТЕ- и ТМ-волн совершенно различна и зависит от толщины активного слоя. Показано, что расходимость излучения в направлении, перпендикулярном плоскости $p-n$ -перехода, определяется произведением интеграла Фурье от поля в волноводе и углового фактора, причем последний часто не учитывают. Приведены приближенные выражения для коэффициента оптического ограничения и расходимости луча при малых толщинах активного слоя.

Приведенные в настоящей главе численные примеры относятся к распространению излучения с определенной фиксированной длиной волны в волноводах на основе $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$. Аналогичные зависимости получаются и для излучения на других длинах волн, а также для случаев, когда Al содержится и в активном слое. Выражения, описывающие систему $\text{Al}_x\text{Ga}_{1-x}\text{As}$, даны в форме, допускающей их применение к другим полупроводниковым твердым растворам при условии использования соответствующих значений физических параметров.

ЛИТЕРАТУРА

1. Marcuse D., Light Transmission Optics, Van Nostrand Reinhold, New York, 1972. (Имеется перевод: Маркузе Д. Оптические волноводы. — М.: Мир, 1974.)
2. Karany N. S., Burke J. J., Optical Waveguides, Academic Press, New York, 1972.
3. Marcuse D., Theory of Dielectric Optical Waveguides, Academic Press, New York, 1974.
4. Arnaud J. A., Beam and Fider Optics, Academic Press, New York, 1976.
5. Magid L. M., Electromagnetic Fields, Energy and Waves, Wiley, New York, 1972.
6. Magid L. M., Electromagnetic Fields, Energy and Waves, Wiley, New York, 1972, chapter 3.
7. Ramo S., Whinnery J. R., Fields and Waves in Modern Radio, Wiley, New York, 1953, p. 277. (Имеется перевод старого издания: Рамо С., Уиннери Дж. Поля и волны в современной радиотехнике. — М. — Л.: Гостехиздат, 1948.)
8. Cheroff G., Stern F., Triebwasser S., Appl. Phys. Lett., 2, 173 (1963).
9. Yarov A., Leite R. C. C., Appl. Phys. Lett., 2, 55 (1963).
10. McWhorter A. L., Solid-State Electron., 6, 417 (1963).
11. Hall R. N., Olechna D. J., Journ. Appl. Phys., 34, 2565 (1963).
12. Lascher G. J., IBM Journ., 7, 58 (1963).
13. Anderson W. W., IEEE Journ. Quantum Electron., QE-1, 228 (1965).
14. Sell D. D., Casey H. C. Jr., Wecht K. W., Journ. Appl. Phys., 45, 2650 (1974).
15. Marple D. T. F., Journ. Appl. Phys., 35, 1241 (1964).
16. Nash F. R., Journ. Appl. Phys., 44, 4696 (1973).
17. Strzalkowski I., Joshi S., Crowell C. R., Appl. Phys. Lett., 28, 350 (1976).
18. Vrehan Q. H. F., Journ. Phys. Chem. Solids, 29, 129 (1968).

19. Paoli T. L., IEEE Journ. Quantum Electron., **QE-13**, 662 (1977).
20. Cook D. D., Nash F. R., Journ. Appl. Phys., **46**, 1660 (1975).
21. Hayashi I., Panish M. B., Journ. Appl. Phys., **41**, 150 (1970).
22. Kressel H., Nelson H., RCA Rev., **30**, 106 (1969).
23. Byer N. E., Butler J. K., IEEE Journ. Quantum Electron., **QE-6**, 291 (1970).
24. Hayashi I. et al., Appl. Phys. Lett., **17**, 109 (1970).
25. Алферов Ж. И. и др., ФТП, **4**, 1826 (1970).
26. Adams M. J., Cross M., Solid-State Electron., **14**, 865 (1971).
27. Butler J. K., Journ. Appl. Phys., **42**, 4447 (1971).
28. Reinhart F. K., Hayashi I., Panish M. B., Journ. Appl. Phys., **42**, 4466 (1971).
29. Butler J. K., Kressel H., Ladany I., IEEE Journ. Quantum Electron., **QE-11**, 402 (1975).
30. Streifer W., Burnham R. D., Scifres D. R., IEEE Journ. Quantum Electron., **QE-12**, 177 (1976).
31. Kaplan W., Advanced Calculus, Addison-Wisley, Reading, Massachusetts, 1952, p. 146.
32. Weinberger H. F., Partial Differential Equations, Blaisdell, Waltham, Massachusetts, 1965, p. 63.
33. Magid L. M., Electromagnetic Fields, Energy and Waves, Wiley, New York, 1972, p. 107.
34. Ramo S., Whinnery J. R., Fields and Waves in Modern Radio, Wiley, New York, 1953, p. 192.
35. Marcuse D., Light Transmission Optics, Van Nostrand Reinhold, New York, 1972, p. 305. (Имеется перевод: Маркузе Д. Оптические волноводы. — М.: Мир, 1974.)
36. Casey H. C., Jr., Sell D. D., Panish M. B., Appl. Phys. Lett., **24**, 63 (1974).
37. Fern R. E., Onton A., Journ. Appl. Phys., **42**, 3499 (1971).
38. Clark G. D., Holonyak N., Jr., Phys. Rev., **156**, 913 (1967).
39. Casey H. C., Jr., Journ. Appl. Phys., **45**, 2766 (1974).
40. Afromowitz M. A., Solid State Commun., **15**, 59 (1974).
41. Casey H. C., Jr., Sell D. D., Wecht K. W., Journ. Appl. Phys., **46**, 250 (1975).
42. Collin R. E., Field Theory of Guided Waves, McGraw-Hill, New York, 1960, p. 470.
43. Dumke W. P., IEEE Journ. Quantum Electron., **QE-11**, 400 (1975).
44. Tien P. K., Appl. Opt., **10**, 2395 (1971).
45. Kogelnik H., Sosnowski T. P., Weber H. P., IEEE Journ. Quantum Electron., **QE-9**, 795 (1973).
46. Kogelnik H., Weber H. P., Journ. Opt. Soc. Am., **64**, 174 (1974).
47. Lotsch H. K. V., Journ. Opt. Soc. Am., **58**, 551 (1968).
48. Goos F., Haenchen H., Ann. Phys., **1**, 333 (1947).
49. Burke J. J., Opt. Sci. Newsletter (Univ. Arizona, Tucson), **5**, 31 (1971).
50. Marcuse D., Theory of Dielectric Optical Waveguides, Academic Press, New York, 1974, p. 3.
51. Marcuse D., Theory of Dielectric Optical Waveguides, Academic Press, New York, 1974, p. 7.
52. Cohen L. G., Bell Syst. Tech. Journ., **51**, 573 (1972).
53. Marcuse D., Light Transmission Optics, Van Nostrand Reinhold, New York, 1972, p. 44. (Имеется перевод: Маркузе Д. Оптические волноводы. — М.: Мир, 1974.)
54. Mathews J., Walker R. L. Mathematical Methods of Physics, Benjamin, New York, 1964, p. 85. (Имеется перевод: Мэтьюз Дж., Уокер Р. Математические методы физики. — М.: Атомиздат, 1972.)
55. Kirkby R. A., Thomson G. H. B., Opto-Electron., **4**, 323 (1972).
56. Hockham G. A., Electron. Lett., **9**, 389 (1973).
57. Lewin L., Electron. Lett., **10**, 134 (1974).
58. Lewin L., Journ. Appl. Phys., **46**, 2323 (1975).

59. *Butler J. K., Loroofchi J.*, IEEE Journ. Quantum Electron., **QE-10**, 809 (1974).
60. *Schlosser W. O.*, частное сообщение.
61. *Casey H. C., Jr., Panish M. B., Merz J. L.*, Journ. Appl. Phys., **44**, 5470 (1973).
62. *Thomson G. H. B., Kirkby P. A.*, IEEE Journ. Quantum Electron., **QE-9**, 311 (1973).
63. *Reinhart F. K., Hayashi I., Panish M. B.*, Journ. Appl. Phys., **42**, 4466 (1971).
64. *Gordon E. I.*, IEEE Journ. Quantum Electron., **QE-9**, 772 (1973).
65. *Krupka D. C.*, IEEE Journ. Quantum Electron., **QE-11**, 390 (1975).
66. *McKenna J.*, не опубликовано.
67. *Ikegami T.*, IEEE Journ. Quantum Electron., **QE-8**, 470 (1972).
68. *Panish M. B. et al.*, Appl. Phys. Lett., **22**, 590 (1973).
69. *Thompson G. H. B., Kirkby P. A.*, Electron. Lett., **9**, 295 (1973).
70. *Casey H. C. et al.*, Journ. Appl. Phys., **45**, 322 (1974).
71. *Lockwood H. F. et al.*, Appl. Phys. Lett., **17**, 499 (1970).
72. *Kressel H., Lockwood H. F., Hawrylo F. L.*, Appl. Phys. Lett., **18**, 43 (1971).
73. *Kressel H., Lockwood H. F., Hawrylo F. L.*, Journ. Appl. Phys., **43**, 561 (1972).
74. *Paoli T. L., Hakki B. W., Miller B. J.*, Journ. Appl. Phys., **44**, 1276 (1973).
75. *Smith G. E.*, IEEE Journ. Quantum Electron., **QE-4**, 288 (1968).
76. *Butler J. K., Kressel H.*, Journ. Appl. Phys., **43**, 3403 (1972).
77. *Hakki B. W.*, Journ. Appl. Phys., **45**, 288 (1974).
78. *Krupka D. C.*, частное сообщение.
79. *Kogelnik H., Shank C. V.*, Appl. Phys. Lett., **18**, 152 (1971).
80. *Kogelnik H., Shank C. V.*, Journ. Appl. Phys., **43**, 2327 (1972).
81. *Kogelnik H.*, Bell Syst. Tech. Journ., **48**, 2909 (1969).
82. *Casey H. C., Jr., Somekh S., Hegems M.*, Appl. Phys. Lett., **27**, 142 (1975).
83. *Aiki K. et al.*, Appl. Phys. Lett., **27**, 145 (1975).
84. *Reinhart F. K., Logan R. A., Shank C. V.*, Appl. Phys. Lett., **27**, 45 (1975).
85. *Yariv A.*, IEEE Journ. Quantum Electron., **QE-9**, 919 (1973).
86. *DeWames R. E., Hall W. F.*, Journ. Appl. Phys., **44**, 3638 (1973).
87. *Cordero R. F., Wang S.*, Appl. Phys. Lett., **24**, 474 (1974).
88. *Wang S.*, IEEE Journ. Quantum Electron., **QE-10**, 413 (1974).
89. *Kittel C.*, Introduction to Solid State Physics, Wiley, New York, 1968, p. 38. (Имеется перевод: Киттель Ч. Введение в физику твердого тела. — М.: Наука, 1978.)
90. *Chinn S. R.*, IEEE Journ. Quantum Electron., **QE-9**, 574 (1973).
91. *Streifer W., Burnham R. D., Scifres D. R.*, IEEE Journ. Quantum Electron., **QE-11**, 154 (1975).
92. *Streifer W., Burnham R. D., Scifres D. R.*, не опубликовано.
93. *Magid L. M.*, Electromagnetic Fields, Energy, and Waves, Wiley, New York, 1972, p. 339.
94. *Yariv A., Gover A.*, Appl. Phys. Lett., **26**, 537 (1975).
95. *Shank C. V., Schmidt R. V.*, Appl. Phys. Lett., **23**, 154 (1973).
96. *Garvin H. L. et al.*, Appl. Opt., **12**, 455 (1973).
97. *Somekh S., Casey H. C., Jr., Hegems M.*, Appl. Opt., **15**, 1905 (1976).
98. *Comerford L., Zory P.*, Appl. Phys. Lett., **25**, 208 (1974).
99. *Беляков Л. В. и др.*, ЖТФ, **44**, 1331 (1974).
100. *Tsang W.-T., Wang S.*, Appl. Phys. Lett., **28**, 44 (1976).
101. *Scifres D. R., Burnham R. D., Streifer W.*, Appl. Phys. Lett., **26**, 48 (1975).
102. *Streifer W., Scifres D. R., Burnham R. D.*, IEEE Journ. Quantum Electron., **QE-11**, 867 (1975).
103. *Aiki K., Nakamura M., Umeda J.-I.*, IEEE Journ. Quantum Electron., **QE-12**, 597 (1976).
104. *Streifer W., Scifres D. R., Burnham R. D.*, IEEE Journ. Quantum Electron., **QE-12**, 74 (1976).

ВЫНУЖДЕННОЕ ИЗЛУЧЕНИЕ В ПОЛУПРОВОДНИКАХ

§ 1. ВВЕДЕНИЕ

Вынужденное излучение представляет собой одно из наиболее интересных явлений, которые могут возникать при взаимодействии электромагнитного излучения с веществом. Это явление заключается в том, что фотон взаимодействует с электроном и, прежде чем поглотиться, индуцирует излучение идентичного фотона. Лазерный эффект получается при обеспечении обратной связи, т. е. возвращения части этого излучения в лазер. Теория лазера любого типа может быть развита из соотношений Эйнштейна [1] для скоростей переходов при поглощении и при вынужденном и спонтанном излучении. Однако характер вынужденного излучения в полупроводниках отличается от характера вынужденного излучения в газовых лазерах или в других твердотельных лазерах, что приводит к некоторому отличию в терминологии. В полупроводниках оптические переходы происходят между распределенными совокупностями энергетических уровней в зонах, в то время как в других лазерах переходы происходят обычно между дискретными энергетическими уровнями. Кроме того, в инжекционном лазере электроны тока накачки преобразуются с высокой квантовой эффективностью непосредственно в фотоны. В этой главе выводятся выражения, необходимые для вычисления коэффициента усиления в полупроводнике, а затем находятся и обсуждаются соотношения между коэффициентом усиления, потерями и плотностью порогового тока.

В следующем параграфе мы сначала выведем выражение для спектральной плотности удельной энергии фотонов излучения абсолютно черного тела, которое затем используем для получения соотношений Эйнштейна [1]. Эти соотношения показывают, что вероятности поглощения и вынужденного излучения равны друг другу и связаны с вероятностью спонтанного излучения. Соотношения Эйнштейна приводят к необходимому условию вынужденного излучения, полученному Бернардом и Дюрафуром [2]; это условие требует, чтобы расстояние между квазиуровнями Ферми превышало энергию излучаемого фотона. Из соотношений Эйнштейна мы получим выражения для коэффициента поглощения, скорости спонтанного излучения и суммарной скорости вынужденного излучения. Кроме того, мы выведем соотношения между коэффициентом поглощения и ско-

ростью спонтанного излучения, а также соотношение, связывающее скорости спонтанного и вынужденного излучений.

Соотношения Эйнштейна справедливы для любой материальной системы. Различия между разными веществами заключаются в вероятности перехода, которая должна быть оценена квантовомеханически. В § 3 этой главы для определения вероятности перехода используется «золотое правило» Ферми, представляющее собой результат применения теории возмущений, зависящих от времени, к взаимодействию электронов в твердом теле с электромагнитным излучением [3]. В соответствии с этим правилом вероятность перехода можно записать в виде квадрата матричного элемента, в который входят волновые функции начального и конечного состояний для данного перехода.

В § 4 приводятся выражения для коэффициента поглощения и скоростей спонтанного и вынужденного излучений в полупроводниках. Эти выражения требуют вычисления матричного элемента и плотности состояний в зоне проводимости и валентной зоне. Для обычно встречающихся концентраций примеси в активных областях полупроводниковых лазеров плотность состояний в зоне проводимости и валентной зоне зависит от концентрации примеси, что приводит к образованию «хвостов» зон внутри запрещенной зоны. Представление хвостов зон моделями Кейна [4] и Гальперина и Лэкса [5] дано в § 5 этой главы.

Матричный элемент рассматривается в § 6. Во многих случаях предполагают, что он определяется так называемым правилом k -отбора, которое требует, чтобы в прямозонных полупроводниках начальное и конечное состояния характеризовались одним и тем же волновым вектором. В противном случае матричный элемент равен нулю. Для полупроводников с примесями правило k -отбора не выполняется, и матричный элемент зависит от энергии. В § 6 выводится матричный элемент для перехода между параболической зоной и мелким уровнем примеси. Впервые вычисленный Иглсом [6] и позднее Думке [7], этот матричный элемент уменьшается с ростом энергии фотонов при удалении состояния в параболической зоне от края зоны. Он был использован Лэшером и Стерном [8] для первого детального вычисления коэффициента усиления в GaAs. Используя модель Кейна [4] для хвоста зоны и этот матричный элемент, Стерн [9] показал, какое влияние хвосты зон оказывают на коэффициент усиления. Вэнь [10] для вычисления коэффициента усиления с тем же матричным элементом использовал модель хвоста зоны, предложенную Гальперином и Лэксом [5]. Эти результаты показали, что при использовании такой сравнительно простой модели можно получить правильное представление о характере процесса. Стерн [11] предложил значительно

более точное, но более сложное выражение для матричного элемента. При выводе этого выражения вместе с предложенной Кейном [4] гауссовой формой хвоста зоны, аппроксимирующей плотность состояний в хвосте зоны в модели Гальперина — Лэкса, он использовал также специальным образом определенные волновые функции, описанные в § 6.

Следует подчеркнуть, что вычисление коэффициента усиления $g(E)$ представляет собой просто вычисление коэффициента поглощения $\alpha(E)$ как функции энергии фотона E . При высоком уровне накачки и энергиях фотонов, меньших чем расстояние между квазиуровнями Ферми, $\alpha(E)$ становится отрицательным. Этот диапазон энергий с отрицательным коэффициентом поглощения $\alpha(E)$ является спектральным диапазоном усиления, т. е. $g(E) = -\alpha(E)$. Сравнение в § 7 этой главы экспериментальных и вычисленных значений $\alpha(E)$ показывает, насколько хорошо численная оценка $\alpha(E)$, полученная с использованием матричного элемента Стерна [11], согласуется с экспериментом. Из этого расчета следует, что с ростом уровня накачки коэффициент усиления насыщается в области низких энергий и его максимум сдвигается в сторону высоких энергий.

В § 8 мы получим выражение для пороговой плотности тока $J_{\text{пор}}$. При этом используется хорошо известное условие, что на пороге усиления должно быть равно сумме всех потерь. На пороговый ток помимо радиационных потерь на зеркалах резонатора Фабри — Перо может оказывать влияние поглощение на свободных носителях, обусловленное инжектированием носителей и легированием. Интересно отметить, что для непрямозонных полупроводников можно ожидать более быстрого по сравнению с ростом коэффициента усиления увеличения поглощения на свободных носителях при высоком уровне накачки. Таким образом, вынужденное излучение становится маловероятным [12]. Поэтому в качестве активных слоев в лазерах на гетероструктурах можно использовать только прямозонные полупроводники. Плотности порогового тока при комнатной температуре, определенные на основе вычисленных в § 8 коэффициентов усиления, представлены в § 4 и 5 гл. 7 и прекрасно согласуются с экспериментальными значениями для GaAs—Al_xGa_{1-x}As ДГС- и ДГС — РО-лазеров.

§ 2. СООТНОШЕНИЯ МЕЖДУ ПОГЛОЩЕНИЕМ, ВЫНУЖДЕННЫМ ИЗЛУЧЕНИЕМ И СПОНТАННЫМ ИЗЛУЧЕНИЕМ

Вводные замечания

Простейшее описание взаимодействия света со свободными электронами в полупроводнике основано на соотношениях Эйн-

штейна [1], которые связывают скорости поглощения и вынужденного и спонтанного излучения. Хотя соотношения Эйнштейна были получены для атомной системы в свободном пространстве с двумя дискретными энергетическими уровнями, в дальнейшем они были распространены на широкополосные спектры [13]. При этом они оказались полезными и для полупроводников, в которых состояния распределены в валентной зоне и зоне проводимости. В этом приближении излучение абсолютно черного тела рассматривается как поле излучения. Понятие излучения абсолютно черного тела полезно также использовать для установления связи излучения, обусловленного внешним возбуждением, с измеряемым спектром поглощения. Распределения плотности фотонов и энергии для излучения абсолютно черного тела выводятся ниже и затем используются в соотношениях между поглощением и вынужденным и спонтанным излучениями.

Излучение абсолютно черного тела

Излучение абсолютно черного тела является равновесным излучением внутри полости при постоянной температуре и представляет определенный интерес, поскольку оно связано с разнообразными неравновесными процессами поглощения и излучения. Излучение абсолютно черного тела подробно рассматривается во многих работах, например в работе Кестина и Дорфмана [14]. Для излучения абсолютно черного тела нужно определить число фотонов на единицу объема и единицу частоты и число фотонов на единицу энергии и единицу объема, которые являются распределениями плотности фотонов. Для нахождения распределения плотности фотонов необходимы две величины. Одна из них представляет собой плотность состояний или число разрешенных решений (мод или состояний), получаемых из уравнений Максвелла. Другая величина — вероятность того, что фотон займет одно из указанных состояний; она определяется законом распределения Бозе — Эйнштейна.

Рассмотрим кубический замкнутый объем со стороной L и предположим, что допустимые решения волнового уравнения не зависят от граничных условий до тех пор, пока все размеры этого объема много больше рассматриваемых длин волн. Это предположение не просто доказать, и обычно оно принимается без доказательства. Джойс [15] довольно просто продемонстрировал независимость решений от граничных условий при выполнении указанных требований. Простейшими граничными условиями являются периодические условия, требующие, чтобы минимальный размер образца содержал много длин волн. В гл. 2 одномерное волновое уравнение (в котором β заменя-

ется на k_x) записывается в виде (2.2.25)

$$(\partial^2 \mathcal{E}_x / \partial z^2) + k_x^2 \mathcal{E}_x = 0,$$

где (2.2.30)

$$k_x = 2\pi/\lambda$$

— постоянная распространения, или волновой вектор. Волна, распространяющаяся в направлении $+z$, определяется выражением (2.2.27)

$$\mathcal{E}_x(z, t) = A \cos(\omega t - k_x z).$$

В соответствии с периодическим граничным условием имеем

$$\mathcal{E}_x(0, t) = \mathcal{E}_x(L, t). \quad (3.2.1)$$

Это равенство ограничивает k_x следующими дискретными значениями:

$$k_x = 2\pi m_x / L, \quad (3.2.2)$$

где m_x — целые числа $(0, \pm 1, \pm 2, \pm 3, \dots)$.

В случае замкнутого объема трехмерное волновое уравнение (2.4.8) с периодическими граничными условиями приводит к ограничениям на волновой вектор, аналогичным одномерному случаю:

$$k_x = 2\pi m_x / L, \quad k_y = 2\pi m_y / L, \quad k_z = 2\pi m_z / L; \quad (3.2.3)$$

здесь m_x, m_y, m_z — целые числа $(0, \pm 1, \pm 2, \pm 3, \dots)$. Эти дискретные значения k_x, k_y и k_z приводят к определенному набору конфигураций поля \mathcal{E} в объеме, которые называются «модами». Корпускулярная природа электромагнитного излучения описывается фотонами, энергия которых дается выражением

$$E = h\nu = |\mathbf{p}|c, \quad (3.2.4)$$

а импульс \mathbf{p} имеет следующие компоненты:

$$p_x = \hbar k_x, \quad p_y = \hbar k_y, \quad p_z = \hbar k_z, \quad (3.2.5)$$

где h — постоянная Планка, а $\hbar = h/2\pi$. Дискретные значения k_x, k_y, k_z приводят к дискретным значениям энергии фотонов, поэтому допустимые решения также называются «состояниями». Задача теперь состоит в том, чтобы подсчитать число этих состояний и вывести выражение для их числа на единицу объема, которое представляет собой плотность состояний. Определение плотности состояний встречается в различных областях физики (см., например, в работе Мандла [16]).

Плотность состояний выводится с помощью представления о волновом векторе \mathbf{k} , поэтому полезно ввести понятие \mathbf{k} -пространства, в котором любой вектор \mathbf{k} определяется выражением

$$\mathbf{k} = a_x k_x + a_y k_y + a_z k_z, \quad (3.2.6)$$

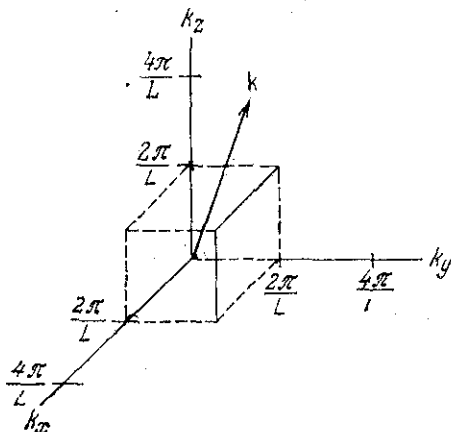


Рис. 3.2.1. Представление k -пространства через k_x , k_y и k_z , определенные выражениями (3.2.3) для периодических граничных условий.

где \mathbf{a}_x , \mathbf{a}_y , \mathbf{a}_z — единичные векторы в x -, y - и z - направлениях. На рис. 3.2.1 волновой вектор \mathbf{k} представлен в единицах k_x , k_y и k_z , определяемых выражениями (3.2.3), а единичный объем в k -пространстве изображен в виде куба, показанного штриховыми линиями, и дается формулой

$$\text{единичный объем в } k\text{-пространстве} = (2\pi/L)^3. \quad (3.2.7)$$

Допустимыми являются только те значения \mathbf{k} , которые определяются выражениями (3.2.3). Из этого следует, что число допустимых значений \mathbf{k} в любом объеме V_k k -пространства равно числу кубов со стороной $2\pi/L$, содержащихся в этом объеме.

Плотность состояний может быть определена из числа состояний, приходящихся на тонкий сферический слой k -пространства, заключенный между сферами, радиусы которых равны k и $k + dk$. Объем такого сферического слоя равен $4\pi k^2 dk$. Поскольку единица плотности состояний в k -пространстве есть величина, обратная единичному объему, определяемому формулой (3.2.7), искомое число состояний равно объему, умноженному на плотность, т. е.

$$dN(k) = 2(L/2\pi)^3 4\pi k^2 dk; \quad (3.2.8)$$

здесь множитель 2 учитывает два различных состояния поляризации фотона с заданным импульсом. В гл. 2 эти поляризации определены как поперечная электрическая (ТЕ) и поперечная магнитная (ТМ). Плотность состояний есть число состояний в единице объема твердого тела $V = L^3$ и равна

$$dN(k) = (k^2/\pi^2) dk. \quad (3.2.9)$$

В случае диэлектрика волновой вектор записывается в виде

$$k = 2\pi/\lambda = (2\pi/c) \bar{n} \nu, \quad (3.2.10)$$

где \bar{n} — показатель преломления, а ν — частота. Таким образом,

$$dk = 2\pi (\bar{n}/c) d\nu [1 + (\nu/\bar{n}) (d\bar{n}/d\nu)]; \quad (3.2.11)$$

здесь член в квадратных скобках представляет собой дисперсию показателя преломления и равен единице для свободного пространства. С учетом (3.2.10) и (3.2.11) уравнение (3.2.9) принимает вид

$$dN(\nu) = (8\pi\bar{n}^3 \nu^2/c^3) [1 + (\nu/\bar{n}) (d\bar{n}/d\nu)] d\nu. \quad (3.2.12)$$

Среднее число фотонов $\langle n_i \rangle$ в данном состоянии определяется законом распределения Бозе — Эйнштейна, который применим к системе одинаковых неразличимых частиц, обладающих целым спином [17]:

$$\langle n_i \rangle = [\exp(h\nu_i/kT) - 1]^{-1}. \quad (3.2.13)$$

Следовательно, распределение плотности фотонов $dD(\nu)$ дается произведением $\langle n_i \rangle dN(\nu)$ и его можно записать как

$$dD(\nu) = \frac{8\pi\bar{n}^3 \nu^2}{c^3} \frac{1 + (\nu/\bar{n}) (d\bar{n}/d\nu)}{\exp(h\nu/kT) - 1} d\nu. \quad (3.2.14)$$

Распределение плотности фотонов излучения абсолютно черного тела нередко записывается в виде распределения по энергиям, которое представляет собой энергию фотона $h\nu$, умноженную на распределение плотности фотонов $dD(\nu)$. В выражении (3.2.14) распределение плотности фотонов записано через число фотонов на единицу объема. Для дальнейших расчетов более удобно представить распределение плотности фотонов как число фотонов, приходящихся на единицу объема и единицу энергии. Используя соотношение (3.2.4) и записывая $d\nu = dE/h$, получаем

$$dD(E) = \frac{8\pi\bar{n}^3 E^2}{h^3 c^3} \frac{1 + (E/\bar{n}) (d\bar{n}/dE)}{\exp(E/kT) - 1} dE. \quad (3.2.15)$$

При этом спектральную плотность относительно удельной энергии E можно обозначить через $P(E)$ и определить ее таким образом, что $dD(E) = P(E)dE$ и

$$P(E) = \frac{8\pi\bar{n}^3 E^2 [1 + (E/\bar{n}) (d\bar{n}/dE)]}{h^3 c^3 [\exp(E/kT) - 1]}. \quad (3.2.16)$$

Спектральная плотность имеет размерность числа фотонов, приходящихся на единицу объема и единицу энергии. Дисперсионный член $[1 + (E/\bar{n}) (d\bar{n}/dE)]$ обычно полагают равным единице.

Соотношения Эйнштейна

Разрешенные электронные состояния в полупроводниках в отличие от двух резких уровней в атомной системе непрерывно распределены внутри валентной зоны и зоны проводимости. На рис. 3.2.2 изображен переход электрона из состояния с энергией E_1 в валентной зоне в состояние с энергией E_2 в зоне проводимости, происходящий вследствие поглощения фотона с энергией $E_2 - E_1 = E_{21} = \hbar\omega$. Скорость этого процесса зависит от таких факторов:

- 1) вероятность перехода B_{12} ;
- 2) вероятность того, что состояние с E_1 занято электроном, f_1 ;
- 3) вероятность того, что состояние с E_2 свободно, $(1 - f_2)$;
- 4) плотность фотонов с энергией E_{21} $P(E_{21})$. При этом скорость переходов вверх может быть записана в виде

$$\lambda_{12} = B_{12} f_1 [1 - f_2] P(E_{21}). \quad (3.2.17)$$

Вероятность того, что состояние с E_1 занято электроном, определяется функцией распределения Ферми — Дирака, описывающей неразличимые, тождественные частицы с полуцелым спином, подчиняющиеся принципу исключения Паули. Она равна [17]

$$f_1 = (\exp [(E_1 - F_1)/kT] + 1)^{-1}; \quad (3.2.18)$$

здесь F_1 — квазиуровень Ферми для электронов валентной зоны в отсутствие равновесия. Для зоны проводимости имеем аналогично

$$f_2 = (\exp [(E_2 - F_2)/kT] + 1)^{-1}; \quad (3.2.19)$$

здесь F_2 — квазиуровень Ферми для электронов зоны проводимости в отсутствие равновесия.

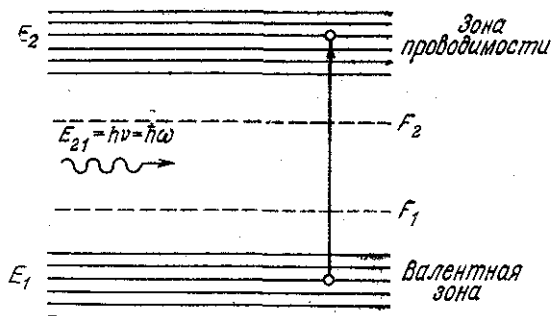


Рис. 3.2.2. Переход электрона с уровня E_1 на уровень E_2 при поглощении фотона с энергией E_{21} . Квазиуровни Ферми для валентной зоны и зоны проводимости обозначены соответственно через F_1 и F_2 .

Кроме участия в процессе поглощения фотоны с энергией E_{21} могут индуцировать переход электрона из состояния с E_2 в состояние с E_1 с испусканием аналогичного фотона. Для этого процесса скорость переходов вниз определяется выражением

$$\tau_{21} = B_{21} f_2 [1 - f_1] P(E_{21}), \quad (3.2.20)$$

где B_{21} — вероятность перехода, f_2 — вероятность того, что состояние с E_2 занято электроном, $[1 - f_1]$ — вероятность того, что состояние с E_1 свободно.

Кроме того, электроны из состояния с E_2 могут переходить в состояние с E_1 спонтанно, без взаимодействия с полем излучения $P(E_{21})$. Вероятность этого процесса равна A_{21} , а скорость спонтанного излучения определяется следующим образом:

$$\tau_{21} (\text{спонт.}) = A_{21} f_2 [1 - f_1]. \quad (3.2.21)$$

Ряд полезных соотношений можно получить, исходя из выражений (3.2.17), (3.2.20) и (3.2.21).

При тепловом равновесии скорость переходов вверх должна равняться общей скорости переходов вниз, т. е.

$$\tau_{12} = \tau_{21} + \tau_{21} (\text{спонт.}). \quad (3.2.22)$$

Кроме того, $F_1 = F_2$. Используя выражения (3.2.17), (3.2.20) и (3.2.21), соотношение (3.2.22) можно переписать в виде

$$P(E_{21}) = \frac{A_{21} f_2 [1 - f_1]}{B_{12} f_1 [1 - f_2] - B_{21} f_2 [1 - f_1]}. \quad (3.2.23)$$

Дальнейшие преобразования этого соотношения с учетом выражений (3.2.16), (3.2.18) и (3.2.19) дают

$$\frac{8\pi \bar{n}^3 E_{21}^2}{h^3 c^3 [\exp(E_{21}/kT) - 1]} = \frac{A_{21}}{B_{12} \exp(E_{21}/kT) - B_{21}}. \quad (3.2.24)$$

Здесь множитель $[1 + (h\nu/\bar{n}) (d\bar{n}/d\nu)]$ положен равным единице. Следует отметить, что часто вместо распределения плотности фотонов по энергии используют распределение плотности энергии излучения по частоте. В этом случае [1, 13, 14] левая часть (3.2.24) принимает вид

$$\frac{8\pi \bar{n}^3 h\nu^3}{c^3 [\exp(h\nu/kT) - 1]}.$$

Соотношение (3.2.24) можно переписать в виде

$$(8\pi \bar{n}^3 E_{21}^2 / h^3 c^3) [B_{12} \exp(E_{21}/kT) - B_{21}] = A_{21} \exp(E_{21}/kT) - A_{21} \quad (3.2.25)$$

и разбить на две части — зависящую и независящую от температуры. Для независящей от температуры части получаем

$$A_{21} = (8\pi \bar{n}^3 E_{21}^2 / h^3 c^3) B_{21}, \quad (3.2.26)$$

а для зависящей от температуры части с учетом (3.2.26) находим

$$B_{12} = B_{21}. \quad (3.2.27)$$

Соотношения (3.2.26) и (3.2.27) называются соотношениями Эйнштейна. Они определяют связь вероятности спонтанного излучения с вероятностями поглощения и вынужденного излучения.

Необходимое условие вынужденного излучения в полупроводниках

Необходимое условие вынужденного излучения было получено Бернарром и Дюрафуром [2] в очень простой форме. Процесс вынужденного излучения преобладает над процессом поглощения в том случае, когда вероятность стимулированного фотоном перехода электрона из зоны проводимости в валентную зону больше вероятности обратного перехода электрона (из валентной зоны в зону проводимости) при поглощении фотона. Это условие означает, что ϵ_{21} [выражение (3.2.20)] должно превосходить ϵ_{12} [выражение (3.2.17)], т. е.

$$B_{21}f_2[1 - f_1]P(E_{21}) > B_{12}f_1[1 - f_2]P(E_{21}). \quad (3.2.28)$$

Поскольку $B_{12} = B_{21}$, это соотношение дает

$$f_2[1 - f_1] > f_1[1 - f_2], \quad (3.2.29)$$

где f_1 и f_2 определяются формулами (3.2.18) и (3.2.19) соответственно. Из соотношения (3.2.29) следует

$$\exp[(F_2 - F_1)/kT] > \exp[(E_2 - E_1)/kT], \quad (3.2.30)$$

или

$$F_2 - F_1 > E_2 - E_1 \quad (3.2.31)$$

Следовательно, разделение квазиуровней Ферми должно превосходить энергию излучаемого фотона, чтобы скорость вынужденных переходов вниз превышала скорость переходов вверх, происходящих вследствие поглощения.

Скорость вынужденного излучения

Часто используют величину, называемую суммарной скоростью вынужденного излучения r_{21} (вынужд.) и представляющую собой разность между скоростями переходов вниз ϵ_{21} и вверх ϵ_{12} . Из соотношений (3.2.17) и (3.2.20) находим

$$r_{21} (\text{вынужд.}) = B_{21}f_2[1 - f_1]P(E_{21}) - B_{12}f_1[1 - f_2]P(E_{21}), \quad (3.2.32)$$

и, поскольку $B_{12} = B_{21}$, после простых преобразований это выражение принимает вид

$$r_{21} (\text{вынужд.}) = B_{12} P(E_{21}) [f_2 - f_1]. \quad (3.2.33)$$

Далее, учитывая соотношения (3.2.16) и (3.2.26), из выражения (3.2.33) получаем

$$r_{21} (\text{вынужд.}) = \frac{A_{21} [f_2 - f_1]}{\exp(E_{21}/kT) - 1}. \quad (3.2.34)$$

Стало общепринятым [18] называть скоростью вынужденного излучения $r_{\text{вынужд.}}(E_{21})$ величину $A_{21}[f_2 - f_1]$ вследствие ее схожести по форме записи со скоростью спонтанного излучения, определяемой выражением (3.2.21):

$$r_{\text{вынужд.}}(E_{21}) = A_{21} [f_2 - f_1]. \quad (3.2.35)$$

Единицей измерения $r_{\text{вынужд.}}$ служит число фотонов в единичном объеме, отнесенное к единице времени и единичному интервалу энергии.

Следует отметить, что скоростью вынужденного излучения называют три довольно близкие величины:

1) i_{21} — скорость переходов вниз, сопровождающихся вынужденным излучением;

2) $r_{21} (\text{вынужд.})$ — суммарную скорость вынужденного излучения, равную разности между скоростями переходов вниз i_{21} и вверх i_{12} ;

3) $r_{\text{вынужд.}}(E_{21})$ — величину, обычно называемую скоростью вынужденного излучения, которая, будучи умноженной на число фотонов в соответствующем состоянии $[\exp(E_{21}/kT) - 1]^{-1}$, дает суммарную скорость вынужденного излучения $r_{21} (\text{вынужд.})$.

Коэффициент поглощения

Взаимодействие фотонов и электронов в твердом теле может быть описано макроскопической величиной — коэффициентом поглощения α . Суммарная скорость поглощения представляет собой разность между скоростями переходов вверх i_{12} и вниз i_{21} . Из соотношений (3.2.17) и (3.2.20) получаем для этой разности выражение

$$r_{12} (\text{погл.}) = B_{12} f_1 [1 - f_2] P(E_{21}) - B_{21} f_2 [1 - f_1] P(E_{21}), \quad (3.2.36)$$

которое при учете равенства $B_{12} = B_{21}$ принимает вид

$$r_{12} (\text{погл.}) = B_{12} [f_1 - f_2] P(E_{21}), \quad (3.2.37)$$

где величина $B_{12} [f_1 - f_2]$ определяет суммарную вероятность поглощения.

Суммарная скорость поглощения r_{12} (погл.) равна коэффициенту поглощения, умноженному на поток фотонов $F(E)$. Этот поток представляет собой определяемое формулой (3.2.16) распределение плотности фотонов $P(E)$, умноженное на групповую скорость v_g . В диэлектрической среде групповая скорость равна

$$v_g = \frac{d\omega}{dk} = \frac{2\pi dE}{\hbar dk} = \frac{c/\bar{n}}{1 + (E/\bar{n})(d\bar{n}/dE)} \quad (3.2.38)$$

[при выводе этой формулы использованы выражения (3.2.11) для dk и (3.2.4), определяющее $\omega = 2\pi E/\hbar$]. Таким образом, коэффициент поглощения может быть записан в следующем виде:

$$\alpha(E_{21}) = \frac{r_{12}(\text{погл.})}{F(E_{21})} = \frac{r_{12}(\text{погл.})}{P(E_{21})v_g} = \frac{B_{12}[f_1 - f_2]}{c/\bar{n}}, \quad (3.2.39)$$

где дисперсионный член $[1 + (E/\bar{n})(d\bar{n}/dE)]$ в v_g положен равным единице.

Коэффициент поглощения можно связать со скоростью вынужденного излучения $r_{\text{вынужд.}}(E_{21})$. Соотношение (3.2.26) позволяет переписать выражение (3.2.35) в виде

$$-r_{\text{вынужд.}}(E_{21}) = (8\pi\bar{n}^3 E_{21}^2 / \hbar^3 c^3) B_{12}[f_1 - f_2], \quad (3.2.40)$$

так что для $\alpha(E)$ получаем формулу

$$\alpha(E_{21}) = -(\hbar^3 c^2 / 8\pi\bar{n}^2 E_{21}^2) r_{\text{вынужд.}}(E_{21}). \quad (3.2.41)$$

Единицей измерения $\alpha(E_{21})$ служит см^{-1} . Из соотношения (3.2.41) следует, что $\alpha(E_{21})$ и $r_{\text{вынужд.}}(E_{21})$ различаются постоянным множителем.

Связь спонтанного излучения с коэффициентом поглощения

Скорость спонтанного излучения можно выразить через коэффициент поглощения, если использовать соотношение Эйнштейна (3.2.26). Подставляя выражение (3.2.37) для r_{12} (погл.) в формулу (3.2.39), получаем

$$B_{12}[f_1 - f_2]P(E_{21}) = \alpha(E_{21})P(E_{21})v_g, \quad (3.2.42)$$

откуда с учетом соотношений (3.2.21) и (3.2.26) имеем

$$r_{21}(\text{спонт.}) = (8\pi\bar{n}^3 E_{21}^2 / \hbar^3 c^3) B_{21}f_2[1 - f_1] = r_{\text{спонт.}}(E_{21}). \quad (3.2.43)$$

Соотношение (3.2.42) позволяет исключить B_{21} из формулы (3.2.43). При $v_g = c/\bar{n}$ имеем

$$r_{\text{спонт.}}(E_{21}) = \frac{8\pi\bar{n}^2 E_{21}^2}{\hbar^3 c^2} \alpha(E_{21}) \frac{f_2[1 - f_1]}{[f_1 - f_2]}, \quad (3.2.44)$$

или

$$r_{\text{спонт}}(E_{21}) = \frac{8\pi n^2 E_{21}^2 \alpha(E_{21})}{h^3 c^2 \{\exp\{[E_{21} - (F_2 - F_1)]/kT\} - 1\}}. \quad (3.2.45)$$

Единицы измерения $r_{\text{спонт}}(E_{21})$ те же, что и для $r_{\text{вынужд}}(E_{21})$. Замена обозначения (3.2.43) проведена для поддержания единства в системе обозначений.

Связь между скоростями спонтанного и вынужденного излучений

Соотношения (3.2.41) и (3.2.45) связывают $r_{\text{вынужд}}(E_{21})$ и $r_{\text{спонт}}(E_{21})$ с $\alpha(E_{21})$. Поэтому легко найти связь между скоростью вынужденного излучения, определяемой выражением (3.2.35), и скоростью спонтанного излучения:

$$r_{\text{спонт}}(E_{21}) = - \frac{r_{\text{вынужд}}(E_{21})}{\exp\{[E_{21} - (F_2 - F_1)]/kT\} - 1}. \quad (3.2.46)$$

или

$$r_{\text{вынужд}}(E_{21}) = r_{\text{спонт}}(E_{21}) \{1 - \exp\{[E_{21} - (F_2 - F_1)]/kT\}\}. \quad (3.2.47)$$

Соотношения (3.2.41), (3.2.45) и (3.2.47) показывают, что величины $\alpha(E)$, $r_{\text{спонт}}(E)$ и $r_{\text{вынужд}}(E)$ связаны друг с другом, и если известна одна из них, то с помощью этих соотношений можно определить две другие. Для того чтобы использовать эти соотношения, необходимо определить экспериментально одну из величин, а именно $\alpha(E)$, или вычислить вероятности переходов B_{12} или A_{21} . Ниже в этой главе мы рассмотрим выражения для $\alpha(E)$, $r_{\text{спонт}}(E)$ и $r_{\text{вынужд}}(E)$, содержащие вероятности переходов и плотности состояний, характерные для полупроводников.

§ 3. ВЕРОЯТНОСТЬ ПЕРЕХОДА

Золотое правило Ферми

Если известен $\alpha(E)$, то, как следует из соотношения (3.2.45), по нему можно определить скорость спонтанного излучения $r_{\text{спонт}}(E)$. В свою очередь $r_{\text{вынужд}}(E)$ связано с $r_{\text{спонт}}(E)$ соотношением (3.2.47), и, следовательно, для определения обеих скоростей необходимо знать только $\alpha(E)$. Кроме того, вычисленные значения $\alpha(E)$ легко проверить экспериментальными измерениями. В силу этих причин при последующем анализе мы будем рассматривать только коэффициент поглощения $\alpha(E)$. При наличии двух дискретных уровней он определяется выражением (3.2.39):

$$\alpha(E_{21}) = B_{12} [f_1 - f_2] \bar{n}/c.$$

В этом выражении неизвестна вероятность перехода B_{12} , причем она, вероятно, зависит от различных свойств системы, связанных с взаимодействием между электронами в твердом теле и электромагнитным излучением. Для рассмотрения этого взаимодействия необходимо обратиться к квантовой механике.

Взаимодействие электронов в полупроводнике с электромагнитным излучением описывается с помощью теории возмущений, зависящих от времени [3]. В этом случае сначала определяют характеристики системы в отсутствие излучения, а затем вычисляется их изменение, происходящее вследствие взаимодействия системы с излучением. Если возмущающие воздействия дают сходящийся ряд теории возмущений, то могут быть получены решения для реальной задачи. Вместо того чтобы проводить трудоемкий непосредственный анализ этого вопроса, мы приведем хорошо известное соотношение, получающееся в результате такого анализа и называемое «золотым правилом» Ферми. Решение уравнения Шредингера с периодическим по времени возмущением

$$H^I(\mathbf{r}, t) = H^I(\mathbf{r}) \cos \omega t \quad (3.3.1)$$

приводит к выражению для вероятности перехода в виде [3]

$$B_{12} = (\pi/2\hbar) |\langle \Psi_1^*(\mathbf{r}, t) | H^I | \Psi_2(\mathbf{r}, t) \rangle|^2. \quad (3.3.2)$$

Это соотношение и называется золотым правилом Ферми. Здесь \mathbf{r} — трехмерный радиус-вектор, $\Psi_1^*(\mathbf{r}, t)$ — функция, комплексно сопряженная с волновой функцией начального состояния, H^I — гамильтониан взаимодействия, $\Psi_2(\mathbf{r}, t)$ — волновая функция конечного состояния. Вероятность перехода, полученная из золотого правила Ферми, была дана Стерном [19] и Бейбом и Уильямсом [23].

Величину $\langle \Psi_1^*(\mathbf{r}, t) | H^I | \Psi_2(\mathbf{r}, t) \rangle$ обычно называют матричным элементом гамильтониана взаимодействия H^I , соответствующим переходу между состояниями $\Psi_1(\mathbf{r}, t)$ и $\Psi_2(\mathbf{r}, t)$. Как мы увидим в дальнейшем, H^I представляет собой известный дифференциальный оператор ∇ . Поэтому матричный элемент определяется скалярным произведением волновой функции начального состояния $\Psi_1^*(\mathbf{r}, t)$ и волновой функции, получающейся в результате воздействия оператора H^I (дифференцирования) на волновую функцию конечного состояния $\Psi_2(\mathbf{r}, t)$, и равен [3]

$$\langle \Psi_1^*(\mathbf{r}, t) | H^I | \Psi_2(\mathbf{r}, t) \rangle = \int_V \Psi_1^*(\mathbf{r}, t) H^I \Psi_2(\mathbf{r}, t) d^3r. \quad (3.3.3)$$

Остальная часть параграфа посвящена выводу выражения, описывающего входящий в матричный элемент (3.3.3) гамильтониан взаимодействия.

Гамильтониан взаимодействия

В этой части параграфа мы рассмотрим величины, входящие в гамильтониан взаимодействия, стоящий в формуле (3.3.3). Очень подробное обсуждение идей, которые будут кратко изложены ниже, можно найти в работе Уайта [3].

В квантовой механике частица описывается своей волновой функцией $\Psi(\mathbf{r}, t)$. Эта функция состояния содержит всю информацию о системе. Поведение $\Psi(\mathbf{r}, t)$ как функции пространственных координат и времени определяется волновым уравнением Шредингера

$$H_{\text{оп}}\Psi(\mathbf{r}, t) = -(\hbar/i)[\partial\Psi(\mathbf{r}, t)/\partial t], \quad (3.3.4)$$

где $H_{\text{оп}}$, как будет показано ниже, представляет собой оператор, соответствующий гамильтониану системы.

Гамильтониан системы связан с ее полной энергией. Классический гамильтониан частицы является консервативной величиной и представляет собой сумму кинетической и потенциальной энергий

$$H = (p^2/2m) + V(\mathbf{r}), \quad (3.3.5)$$

где m — масса частицы. Консервативной называется сила, которая полностью определяется однозначным потенциалом $V(\mathbf{r})$ и имеет вид $\mathbf{F} = -\nabla V(\mathbf{r})$. Сила консервативна в том случае, когда потенциал частицы в данной точке не зависит от траектории, по которой частица в эту точку попала. Это условие означает, что потенциал частицы не зависит от ее скорости в этой точке. При исследовании взаимодействия частицы с электромагнитным полем необходимо рассмотреть действие неконсервативной силы на электрон.

Такой математический оператор, как $H_{\text{оп}}$, может быть связан с физически наблюдаемыми величинами, как, например, с $\Psi(\mathbf{r}, t)$. Исходя из свойств оператора можно изучить поведение функции $\Psi(\mathbf{r}, t)$. Так, в одномерном случае выражение $\partial\Psi(x, t)/\partial x$ можно рассматривать как составленное из оператора $\partial/\partial x$ и операнда $\Psi(x, t)$. Импульсу p соответствует оператор [3]

$$p \rightarrow (\hbar/i)\nabla, \quad (3.3.6)$$

так что оператор $H_{\text{оп}}^0$, связанный с классическим гамильтонианом, принимает вид

$$H_{\text{оп}}^0 = -(\hbar^2\nabla^2/2m) + V(\mathbf{r}). \quad (3.3.7)$$

С учетом этого выражения волновое уравнение (3.3.4) запишется следующим образом:

$$[-(\hbar^2/2m)\nabla^2 + V(\mathbf{r})]\Psi(\mathbf{r}, t) = -(\hbar/i)[\partial\Psi(\mathbf{r}, t)/\partial t]. \quad (3.3.8)$$

Уравнение (3.3.8) можно разделить на два уравнения, одно из которых определяет координатную зависимость $\Psi(\mathbf{r}, t)$, а другое — временную. Метод разделения переменных уже применялся в § 4 гл. 2 при решении волнового уравнения. Как и в задаче о волноводном распространении излучения, применение граничных условий приводит к уравнениям на собственные значения. Собственные решения уравнения (3.3.8) получаются в виде произведения собственной функции, зависящей от пространственных переменных, на гармоническую функцию от времени:

$$\Psi_n(\mathbf{r}, t) = \psi_n(\mathbf{r}) \exp(j\omega_n t). \quad (3.3.9)$$

Далее рассмотрим гамильтониан электрона, взаимодействующего с электромагнитным полем (см. работу [3] гл. 2). На электрон с зарядом q , движущийся со скоростью \mathbf{v} в электромагнитном поле, действует сила Лоренца

$$\mathbf{F} = q(\mathcal{E} + \mathbf{v} \times \mathcal{B}). \quad (3.3.10)$$

Поскольку сила \mathbf{F} зависит от \mathbf{v} , она не является консервативной и не может быть представлена в виде градиента скалярного поля. Действие этой неконсервативной силы удобно описывать, представив электромагнитное поле посредством векторного поля \mathbf{A} . Это поле находится из уравнений Максвелла (2.2.1), (2.2.7):

$$\begin{aligned} \nabla \times \mathcal{E} &= -\partial \mathcal{B} / \partial t, \\ \nabla \cdot \mathcal{B} &= 0. \end{aligned}$$

Из векторной алгебры известно, что дивергенция ротора равна нулю, т. е.

$$\nabla \cdot \nabla \times \mathbf{G} = 0. \quad (3.3.11)$$

Поскольку $\nabla \cdot \mathcal{B} = 0$, вектор \mathcal{B} можно представить в виде ротора некоторого вектора \mathbf{A} :

$$\mathcal{B} = \nabla \times \mathbf{A}. \quad (3.3.12)$$

Тогда уравнение (2.2.1) перепишется в виде

$$\nabla \times \mathcal{E} = -(\partial / \partial t) \nabla \times \mathbf{A}, \quad (3.3.13)$$

или

$$\nabla \times (\mathcal{E} + (\partial / \partial t) \mathbf{A}) = 0. \quad (3.3.14)$$

Из векторной алгебры известно также, что если ротор вектора равен нулю, то этот вектор должен являться градиентом некоторого потенциала φ . Таким образом,

$$\mathcal{E} + (\partial / \partial t) \mathbf{A} = -\nabla \varphi. \quad (3.3.15)$$

Вектор \mathbf{A} и скаляр φ могут быть выбраны так, что $\varphi = 0$ и

$$\mathcal{E} = -\partial\mathbf{A}/\partial t. \quad (3.3.16)$$

Гамильтониан электрона, движущегося в неконсервативном потенциале, определяется хорошо известным выражением (см. [3] гл. 2)

$$H = (1/2m)(\mathbf{p} - q\mathbf{A})^2 + V(\mathbf{r}), \quad (3.3.17)$$

где \mathbf{A} определяется соотношением (3.3.12), а \mathcal{E} выражается через \mathbf{A} по формуле (3.3.16).

Таким образом, взаимодействующий с электромагнитным полем электрон описывается гамильтонианом

$$H = (1/2m)(\mathbf{p}^2 - 2q\mathbf{A} \cdot \mathbf{p} + q^2\mathbf{A}^2) + V(\mathbf{r}). \quad (3.3.18)$$

Пренебрегая членом $q^2\mathbf{A}^2$ и учитывая соотношение (3.3.6), получаем оператор $H_{\text{оп}}$ в следующем виде:

$$H_{\text{оп}} = (1/2m)(-\hbar^2\nabla^2 - 2(q\hbar/j)\mathbf{A} \cdot \nabla) + V(\mathbf{r}). \quad (3.3.19)$$

Сравнение этого выражения с выражением для классического гамильтониана (3.3.7) дает

$$H_{\text{оп}} = H_{\text{оп}}^0 + H_{\text{оп}}^I, \quad (3.3.20)$$

где

$$H_{\text{оп}}^I = - (q\hbar/jm)\mathbf{A} \cdot \nabla, \quad (3.3.21)$$

или

$$H_{\text{оп}}^I = - (q/m)\mathbf{A} \cdot \mathbf{p}. \quad (3.3.22)$$

Выражение (3.3.21) является операторной формой энергии взаимодействия (3.3.22).

Матричный элемент

Выражение для матричного элемента (3.3.3) с определяемым выражением (3.3.22) H^I можно упростить, если векторный потенциал выразить через вектор электрического поля, который для бегущей волны может быть записан в следующем виде [см. выражение (2.2.27)]:

$$\mathcal{E} = \mathbf{a}_x \mathcal{E}_0 \cos(\omega t - kz) = \mathbf{a}_x \mathcal{E}_0 \exp[j(\omega t - kz)], \quad (3.3.23)$$

где подразумевается, как и в § 2 гл. 2, что берется действительная часть $\exp[j(\omega t - kz)]$. Электрическое поле, описываемое этим выражением, поляризовано параллельно оси x и имеет волновой вектор \mathbf{k} . Для кубических кристаллов, которые рассматриваются здесь, направление поляризации несущественно. Векторный потенциал \mathbf{A} связан с \mathcal{E} соотношением (3.3.16), и, следовательно,

$$\mathbf{A} = (j\mathbf{a}_x \mathcal{E}_0/\omega) \exp[j(\omega t - kz)] \quad (3.3.24)$$

и

$$|\mathbf{A}|^2 = |\mathbf{A} \cdot \mathbf{A}^*| = \mathcal{E}_0^2 / \omega^2. \quad (3.3.25)$$

Для определения амплитуды вектора электрического поля \mathcal{E}_0 выразим поток электромагнитного излучения, который определяется действительной частью вектора Пойнтинга, через энергию фотонов $E = \hbar\omega$. Для \mathcal{E} , определяемого выражением (3.3.23), из соотношения (2.2.14) получаем следующее выражение для магнитного вектора \mathcal{H} :

$$\mathcal{H}_y = (a_y \mathcal{E}_0 k / \mu_0 \omega) \exp[j(\omega t - kz)]. \quad (3.3.26)$$

Тогда из формулы (2.2.57), определяющей вектор Пойнтинга в виде

$$|\mathbf{S}| = \frac{1}{2} \operatorname{Re}(\mathcal{E} \times \mathcal{H}^*),$$

и выражений (3.3.23) и (3.3.26) имеем

$$|\mathbf{S}| = \frac{1}{2} (\mathcal{E}_0^2 k / \mu_0 \omega). \quad (3.3.27)$$

С другой стороны, поток энергии, переносимой фотоном, равен энергии фотона $\hbar\omega$, умноженной на его групповую скорость. Следовательно,

$$\hbar\omega c / \bar{n} = \frac{1}{2} \bar{n} \epsilon_0 \mathcal{E}_0^2 c, \quad (3.3.28)$$

где выражение в правой части равенства получено из соотношения (3.3.27) подстановкой $k = \bar{n}\omega/c$ и $c^2 = 1/\mu_0\epsilon_0$. Из формул (3.3.25) и (3.3.28) следует

$$|\mathbf{A}|^2 = 2\hbar/\epsilon_0 \bar{n}^2 \omega. \quad (3.3.29)$$

В более общем случае, когда вектор поляризации волны \mathbf{n} направлен произвольным образом в пространстве, $|\mathbf{A}|^2$ определяется тем же выражением (3.3.29), а для \mathbf{A} справедливо выражение

$$\mathbf{A} = (\mathbf{n} \mathcal{E}_0 / \omega) \exp[j(\omega t - \mathbf{k} \cdot \mathbf{r})]. \quad (3.3.30)$$

Теперь, используя выражения (3.3.9), (3.3.22) и (3.3.30) для волновых функций, оператора взаимодействия и векторного потенциала соответственно, перепишем формулу (3.3.2) для вероятности перехода в следующем виде:

$$B_{12} = \frac{\pi}{2\hbar} |\langle \psi_1^*(\mathbf{r}) \exp(-j\omega_1 t) | - \\ - \mathbf{n} \frac{q}{m} \left(\frac{2\hbar}{\epsilon_0 \bar{n}^2 \omega} \right)^{1/2} \exp[j(\omega t - \mathbf{k} \cdot \mathbf{r})] \cdot \mathbf{p} | \psi_2(\mathbf{r}) \exp(j\omega_2 t) \rangle|^2. \quad (3.3.31)$$

Для частот, удовлетворяющих условию $\omega = \omega_2 - \omega_1$, гармоническая зависимость от времени исчезает. В этом случае имеем

$$B_{12} = \frac{\pi q^2 \hbar}{m^2 \epsilon_0 \bar{n}^2 \hbar \omega} |\langle \Psi_1^*(\mathbf{r}) | \mathbf{p} | \Psi_2(\mathbf{r}) \rangle|^2, \quad (3.3.32)$$

где множитель $\exp(-j\mathbf{k} \cdot \mathbf{r})$ положен равным единице, поскольку те значения \mathbf{r} , при которых $\Psi_2^*(\mathbf{r})$ и $\Psi_1(\mathbf{r})$ дают сколь угодно существенный вклад в интеграл, много меньше λ , т. е. $\mathbf{k} \cdot \mathbf{r} \ll 1$. Импульс \mathbf{p} берется в операторной форме (3.3.6). Выражение в угловых скобках называется матричным элементом импульса M и определяется соотношением, аналогичным соотношению (3.3.3):

$$M = \langle \Psi_1^*(\mathbf{r}) | \mathbf{p} | \Psi_2(\mathbf{r}) \rangle = \int_V \Psi_1^*(\mathbf{r}) \mathbf{p} \Psi_2(\mathbf{r}) d^3r. \quad (3.3.33)$$

Таким образом, выражение для вероятности перехода принимает вид

$$B_{12} = \frac{\pi q^2 \hbar}{m^2 \epsilon_0 \bar{n}^2 \hbar \omega} |M|^2. \quad (3.3.34)$$

Определяемую соотношением (3.3.33) величину часто называют дипольным матричным элементом, так как она соответствует члену типа $\mathcal{E} \cdot \mathbf{d}$, где \mathbf{d} — дипольный момент. Хотя это не совсем очевидно, все же можно показать [3], что соотношение (3.3.33) определяет взаимодействие диполя с электрическим полем, поэтому название «дипольный матричный элемент» столь же употребительно, как и «матричный элемент импульса». Величину $2|M|^2/\hbar\omega t$ называют силой осциллятора [19].

§ 4. СКОРОСТИ ОПТИЧЕСКОГО ПОГЛОЩЕНИЯ И ИЗЛУЧЕНИЯ В ПОЛУПРОВОДНИКАХ

Плотность состояний

Рассмотренные в § 2 и 3 настоящей главы скорости поглощения и излучения описывают переходы в замкнутой системе между двумя дискретными энергетическими уровнями, разделенными энергетическим интервалом $\hbar\omega_{21}$. Теперь необходимо распространить полученные выражения на случай переходов в полупроводнике, когда имеется множество состояний в валентной зоне и зоне проводимости. Используем одноэлектронное приближение и переходы будем рассматривать как изменение состояния одного электрона. Элементарное обсуждение одноэлектронного приближения приведено во многих книгах, например в работах Блэйкмора [21], Мак-Келви [22] и Бьюба [23], поэтому мы не будем останавливаться на этом вопросе.

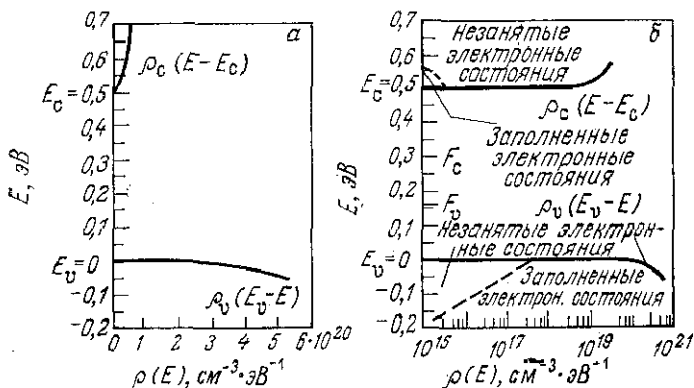


Рис. 3.4.1. *а* — зависимость плотности состояний в параболической зоне от энергии для $m_n = 0,07m_0$ и $m_p = 0,5m_0$; *б* — распределение дырок в валентной зоне и электронов в зоне проводимости для $p = 1 \cdot 10^{16} \text{ см}^{-3}$ и $n = 1 \cdot 10^{14} \text{ см}^{-3}$ при $T = 297 \text{ К}$

Плотность одноэлектронных состояний, приходящихся на единичный интервал энергии, в зоне проводимости определяется известным параболическим выражением [22]

$$\rho_c(E - E_c) = (2\pi^2)^{-1} (2m_n/\hbar^2)^{3/2} (E - E_c)^{1/2}, \quad (3.4.1)$$

где m_n — эффективная масса электронов в зоне проводимости, а E_c — край зоны проводимости. Аналогично для валентной зоны [22]

$$\rho_v(E_v - E) = (2\pi^2)^{-1} (2m_p/\hbar^2)^{3/2} (E_v - E)^{1/2}, \quad (3.4.2)$$

где m_p — эффективная масса дырок в валентной зоне, а E_v — край валентной зоны. На рис. 3.4.1, *а* в качестве примера приведена зависимость плотности состояний от энергии в зоне проводимости с эффективной массой электрона $m_n = 0,07 m_0$, где m_0 — масса свободного электрона, и в валентной зоне с эффективной массой дырки $m_p = 0,5 m_0$. Ширина запрещенной зоны E_g равна энергетическому расстоянию $(E_c - E_v)$ между валентной зоной и зоной проводимости. Выбранные значения эффективных масс соответствуют арсениду галлия, который подробно рассматривается ниже в этой главе. При построении графиков принято значение ширины запрещенной зоны $0,5 \text{ эВ}$.

Коэффициент поглощения

В выражении (3.2.17) скорость переходов вверх между состояниями 1 и 2 полагалась пропорциональной вероятности f_1 того, что состояние с E_1 занято электроном. Скорость переходов между валентной зоной и зоной проводимости полупроводника

зависит от плотности заполненных состояний в валентной зоне $\rho_1(E)f_1$. Аналогично вероятность $[1 - f_2]$ того, что состояние с E_2 свободно, переходит в плотность незаполненных состояний в зоне проводимости $\rho_2(E)(1 - f_2)$. Вероятности заполнения определяются выражениями (3.2.18) и (3.2.19). Плотности заполненных и незаполненных состояний записываются в виде

$$\rho_1(E)f_1 = \rho_v(E_v - E)f_v, \quad (3.4.3)$$

$$\rho_2(E)[1 - f_2] = \rho_c(E - E_c)[1 - f_c]. \quad (3.4.4)$$

На рис. 3.4.1, б приведены эти плотности состояний для значений эффективных масс, использованных при построении графиков на рис. 3.4.1, а. В валентной зоне $p = 1 \cdot 10^{16} \text{ см}^{-3}$, а в зоне проводимости $n = 1 \cdot 10^{14} \text{ см}^{-3}$. Кривые на рис. 3.4.1, б, соответствующие плотностям заполненных и незаполненных состояний, строились по данным, представленным на рис. 3.4.1, а. При этом значения $\rho_v(E_v - E)$ и $\rho_c(E - E_c)$ умножались соответственно на значения f_v , определенные по формуле (3.2.18), и f_c , определенные по формуле (3.2.19). Положение квазиуровней Ферми определялось способом, описанным в § 3 гл. 4. Для того чтобы одновременно представить на графике плотности заполненных и незаполненных состояний, необходимо использовать полулогарифмический масштаб.

Коэффициент поглощения пропорционален разности между скоростями переходов вверх и вниз и для двух дискретных уровней определяется выражением (3.2.39). В отличие от двухуровневой системы в полупроводнике существует континуум состояний в валентной зоне и зоне проводимости, характеризующийся плотностями состояний $\rho_c(E - E_c)$ и $\rho_v(E_v - E)$, которые определяются формулами (3.4.1) и (3.4.2). Поэтому коэффициент поглощения фотонов с энергией $\hbar\omega$ должен равняться сумме коэффициентов поглощения этих фотонов для всех пар энергетических уровней, разделенных энергией $\hbar\omega$. Эту сумму можно представить в виде интеграла

$$\alpha(\hbar\omega) = \int_{-\infty}^{\infty} \frac{B_{12}}{c/\bar{n}} [f_1 - f_2] \rho_v(E_v - E) \rho_c(E - E_c) \delta(E_2 - E_1 - \hbar\omega) dE. \quad (3.4.5)$$

Интеграл (3.4.5) берется от произведения $\rho_v(E_v - E_1)$ на $\rho_c(E_2 - E_c)$ по множеству значений E_2 и E_1 в зоне проводимости и в валентной зоне, разделенных энергией, равной энергии фотона $\hbar\omega$. Именно на такой выбор уровней при интегрировании и указывает дельта-функция $\delta(E_2 - E_1 - \hbar\omega)$.

Поскольку большая часть статей по этому вопросу написана Стерном, мы будем использовать его обозначения. В интеграле, входящем в формулу (3.4.5), потолок валентной зоны удобно

сдвинуть вверх на величину $\hbar\omega$, как это показано в следующей части настоящего параграфа, в которой рассматривается скорость спонтанного излучения. При этом энергию фотона $\hbar\omega$ обозначают через E , энергетическую переменную, по которой происходит интегрирование, — через E' и вводят обозначение $E'' = E' - \hbar\omega$. Краю зоны проводимости E_c будет соответствовать $E' = 0$, а плотности состояний в зоне проводимости и в валентной зоне будут иметь аргументами E' и E'' соответственно. С учетом выражения (3.3.34) для B_{21} формула (3.4.5), определяющая коэффициент поглощения, запишется в следующем виде [8]:

$$\alpha(E) = (\pi q^2 \hbar / \epsilon_0 m^2 c \hbar E) \int_{-\infty}^{\infty} \rho_c(E') \rho_v(E'') |M(E', E'')|^2 [f(E'') - f(E')] dE', \quad (3.4.6)$$

где для системы единиц СГС ϵ_0 надо заменить на $1/4\pi$. Если имеется несколько валентных зон, как, например, в полупроводниках $A^{III}B^V$, то должны учитываться зоны и с тяжелыми, и с легкими дырками. Коэффициент поглощения в этом случае будет равен сумме членов, соответствующих каждой валентной зоне и определяемых формулами, аналогичными формуле (3.4.6).

Скорость спонтанного излучения

Вероятность спонтанного излучения A_{21} связана с B_{12} соотношениями Эйнштейна (3.2.26) и (3.2.27). Из выражения (3.3.34) для B_{12} получаем выражение для A_{21} в виде

$$A_{21} = \frac{4\pi \hbar q^2 E_{21}}{m^2 \epsilon_0 \hbar^2 c^3} |M|^2. \quad (3.4.7)$$

Скорость спонтанного излучения при переходах между двумя дискретными уровнями E_2 и E_1 определена выражением (3.2.21). Так же, как и коэффициент поглощения, скорость спонтанного излучения в полупроводнике зависит от плотности занятых состояний в зоне проводимости и пустых состояний в валентной зоне и может быть представлена в виде [8]

$$r_{\text{спонт}}(E) = (4\pi \hbar q^2 E / m^2 \epsilon_0 \hbar^2 c^3) \times \\ \times \int_{-\infty}^{\infty} \rho_c(E') \rho_v(E'') |M(E', E'')|^2 f(E') [1 - f(E'')] dE', \quad (3.4.8)$$

где, как и в выражении (3.4.6), энергия излучаемого фотона $E = E' - E''$. На рис. 3.4.2, а указаны плотности тех пустых состояний в валентной зоне (дырок) $p(E) = \rho_v(E_v - E) [1 - f_v]$ и занятых состояний в зоне проводимости (электронов) $n(E) = \rho_c(E - E_c) f_c$, которые разделены энергией $\hbar\omega$. Для вычисления интеграла в выражении (3.4.8) кривую, характеризующую

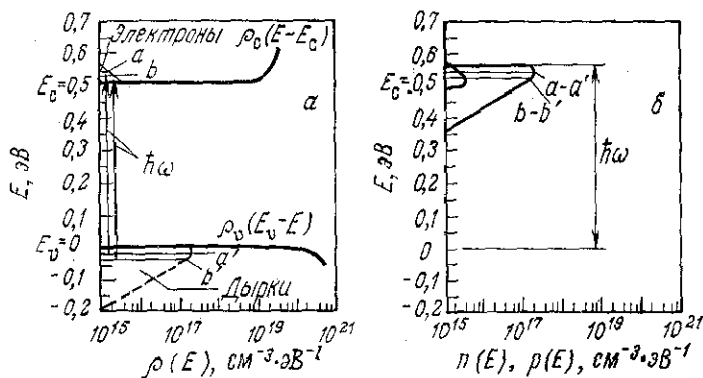


Рис. 3.4.2. *a* — пустые состояния в валентной зоне (дырки) и занятые состояния в зоне проводимости (электроны) при $T = 297$ К. Интеграл в выражении (3.4.8) суммирует излучение при переходах между парами состояний $a - a'$ и $b - b'$, разделенных энергией $\hbar\omega$; *b* — картина, поясняющая вычисление скорости спонтанного излучения по формуле (3.4.8). Кривая, описывающая распределение дырок, сдвинута вверх по оси энергий на величину $\hbar\omega$. Интегрирование происходит в области перекрытия кривых, соответствующих распределениям электронов и дырок.

распределение дырок, сдвигают вверх по оси энергий на величину $\hbar\omega$, как это продемонстрировано на рис. 3.4.2, *b*. Взаимное расположение кривых $p(E)$ и $n(E)$ будет различным при разных значениях энергии фотонов $\hbar\omega$.

Скорость вынужденного излучения

Скорость вынужденного излучения при переходах между двумя дискретными уровнями E_1 и E_2 определена выражением (3.2.35), которое содержит вероятность спонтанного излучения A_{21} . Эта скорость зависит от разности между плотностями занятых состояний в зоне проводимости и в валентной зоне и с учетом выражения (3.4.7) для A_{21} может быть записана следующим образом:

$$r_{\text{вынужд}}(E) = \\ = (4\pi\bar{n}q^2E/m^2\varepsilon_0\hbar^2c^3) \int_{-\infty}^{\infty} \rho_c(E')\rho_v(E'') |M(E', E'')|^2 \{f(E') - f(E'')\} dE'. \quad (3.4.9)$$

За исключением множителей, содержащих фермиевские функции, выражения (3.4.8), определяющее скорость спонтанного излучения, и (3.4.9), определяющее скорость вынужденного излучения, совпадают. Суммарная скорость вынужденного излучения $r_E(\text{вынужд.})$ [см. выражение (3.2.34)] определяется произведением выражения (3.4.9) и числа фотонов в состоянии

с энергией E , которое равно $[\exp(E/kT) - 1]^{-1}$. При выводе выражений (3.4.6), (3.4.8) и (3.4.9) предполагалось, что носители в зоне проводимости и в валентной зоне находятся в состоянии квазиравновесия. В ряде работ с целью объяснения полученных в них отдельных неожиданных результатов предполагалось отклонение состояния носителей от квазиравновесного [24—27]. Однако эквивалентность при комнатной температуре усредненных по времени спектров и спектров, снятых с временным разрешением, указывает на то, что квазиравновесие в зоне проводимости и в валентной зоне все же существует [28].

Проблема вычисления $\alpha(E)$, $r_{\text{спонт}}(E)$ и $r_{\text{вынужд}}(E)$ сводится к нахождению $\rho_c(E')$, $\rho_v(E'')$ и $|M(E', E'')|^2$. При высоких концентрациях примеси или высоких уровнях накачки, характерных для полупроводниковых лазеров, плотность состояний зависит от концентраций и не может быть представлена простыми параболическими выражениями (3.4.1) и (3.4.2). Зависимость плотности состояний от концентрации примесей рассматривается в следующем параграфе, а проблема вычисления матричного элемента обсуждается в § 6 настоящей главы.

§ 5. ЗАВИСИМОСТЬ ПЛОТНОСТИ СОСТОЯНИЙ ОТ КОНЦЕНТРАЦИИ ПРИМЕСЕЙ

Вводные замечания

Примеси в полупроводниках обычно рассматриваются как локализованные уровни с энергией, фиксированной по отношению к валентной зоне или зоне проводимости. Однако в 1949 г. Пирсон и Бардин показали [29], что энергия ионизации бора в Si уменьшается при увеличении концентрации примеси. Было доказано, что такое поведение характерно как для доноров [30], так и для акцепторов [31] в GaAs. Поскольку полупроводниковые лазеры часто изготавливаются из сильно легированного материала и работают при больших концентрациях инжектированных носителей, необходимо рассмотреть влияние этих больших концентраций на плотность состояний. При концентрациях примесей, характерных для полупроводниковых лазеров, нельзя описывать примеси локализованными уровнями, энергии которых отделены от краев зон некоторым промежутком. В этом случае произвольное распределение заряженных примесей в кристалле приводит к флуктуациям потенциала, которые создают хвосты плотности состояний в зоне проводимости и в валентной зоне [4, 5, 32].

Доноры с мелкими уровнями в GaAs, такие, как S, Se и Si, имеют энергию ионизации E_D , равную 0,006 эВ при низких концентрациях [33]. Измерения эффекта Холла, выполненные

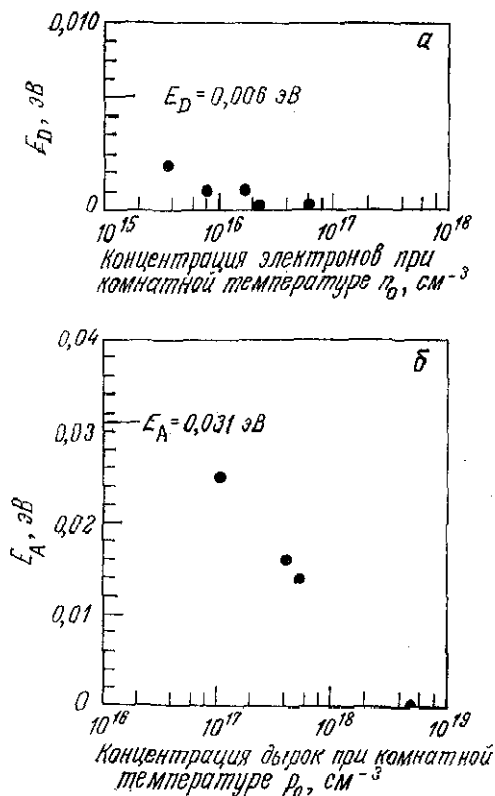


Рис. 3.5.1. *а* — зависимость энергии ионизации мелкого донора от концентрации свободных электронов в GaAs [30]; *б* — зависимость энергии ионизации акцептора (Zn) от концентрации свободных дырок в GaAs [31].

Емельяненко и др. [30], показали, что энергия ионизации доноров стремится к нулю при концентрациях свободных электронов, примерно равных $2 \cdot 10^{16} \text{ см}^{-3}$, как это показано на рис. 3.5.1, *а*. Для Zn, являющегося акцептором с мелким уровнем в GaAs, энергия ионизации при низкой концентрации равна 0,031 эВ [34]. Приведенные на рис. 3.5.1, *б* результаты Эрманиса и Вольфштира [31] показывают, что E_A для Zn становится равной нулю при концентрации дырок в интервале $1 - 5 \times 10^{18} \text{ см}^{-3}$. Эти результаты подчеркивают тот факт, что примеси в полупроводниках, таких, как GaAs, не могут быть описаны обычной моделью локализованных уровней до тех пор, пока их концентрация не становится достаточно низкой.

Когда энергия ионизации донора или акцептора становится равной нулю, концентрация носителей перестает зависеть от температуры [30, 31], и все атомы донора или акцептора (пред-

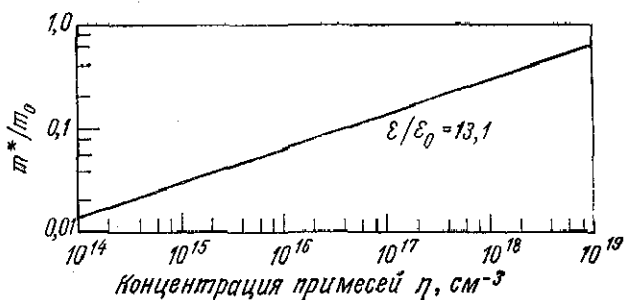


Рис. 3.5.2. Определяемая соотношением (3.5.1) концентрация примесей η , при которой происходит переход в состояние с примесной проводимостью металлического типа, как функция отношения эффективной массы плотности состояний к массе свободного электрона m^*/m_0 .

полагается, что они не образуют комплексов) ионизируются. Поэтому в случае малой компенсации концентрация свободных электронов n_0 будет равна концентрации доноров (а в материале p -типа концентрация свободных дырок p_0 — концентрации акцепторов). Лучшим критерием для определения той концентрации примесей, при которой их энергия ионизации становится равной нулю, является возникновение примесной проводимости металлического типа. Было предсказано [35], что такой переход происходит, когда отношение среднего расстояния r между атомами примеси к радиусу a^* водородоподобного уровня примеси примерно равно трем:

$$r/a^* \approx 3,0. \quad (3.5.1)$$

В этом соотношении r равно $(3/4\pi\eta)^{1/3}$, где η — концентрация примеси, а a^* есть $(\epsilon/\epsilon_0)(m_0/m^*) 0,5 \cdot 10^{-8}$ см, причем m^* соответствует электронной эффективной массе плотности состояний в случае доноров или дырочной эффективной массе плотности состояний в случае акцепторов. На рис. 3.5.2 представлена зависимость m^*/m_0 от η , полученная из соотношения (3.5.1) при $\epsilon/\epsilon_0 = 13,1$ (GaAs, [36]). Большинство обычно используемых полупроводников имеют такое же отношение ϵ/ϵ_0 , но разные эффективные массы. Для p -GaAs с $m_p/m_0 = 0,48$ [37] η , как следует из рис. 3.5.2, равно $\sim 4 \cdot 10^{18}$ см^{-3} , в то время как для n -GaAs с $m_n/m_0 = 0,067$ [37] η равно $\sim 1 \cdot 10^{16}$ см^{-3} . Эти значения находятся в разумном согласии с теми значениями концентраций на рис. 3.5.1, при которых E_D и E_A становятся равными нулю. Оставшаяся часть параграфа посвящена рассмотрению плотности состояний в зоне проводимости и в валентной зоне при концентрациях примесей достаточно высоких, чтобы их энергии ионизации примерно равнялись нулю.

Длина экранирования

При вычислении зависящей от концентрации плотности состояний длина экранирования является одной из главных величин, которые надо определить. При высоких концентрациях примесей свободные носители экранируют кулоновское притяжение (или отталкивание), существующее между ионизованными примесями. Длина экранирования находится из решения уравнения Пуассона в случае экранирования точечного заряда, причем на расстояниях достаточно больших, чтобы не учитывать квантово-механические эффекты. Электронную концентрацию обычно получают путем интегрирования выражения (3.4.3). Для зоны проводимости имеем [21]

$$n = \int_{-\infty}^{\infty} \rho_c(E) f_c(E - F) dE. \quad (3.5.2)$$

Для системы невзаимодействующих электронов, находящихся в состоянии теплового равновесия в поле с переменным в пространстве электростатическим потенциалом $V(r)$, их концентрация в некоторой точке r дается выражением [38]

$$n = \int_{-\infty}^{\infty} f(E - F) \rho[E + qV(r)] dE. \quad (3.5.3)$$

Удобно положить $\epsilon = E + qV$, так что соотношение (3.5.3) примет вид

$$n = \int_{-\infty}^{\infty} f(\epsilon - qV - F) \rho(\epsilon) d\epsilon, \quad (3.5.4)$$

а функция распределения Ферми запишется следующим образом:

$$f(\epsilon - qV - F) = 1 / \{1 + \exp[\epsilon - (F + qV)] / kT\}. \quad (3.5.5)$$

Потенциал при наличии N_D^+ ионизованных доноров определяется из уравнения Пуассона

$$\nabla^2 V(r) = (q/\epsilon) [n(V) - N_D^+] - (q/\epsilon) \delta(r), \quad (3.5.6)$$

где $\nabla^2 = (\partial^2/\partial r^2) + (2/r)(\partial/\partial r)$ для сферически симметричного случая и дельта-функция $q\delta(r) = (q/r^2)\delta(r)$ соответствует положительному точечному заряду в точке $r = 0$.

Для того чтобы решить уравнение (3.5.6), представим $V(r)$ в виде суммы постоянного потенциала V_0 и переменного V_1 : $V(r) = V_0 + V_1$. Уравнение (3.5.6) запишется следующим образом:

$$\nabla^2 (V_0 + V_1) = (q/\epsilon) [n(V_0 + V_1) - N_D^+] - (q/\epsilon) \delta(r). \quad (3.5.7)$$

Когда V_1 равно нулю, $n(V_0) = N_D^+$. Разлагая $n(V_0 + V_1)$ в ряд Тейлора и оставляя только первые два члена, получаем

$$n(V_0 + V_1) \approx N_D^+ + V_1 \left. \frac{\partial n}{\partial V} \right|_{V_0}. \quad (3.5.8)$$

Учет в разложении $n(V_0 + V_1)$ только первых двух членов приводит к «линейному экранированию», поскольку в этом случае мы оставляем только члены, линейные по V_1 . Подстановка выражения (3.5.8) в уравнение (3.5.7) дает

$$\frac{d^2 V_1}{dr^2} + \frac{2}{r} \frac{dV_1}{dr} - \frac{q}{\epsilon} \left(\frac{\partial n}{\partial V} \right) \Big|_{V_0} V_1 = - \frac{q}{\epsilon r^2} \delta(r). \quad (3.5.9)$$

Предполагая, что электростатический потенциал описывается выражением $(q/\epsilon)(1/r)\exp(-r/L_s)$ [39], из уравнения (3.5.9) имеем

$$1/L_s^2 = \frac{q}{\epsilon} \left. \frac{\partial n}{\partial V} \right|_{V_0}. \quad (3.5.10)$$

Из этого соотношения и выражения (3.5.4) следует

$$1/L_s^2 = \frac{q}{\epsilon} \int_{-\infty}^{\infty} \rho(\epsilon) \left[\frac{\partial}{\partial V} f(\epsilon - qV - F) \right]_{V_0} d\epsilon, \quad (3.5.11)$$

откуда для $V_0 = 0$ получаем

$$1/L_s^2 = \frac{q^2}{\epsilon kT} \int_{-\infty}^{\infty} f(E - F) [1 - f(E - F)] \rho(E) dE. \quad (3.5.12)$$

В этой формуле $f(E - F)$ — функция распределения Ферми — Дирака для электронов в зоне проводимости или дырок в валентной зоне [см. выражения (3.2.18), (3.2.19)], а $\rho(E)$ — соответствующая плотность состояний.

При вычислении интеграла в формуле (3.5.12) обычно встречаются две ситуации. Первая соответствует случаю, когда $|F| \ll 1$, так что [21]

$$\int_{-\infty}^{\infty} f_i \rho_i(E) dE = n, \quad (3.5.13)$$

где n — концентрация свободных электронов (дырок). При этом

$$L_s = (kT\epsilon/q^2 n)^{1/2}; \quad (3.5.14)$$

эта величина называется длиной экранирования Дебая. Зависимость L_s от концентрации свободных носителей при $T = 297$ К

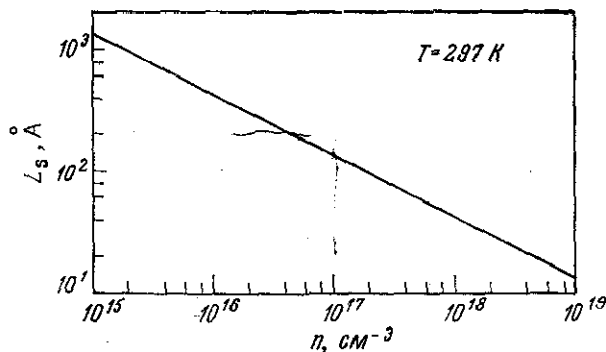


Рис. 3.5.3. Зависимость длины экранирования Дебая от концентрации свободных носителей.

представлена на рис. 3.5.3. Другая ситуация соответствует низким температурам, где $[f_i(1-f_i)]/kT$ равно нулю везде, кроме случая $E = F$, т.е. является дельта-функцией. В этом случае из формулы (3.5.12) следует

$$1/L_s^2 = (q^2/\epsilon) \rho_i(F), \quad (3.5.15)$$

где $\rho_i(F)$ определяется соотношениями (3.4.1) или (3.4.2) при $E = F$. Уровнем Ферми F называется энергетический уровень, вплоть до которого при $T = 0$ К заполнены все состояния в параболической зоне. Он определяется формулой

$$F = (\hbar^2/2m^*) (3\pi^2 n)^{2/3}. \quad (3.5.16)$$

Из формулы (3.5.15) для низкотемпературной длины экранирования получаем

$$L_s = (\pi \epsilon \hbar^2 / m^* q^2)^{1/2} (\pi / 3n)^{1/6}. \quad (3.5.17)$$

Приведенные выше выражения получены в системе единиц МКС; чтобы получить соответствующие выражения в системе единиц СГС, их необходимо умножить на $1/(4\pi)^{1/2}$ (ср. например, выражение (3.5.17) и выражение (5) в работе [4]).

Оценка плотности состояний, зависящей от концентрации

В проводимом ниже обсуждении вопроса о плотности состояний, зависящей от концентрации, основное внимание обращено на определение соответствующих параметров и оценку их величин. В виде примера приводятся типичные значения для GaAs. Во многих работах [4, 5, 32, 40—42] получены выражения, описывающие плотность состояний в полупроводниках, легирован-

ных достаточно сильно, чтобы примесные уровни слились с прилегающими зонами. Выражение (3.5.1) дает критерий для такого слияния примесных уровней с зонами, а определяемая им критическая концентрация примеси представлена графически на рис. 3.5.2 как функция эффективной массы. В предложенной Кейном [4] модели хвостов зон принимается, что носители имеют достаточно низкую кинетическую энергию, позволяющую им следовать за обусловленными наличием ионизованных примесей флуктуациями потенциала. При флуктуациях потенциальной энергии с гауссовым законом распределения, при котором среднеквадратичное отклонение равно $V_{\text{ср. кв}}$, плотность состояний в зоне проводимости полупроводника n -типа имеет вид [4]

$$\rho_c(E') = (2\eta_c)^{1/2} (m_n^{3/2} / \pi^2 \hbar^3) y[(E' - E_c) / \eta_c], \quad (3.5.18)$$

где m_n — эффективная масса плотности состояний в зоне проводимости, E_c — край номинальной (параболической) зоны проводимости, а $\eta_c = \sqrt{2} V_{\text{ср. кв}}$. Безразмерная функция $y[(E' - E_c) / \eta_c]$ определяется выражением

$$y(x) = \pi^{-1/2} \int_{-\infty}^x (x - z)^{1/2} \exp(-z^2) dz; \quad (3.5.19)$$

ее график приведен на рис. 3.5.4. Из выражения (3.5.19) и рис. 3.5.4, *а* следует, что при высоких энергиях ρ_c ведет себя как $(E - E_c)^{1/2}$, в то время как при низких энергиях, как это видно из того же выражения и рис. 3.5.4, *б*, ρ_c имеет гауссову форму $\exp[-(E' - E_c)^2 / \eta_c^2]$. Плотность состояний в валентной зоне полупроводника n -типа определяется выражением, аналогичным выражению (3.5.18), и получается заменой m_n на m_p и изменением знака аргумента у y . Плотность состояний в полупроводнике p -типа получается таким же образом, как и в полупроводнике n -типа.

Если флуктуации потенциала возникают из-за кулоновских потенциалов, произвольным образом размещенных по узлам решетки ионизованных доноров и акцепторов с концентрациями N_D^+ и N_A^- соответственно, то среднеквадратичное отклонение $V_{\text{ср. кв}}$ определяется выражением [4, 43]

$$V_{\text{ср. кв}} = (q^2 / 4\pi\epsilon) [2\pi (N_D^+ + N_A^-) L_s]^{1/2}, \quad (3.5.20)$$

где L_s — длина экранирования, определяемая формулой (3.5.12). В общем случае L_s зависит от концентрации носителей и температуры, а также от конкретного вида зависимости плотности состояний от энергии. Как следует из выражений (3.5.12),

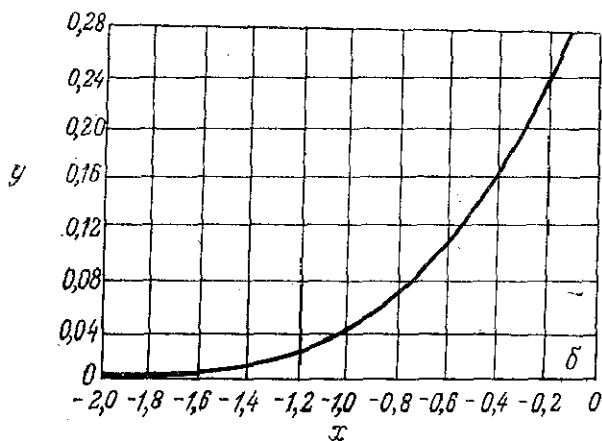
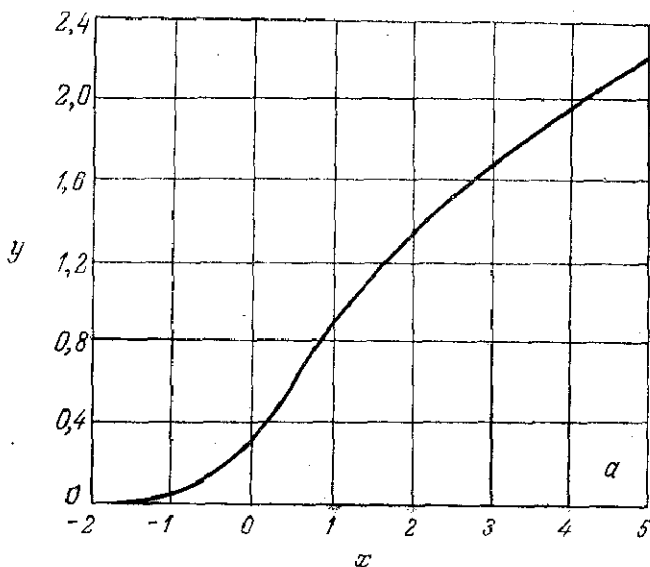


Рис. 3.5.4. Зависимость входящей в модель хвостов зон Кейна безразмерной величины y от $x = (E' - E_c)/\eta_c$ [4].

(3.5.18) и (3.5.20), задачу о нахождении L_s и ρ_c или ρ_v необходимо решать самосогласованным образом. Однако задача облегчается тем, что во многих случаях L_s удовлетворительно описывается выражением (3.5.14) или для нахождения ρ_i и L_s необходима лишь одна итерация.

Для иллюстрации модели хвостов зон Кейна [4] была вычислена зависящая от концентрации плотность состояний ν

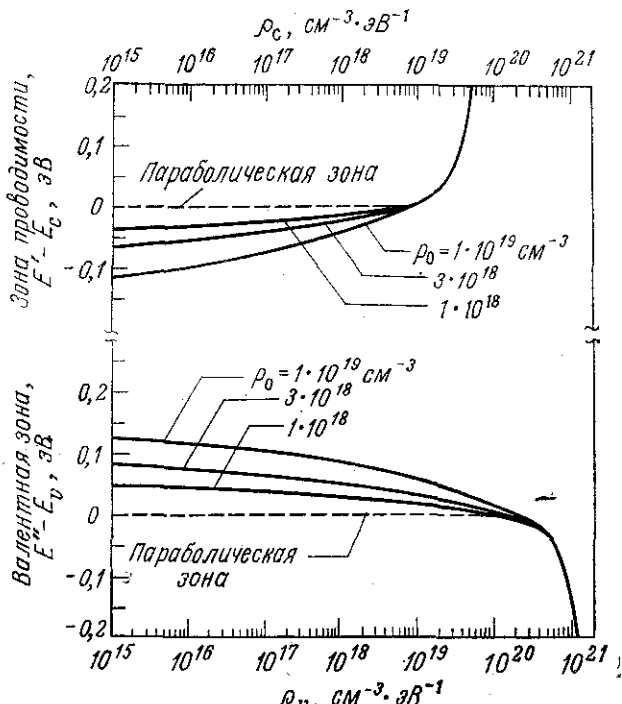


Рис. 3.5.5. Плотность состояний в хвостах зон GaAs, полученная из модели Кейна для указанных концентраций дырок при $T = 297$ К.

GaAs. При этом не совсем ясен выбор значения эффективной массы, поскольку, по-видимому, нет такой теории, описывающей хвосты зон, которая учитывала бы детальную структуру энергетических зон. Поэтому зоны предполагались изотропными, невырожденными и параболическими. При вычислениях, проводимых для GaAs, непараболическость зоны проводимости учитывается заменой эффективной массы $m_n = 0,067 m_0$ [37], соответствующей дну зоны проводимости, на несколько большую электронную массу $m_n = 0,07 m_0$, характерную для типичного значения энергии электрона внутри зоны проводимости. Аналогичная процедура использовалась при выборе эффективной массы легких дырок, где бралось значение $m_{p_l} = 0,085 m_0$. Масса тяжелых дырок m_{p_h} бралась равной $0,55 m_0$. На рис. 3.5.5 показаны результаты вычислений при термически равновесных концентрациях дырок $\rho_0 = N_A^- = 1 \cdot 10^{18}$, $3 \cdot 10^{18}$, $1 \cdot 10^{19}$ и $N_D^+ = 0$, а в табл. 3.5.1 приведены соответствующие значения длины экранирования и входящего в гауссиан параметра η_0 .

Таблица 3.5.1

Параметры для вычисления плотности состояний
в модели Кейна [4]

| Концентрация дырок $p_0, \text{см}^{-3}$ | Длина экранирования $L_s, \text{см}$ [формула (3.5.17)] | Параметр хвоста зоны $\eta_D, \text{эВ}$ |
|---------------------------------------------|------------------------------------------------------------|---------------------------------------------|
| $1 \cdot 10^{18}$ | $17,5 \cdot 10^{-8}$ | $17,1 \cdot 10^{-3}$ |
| $3 \cdot 10^{18}$ | $14,6 \cdot 10^{-8}$ | $27,0 \cdot 10^{-3}$ |
| $1 \cdot 10^{19}$ | $11,9 \cdot 10^{-8}$ | $44,5 \cdot 10^{-3}$ |

Используемая в модели Кейна длина экранирования определяется выражением (3.5.17) [4]. Рис. 3.5.5 показывает, что на расстоянии $\sim 0,05-0,1$ эВ от краев параболических зон еще существует значительная плотность состояний.

Вычисление плотности состояний по формуле (3.5.18) не учитывает туннелирования электронов и дырок через потенциальные барьеры и поэтому дает несколько завышенные значения, если только носители не имеют большой эффективной массы. Кинетическая энергия носителей учтена в работе Гальперина и Лэкса [5], которые привели численные результаты по плотностям состояний в хвосте зоны в ограниченном интервале энергий. Степень уменьшения размытости краев зон из-за кинетической энергии носителей зависит от отношения «кинетической энергии» $\hbar^2/2m^*L_s^2$ (здесь m^* обозначает m_n или m_p) к $V_{\text{ср. кв}}$ [42, 44—46].

Гальперин и Лэкс [5] получили формулу, выражающую плотность состояний в хвосте зоны $\rho_t(E)$ через две функции $a(v)$ и $b(v)$. Эта формула имеет вид

$$\rho_t(E) = (Q^3/E_Q) [a(v)/(\xi')^2] \exp[-b(v)/2\xi'], \quad (3.5.21)$$

где Q — обратная величина длины экранирования, определяемой выражением (3.5.12):

$$Q = 1/L_s. \quad (3.5.22)$$

Величины, входящие в формулу (3.5.21), имеют следующий смысл:

$$E_Q = \hbar^2 Q^2 / 2m^*, \quad (3.5.23)$$

$$v = (E_0 - E) / E_Q, \quad (3.5.24)$$

$$\xi = V_{\text{ср. кв}}^2 = q^4 (N_A^- + N_D^+) / (8\pi e^2 Q), \quad (3.5.25)$$

$$E_0 = -q^2 |N_A - N_D| / \epsilon Q^2. \quad (3.5.26)$$

В последнем выражении $|N_A - N_D|$ — концентрация свободных носителей. Параметр ξ' определялся как ξ/E_Q^2 , что с учетом соотношений (3.5.23) и (3.5.25) дает

$$\xi' = q^4 (N_A^- + N_D^+) (m^*)^2 / 2\pi e^2 \hbar^4 Q^3. \quad (3.5.27)$$

Все эти соотношения записаны в системе единиц МКС. На рис. 3.5.6 приведены графики функций $a(v)$ и $b(v)$, аргументом у которых служит безразмерный параметр v .

Численный расчет $\rho_i(E)$ описан в приложении к работе [47]. Одной из величин, которую надо определить в первую очередь, является Q^2 . Она определяется формулой (3.5.12), в которой неизвестны ни плотность состояний $\rho(E)$, ни уровень Ферми F_c или F_v . Однако, аппроксимируя соответствующим образом величину Q^2 , можно получить приближенные выражения для $\rho(E)$ и F , которые затем использовать для нахождения более точного выражения, описывающего Q^2 . Выбор первоначального выражения, аппроксимирующего Q^2 , зависит от того, является ли температура выше или ниже «температуры вырождения» $T_{\text{выр}}$, которая определяется формулой [48]

$$T_{\text{выр}}(K) = (3/\pi)^{2/3} (h^2/8km^*) \rho^{2/3}, \quad (3.5.28)$$

где ρ — концентрация свободных дырок или электронов в материале p - или n -типа соответственно.

Для параболических зон при $T > T_{\text{выр}}$ аппроксимирующее Q^2 выражение получается из соотношения (3.5.14), а при $T < T_{\text{выр}}$ — из соотношения (3.5.17). Использование в каждом из этих случаев соответствующего выражения для Q^2 позволяет вычислить величины E_0 , ξ' и E_Q . Параметр v вычисляется при значениях энергии E , лежащих в интервале $2E_0 < E < E_0$, так что $v > 0$ (E_0 является отрицательной величиной). Приближенное выражение для плотности состояний в хвосте зоны может быть теперь получено из формулы (3.5.21). Самосогласованные решения для Q^2 , $\rho(E)$ и F получаются в результате двух последовательных итераций. Функция $\rho_i(E)$ в модели Гальперина и Лэкса не описывает состояния непосредственно вблизи дна параболической зоны, поэтому обычно необходима гладкая экстраполяция в ограниченном интервале энергий E .

На конкретном примере сравнение форм хвостов зон, рассчитанных по моделям Кейна [4] и Гальперина и Лэкса [5], было проведено Хуанем [49]. На рис. 3.5.7 приведены результаты его расчетов для образца p -типа. Этот пример показывает, что модель Кейна, в которой пренебрегается кинетической энергией локализации носителей, дает завышенные значения плотности состояний в хвосте зоны. Однако гауссову форму

2, 3505
 $1 + \lg b(\nu) \approx 0,10376 (\lg \nu + 3)$

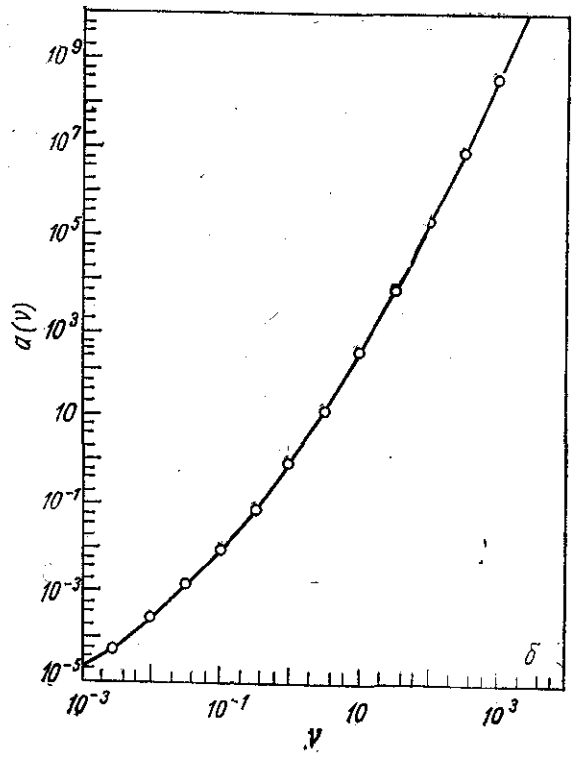
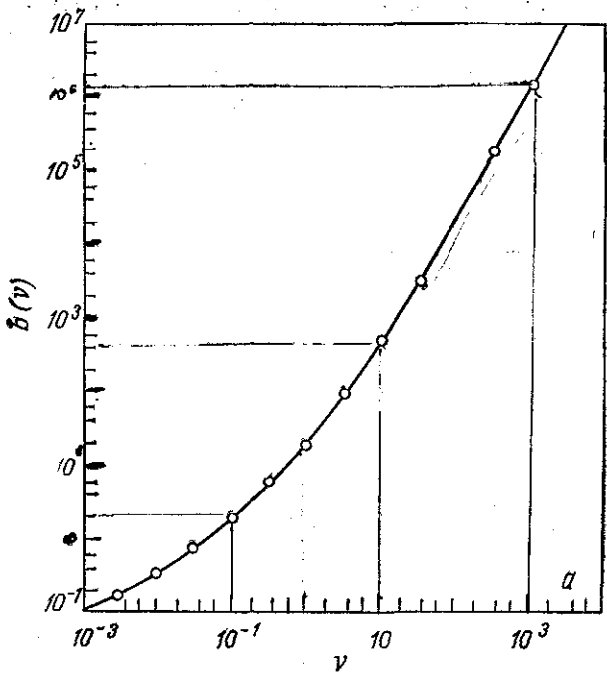


Рис. 3.5.6. Параметры $a(\nu)$ и $b(\nu)$, входящие в выражение (3.5.21), описывающие плотность состояний в хвосте в модели Гальперина и Лэкса [5], как функции $\nu = (E_0 - E)/E_0$.

плотности состояний, получающуюся в модели Кейна, использовать в расчетах значительно легче, чем соответствующее выражение в модели Гальперина — Лэкса.

Стерн [50, 51] предложил для описания плотности состояний выражение, полученное подгонкой в области хвостов зон гауссовой формы плотности состояний модели Кейна к результатам Гальперина — Лэкса. Такую аппроксимацию мы будем в дальнейшем называть моделью хвостов зон Гальперина — Лэкса в гауссовой форме (ГЛГ). Чтобы быть уверенным в том, что при энергиях, существенно превышающих значения, соот-

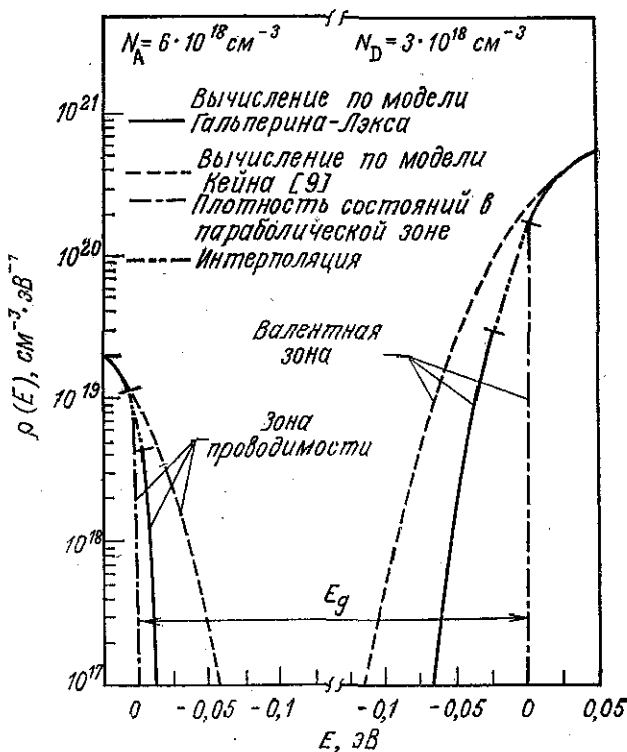


Рис. 3.5.7. Сравнение различных моделей, описывающих плотность состояний, при $N_A = 6 \cdot 10^{18} \text{ см}^{-3}$, $N_D = 3 \cdot 10^{18} \text{ см}^{-3}$ и $T = 300 \text{ К}$. Уровень инжекции в этих расчетах брался примерно равным пороговому при $T = 300 \text{ К}$ для лазера с общими потерями 100 см^{-1} . Ширина запрещенной зоны равна E_g [49].

ветствующие хвосту зоны, полученная плотность состояний будет совпадать с плотностью состояний в невозмущенном кристалле, была использована объемная масса плотности состояний. Характеризующий степень размытости распределения состояний в хвосте зоны параметр η определялся из требования [50, 51], чтобы аппроксимирующее плотность состояний выражение давало тот же результат, что и модель Гальперина — Лэкса, для лежащей в области применимости этой модели энергии, при которой $b(\nu) = 10 \xi'$. Для этой энергии экспоненциальный член в выражении (3.5.21) равен $\exp(-5)$. На рис. 3.5.8 приведены полученные в результате такой подгонки кривые для случая $p_0 = 2 \cdot 10^{18} \text{ см}^{-3}$, $N_A = 1,1 \cdot 10^{19} \text{ см}^{-3}$ и $N_D = 9 \cdot 10^{18} \text{ см}^{-3}$. Другие авторы [47, 49, 52, 53] использовали несколько отличающиеся аппроксимирующие выражения для интерполяции между хвостом плотности состояний в модели Гальперина — Лэкса и плотностью состояний в невозмущенной параболической зоне

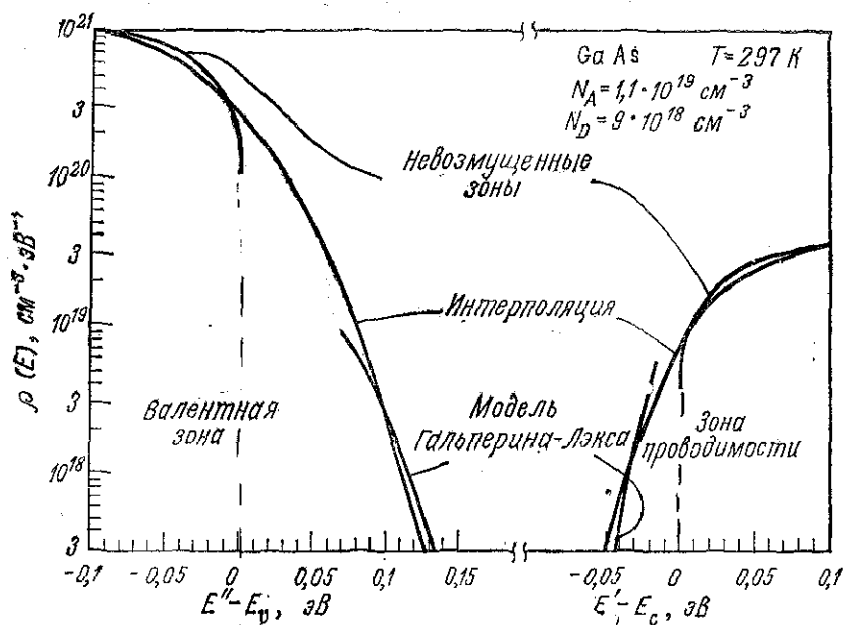


Рис. 3.5.8. Кривые, характеризующие плотность состояний в зоне проводимости и валентной зоне GaAs с концентрацией некомпенсированных акцепторов $2 \cdot 10^{18} \text{ см}^{-3}$. Верхние кривые соответствуют плотностям состояний в невозмущенных зонах, нижние — плотностям состояний в хвостах зон, рассчитанным по теории Гальперина и Лэкса. Соединяющие их кривые являются интерполированными функциями Кейна и соответствуют плотностям состояний, которые используются в модели хвостов зон Гальперина — Лэкса в гауссовой форме [51].

Используемая для нахождения плотности состояний в модели Гальперина — Лэкса длина экранирования сама определяется конкретным видом зависимости плотности состояний от энергии, как это ясно видно из выражения (3.5.12). Поэтому расчет, результаты которого представлены на рис. 3.5.8, проводился самосогласованным образом. Однако при комнатной температуре вид зависимости плотности состояний от энергии относительно слабо влияет на величину L_s , поэтому с малой ошибкой можно пользоваться выражением $L_s = (ekT/\rho q^2)^{1/2}$, соответствующим свободным носителям.

В рассмотренной выше модели Стерна [50, 51] для легких и тяжелых дырок используется одно и то же значение параметра хвоста зоны η . Зона легких дырок должна учитываться в расчетах по оптическому поглощению и излучению, поскольку ее вклад составляет примерно одну треть поглощения (излучения) в слабо легированных образцах, а также влияет на положение края поглощения (на скорость излучения) в сильно легированных образцах. Расщеплением валентной зоны пренебрегается, так как при рассматриваемых энергиях оно не вносит существенного вклада в оптическое поглощение или излучение в GaAs. Дальнейшие детали, касающиеся параметров этой модели, обсуждаются в § 7 настоящей главы, где сравниваются экспериментальные и рассчитанные оптические спектры.

§ 6. ОПТИЧЕСКИЙ МАТРИЧНЫЙ ЭЛЕМЕНТ

Вводные замечания

Из целого ряда вопросов, которые обычно обсуждаются при рассмотрении явления вынужденного излучения, вопросу об оценке величины матричного элемента уделяется, возможно, меньше всего внимания. Чтобы сделать эту оценку, необходимо вычислить интеграл в выражении (3.3.33), а для этого надо знать волновые функции начального и конечного состояний электрона. Вывод выражений, описывающих волновые функции, более уместен в литературе по теории твердого тела, поэтому здесь мы приводим без вывода выражения для соответствующих волновых функций и даем ссылки на работы, содержащие детальные, но довольно сложные выводы этих выражений. К сожалению, физический смысл теряется при необходимых численных расчетах на ЭВМ. Поэтому в настоящем параграфе основное внимание уделяется физической стороне используемых выводов и оценок матричных элементов. Полученные результаты иллюстрируются в § 7 этой главы, где дано сравнение экспериментальных и расчетных значений коэффициента поглощения.

Стандартная теория оптических переходов между валентной зоной и зоной проводимости в прямозонных полупроводниках основывается на так называемом правиле k -отбора [54]. Волновой вектор k_1 волновой функции в валентной зоне и волновой вектор k_2 волновой функции в зоне проводимости должны отличаться между собой на волновой вектор фотона, иначе матричный элемент равен нулю. Поскольку волновой вектор, умноженный на \hbar , дает импульс кристалла, это условие просто выражает закон сохранения импульса при оптическом переходе. Для электрона, имеющего среднюю энергию 0,04 эВ при комнатной температуре, $|k_e|$ примерно составляет $\sim 10^7$ см $^{-1}$, в то время как для инфракрасного фотона с длиной волны 1 мкм $|k_p|$ равно $\sim 6 \cdot 10^4$ см $^{-1}$. Так как $|k_e| \gg |k_p|$, то волновым вектором фотона пренебрегают, и обычно правило k -отбора записывают в виде

$$k_1 \approx k_2. \quad (3.6.1)$$

Таким образом, разрешенные переходы происходят между начальными и конечными состояниями с одинаковыми волновыми векторами и называются «прямыми», или «вертикальными», переходами. Если минимуму зоны проводимости и максимуму валентной зоны соответствуют разные векторы k , то при оптическом переходе для сохранения импульса кристалла необходимо участие фонона, и тогда переход называется «непрямым». Когда в полупроводник добавляются примеси, волновые функции и матричные элементы меняются, и правило k -отбора уже не выполняется [6, 7]. Поэтому рассматриваемые здесь переходы называются оптическими переходами, не подчиняющимися правилу k -отбора; матричный элемент такого перехода зависит от энергии.

Матричный элемент перехода между мелким уровнем и параболической зоной

Вероятно, самым простым случаем, который может быть рассмотрен, является переход между локализованным состоянием, создаваемым мелкой примесью, и параболической зоной. Ряд весьма полезных идей может быть проиллюстрирован этим простым примером. В приближении эффективной массы волновая функция локализованного состояния может быть записана следующим образом [55]:

$$\Psi_1(r) = \psi_{\text{orb}}(r) u_1(r), \quad (3.6.2)$$

где $u_1(r)$ — блоховская функция, соответствующая краю зоны вблизи уровня мелкой примеси. Блоховская функция периодична с периодом решетки и модулирована значительно медленнее

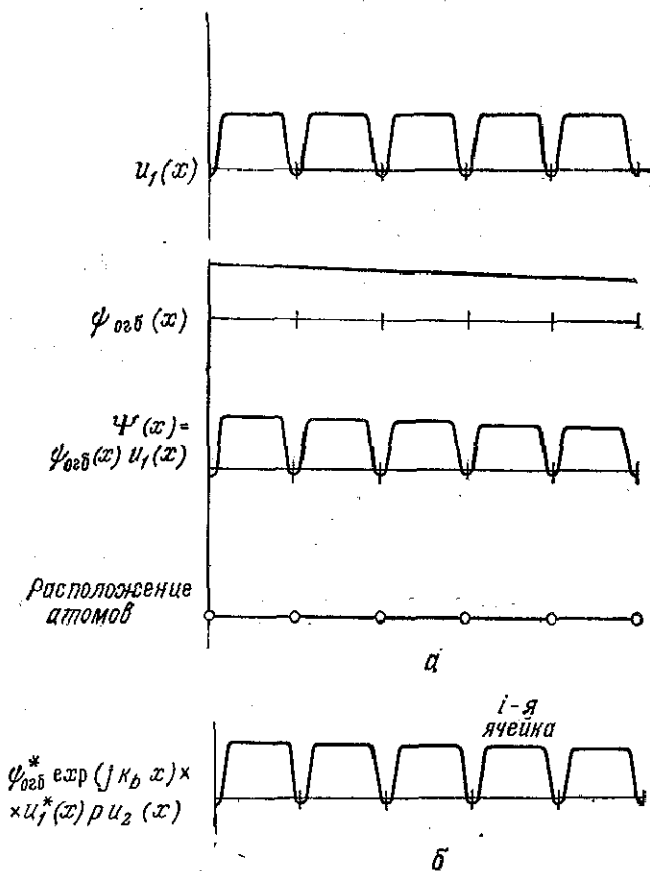


Рис. 3.6.1. *a* — одномерная волновая функция, описывающая локализованное состояние. Блоховская функция, являющаяся периодической функцией с периодом решетки, обозначена через $u_1(x)$, влияние примеси характеризуется функцией $\psi_{0gb}(x)$, а волновая функция Ψ равна произведению $\psi_{0gb}(x)u_1(x)$; *b* — к вычислению первого интеграла в формуле (3.6.5), который равен сумме по всем ячейкам от произведений матричного элемента $u_1^*(x)ru_2(x)$ для каждой ячейки на медленно меняющуюся величину $\psi_{0gb}^*(x)\exp(jk_b x)$.

меняющейся в пространстве огибающей функцией $\psi_{0gb}(\mathbf{r})$, которая характеризует влияние примеси. Огибающая функция $\psi_{0gb}(\mathbf{r})$ экспоненциально спадает в пространстве, но на расстоянии, большом по сравнению с периодом решетки. Она нормирована так, что

$$\int_{-\infty}^{\infty} \psi_{0gb}^*(\mathbf{r}) \psi_{0gb}(\mathbf{r}) d^3\mathbf{r} = 1.$$

Функции Блоха рассматриваются во многих книгах по физике твердого тела (например, в книге Бьюба [56]). На рис. 3.6.1, а схематически изображена одномерная периодическая решетка. Любое соотношение, полученное для функций Блоха в этом одномерном случае, справедливо также для функций Блоха, соответствующих трехмерному случаю. В нижней части рис. 3.6.1, а показано расположение атомов в одномерной решетке. Блоховская функция обозначена через $u(x)$, а $\psi_{\text{орб}}(x)$ представляет собой медленно спадающую экспоненциальную функцию. Волновая функция в параболической зоне берется в виде плоской волны с волновым вектором k_b [56]:

$$\psi_2(\mathbf{r}) = V^{-1/2} u_2(\mathbf{r}) \exp(jk_b \cdot \mathbf{r}), \quad (3.6.3)$$

где V — объем системы, а $u_2(\mathbf{r})$ — блоховская функция в параболической зоне.

Матричный элемент перехода с локализованного примесного (акцепторного) уровня в зону проводимости определяется формулой, получающейся в результате подстановки выражений (3.6.2) и (3.6.3) в формулу (3.3.33):

$$M = V^{-1/2} \int_V \psi_{\text{орб}}^*(\mathbf{r}) u_1^*(\mathbf{r}) \mathbf{p} u_2(\mathbf{r}) \exp(jk_b \cdot \mathbf{r}) d^3\mathbf{r}. \quad (3.6.4)$$

Здесь \mathbf{p} — оператор импульса, определяемый выражением (3.3.6). Поэтому формула (3.6.4) может быть преобразована к виду

$$M = V^{-1/2} \left[\int_V \psi_{\text{орб}}^*(\mathbf{r}) \exp(jk_b \cdot \mathbf{r}) u_1^*(\mathbf{r}) \mathbf{p} u_2(\mathbf{r}) d^3\mathbf{r} + \int_V \psi_{\text{орб}}^*(\mathbf{r}) u_1^*(\mathbf{r}) u_2(\mathbf{r}) \mathbf{p} \exp(jk_b \cdot \mathbf{r}) d^3\mathbf{r} \right]. \quad (3.6.5)$$

Второй интеграл в этом выражении равен нулю, так как функции $u_1(\mathbf{r})$ и $u_2(\mathbf{r})$ ортогональны. Способ оценки первого интеграла поясняется на рис. 3.6.1, б. Этот интеграл равен сумме произведений матричного элемента $u_1^*(\mathbf{r}) \mathbf{p} u_2(\mathbf{r})$, одинакового для всех ячеек, на медленно меняющуюся величину, которая примерно постоянна в каждой ячейке i , причем суммирование ведется по всем ячейкам кристалла:

$$M = V^{-1/2} \sum_i \psi_{\text{орб}}^*(\mathbf{r}) \exp(jk_b \cdot \mathbf{r}) \langle u_1^*(\mathbf{r}) | \mathbf{p} | u_2(\mathbf{r}) \rangle, \quad (3.6.6)$$

или

$$M = V^{-1/2} \langle u_1^*(\mathbf{r}) | \mathbf{p} | u_2(\mathbf{r}) \rangle \int_V \psi_{\text{орб}}^*(\mathbf{r}) \exp(jk_b \cdot \mathbf{r}) d^3\mathbf{r}. \quad (3.6.7)$$

Выражение (3.6.7) получено из выражения (3.6.6) заменой суммы по всем ячейкам на интеграл по объему кристалла.

Интеграл в выражении (3.6.7) может быть назван вкладом в матричный элемент от огибающей функции:

$$M_{орб} = \int_V \psi_{орб}^*(\mathbf{r}) \exp(i\mathbf{k}_b \cdot \mathbf{r}) d^3r. \quad (3.6.8)$$

Величина $\langle u_1^*(\mathbf{r}) | \mathbf{p} | u_2(\mathbf{r}) \rangle$ представляет собой усредненный матричный элемент по блоховским состояниям в зонах в случае отсутствия примесей. Кейн получил выражение, определяющее этот матричный элемент, в виде [57]

$$\langle u_1^*(\mathbf{r}) | \mathbf{p} | u_2(\mathbf{r}) \rangle = -m_0 P / j\hbar \equiv M_b. \quad (3.6.9)$$

Таким образом,

$$M = M_b \cdot M_{орб}. \quad (3.6.10)$$

Для полупроводниковых соединений $A^{III}B^V$ имеем [8, 51]

$$|M_b|^2 = m_0^2 P^2 / 6\hbar^2, \quad (3.6.11)$$

где множитель $1/3$ возникает вследствие усреднения по всем направлениям, а множитель $1/2$ — из-за правила отбора по спину. Выражение (12) работы [57] дает

$$P^2 = \frac{\hbar^2 E_g (E_g + \Delta)}{2m_n \left(E_g + \frac{2}{3} \Delta \right)}, \quad (3.6.12)$$

где E_g — ширина запрещенной зоны, а Δ — энергия спин-орбитального расщепления. С учетом этого соотношения выражение (3.6.11) запишется в виде

$$|M_b|^2 = \frac{m_0^2 E_g (E_g + \Delta)}{12m_n \left(E_g + \frac{2}{3} \Delta \right)}. \quad (3.6.13)$$

Для GaAs с уже использовавшимися выше значениями $m_n = 0,067 m_0$ и $E_g = 1,424$ эВ и с $\Delta = 0,33$ эВ получаем

$$|M_b|^2 = 1,33 m_0 E_g. \quad (3.6.14)$$

Поправки, возникающие вследствие температурных эффектов, а также вклад более высоких зон в эффективную массу, соответствующую дну зоны проводимости, не учитывались.

Для оценки $M_{орб}$, определенного выражением (3.6.8), предположим, что волновая функция локализованного состояния имеет форму 1s-орбитали атома водорода [58]:

$$\psi_{орб}(\mathbf{r}) = \pi^{-1/2} (1/a^*)^{3/2} \exp(-r/a^*), \quad (3.6.15)$$

где $a^* = 4\pi\epsilon\hbar^2/m^*q^2$ — эффективный борковский радиус локализованного состояния. Тогда

$$M_{\text{оргб}} = \frac{V^{-1/2}\pi^{-1/2}}{a^{*3/2}} \int_V \exp(-r/a^*) \exp[i\mathbf{k}_b \cdot \mathbf{r}] d^3\mathbf{r}, \quad (3.6.16)$$

или

$$M_{\text{оргб}} = \frac{V^{-1/2}\pi^{-1/2}}{a^{*3/2}} \int_{r=0}^{\infty} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \exp(-r/a^*) \exp(ik_b r \cos \theta) r^2 \sin \theta dr d\theta d\varphi. \quad (3.6.17)$$

Интегрирование по φ дает множитель 2π . Вводя новую переменную $t = -\cos \theta$ и учитывая, что $\sin \theta d\theta = dt$, а пределы интегрирования $\theta = 0$ и $\theta = \pi$ меняются на $t = -1$ и $t = 1$ соответственно, получаем в результате интегрирования по этой переменной

$$M_{\text{оргб}} = \frac{4\pi^{1/2}V^{-1/2}}{-ik_b a^{*3/2}} \int_0^{\infty} r \exp(-r/a^*) \sin(k_b r) dr. \quad (3.6.18)$$

Вычисление интеграла в этом выражении и использование равенства $|M_{\text{оргб}}|^2 = M_{\text{оргб}}^* M_{\text{оргб}}$ дают

$$|M_{\text{оргб}}|^2 = 64\pi a^{*3} (1 + a^{*2}k_b^2)^{-4} V^{-1}. \quad (3.6.19)$$

Эта формула была получена Игглсом [6] и Думке [7]. Когда переход происходит между мелким акцептором и зоной проводимости, величина k_b равна $(2m_c E'_c/\hbar^2)^{1/2}$, причем E'_c отсчитывается от края зоны проводимости. Матричный элемент $|M|^2$ теперь легко получить, перемножая выражения (3.6.14) и (3.6.19). Поскольку k_b входит в знаменатель выражения (3.6.19), матричный элемент уменьшается при увеличении энергии фотона, так как при этом соответствующее состояние в зоне проводимости удаляется от края зоны. Таким образом, в рассмотренном случае, когда не выполняется правило k -отбора, матричный элемент оптического перехода зависит от энергии фотона.

Матричный элемент, не зависящий от энергии, был использован Лэшером и Стерном [8] в первых детальнейших численных расчетах спектров спонтанного и вынужденного излучений, а также коэффициента усиления в GaAs. Эти начальные расчеты, в которых рассматривались только параболические зоны, послужили отправной точкой для последующих публикаций. Стерн [9] использовал матричный элемент, определяемый произведением выражений (3.6.14) и (3.6.19), и модель хвостов зон Кейна [4] для демонстрации влияния наличия хвостов зон на коэффициент усиления. Хуань [10] модернизировал расчеты Стерна [9], используя при вычислении спонтанных спектров

и коэффициента усиления при высоких уровнях накачки выражение для плотности состояний, получаемое в модели хвостов зон Гальперина — Лэкса [5]. Эти результаты представлены в § 8 настоящей главы, который посвящен рассмотрению вопроса о плотности порогового тока. Расчеты Стерна [9] и Хуаня [10] показывают, что на основе приведенной выше относительно простой модели переходов с несоблюдением правила k -отбора может быть получено много полезных сведений. Стерн [11] предложил затем уточненное, но значительно более сложное выражение для матричного элемента [11]. Оно будет рассмотрено в следующем разделе этого параграфа.

Уточненное выражение для матричного элемента

При рассмотрении сильно легированных полупроводников необходимо учитывать хвосты зон. В области хвостов зон флуктуации потенциала нарушают трансляционную симметрию кристалла и так деформируют волновые функции, что они перестают быть подобными плоским волнам, и поэтому правило k -отбора становится неприменимым. Однако при больших энергиях фотонов, превышающих ширину запрещенной зоны, переходы будут происходить между состояниями невозмущенных зон, где правило k -отбора выполняется. Поэтому Стерн использовал матричный элемент, который при больших энергиях фотонов соответствует переходам, для которых справедливо правило k -отбора, а при малых энергиях описывает не подчиняющиеся этому правилу переходы между локализованными состояниями хвоста зоны и состояниями в невозмущенной зоне, волновые функции которых подобны плоским волнам [11, 51]. Выражение для такого матричного элемента получается при предположении, что огибающая функция имеет вид

$$\psi_{огб} = (\beta^3/\pi)^{1/2} \exp(j\mathbf{k} \cdot \mathbf{r}) \exp(-\beta|\mathbf{r} - \mathbf{r}_i|), \quad (3.6.20)$$

где \mathbf{r}_i — радиус-вектор центра данного состояния, β характеризует степень локализации состояния, а \mathbf{k} , наоборот, является мерой тех свойств этого состояния, которые характерны для плоской волны.

Для вычисления параметров \mathbf{k} и β волновой функции (3.6.20), соответствующей состоянию с энергией E' в возмущенной зоне проводимости, вводится другая энергия E^* . Она определяется так, что число состояний с энергиями, меньшими E^* , при невозмущенной плотности состояний равно числу состояний с энергиями, меньшими E' , при реальной плотности состояний. В этом случае для описания хвоста зоны используется модель Кейна [4] с параметром гауссиана η , выбранным так, чтобы, как это уже было описано в § 5 настоящей главы, плотность состояний для энергии, при которой $b(v) = 10 \xi'$, получа-

лась такая же, как и в модели Гальперина — Лэкса [5]. Разность $E^* - E'$ служит мерой понижения энергии, которое происходит из-за флуктуаций потенциала. Для зоны проводимости параметры \hbar_c и β_c определяются из соотношений

$$\hbar^2 k_c^2 / 2m_c = E^* - E_c, \quad (3.6.21)$$

$$\hbar^2 \beta_c^2 / 2m_c = \hbar_c (E^* - E'), \quad (3.6.22)$$

где E_c соответствует краю номинальной (параболической) зоны проводимости. Коэффициент \hbar_c выбирается так, чтобы величина $\hbar_c (E^* - E')$ равнялась кинетической энергии локализации T [5] для состояния с энергией, при которой значения плотности состояний в гауссовом приближении и в модели Гальперина — Лэкса совпадают.

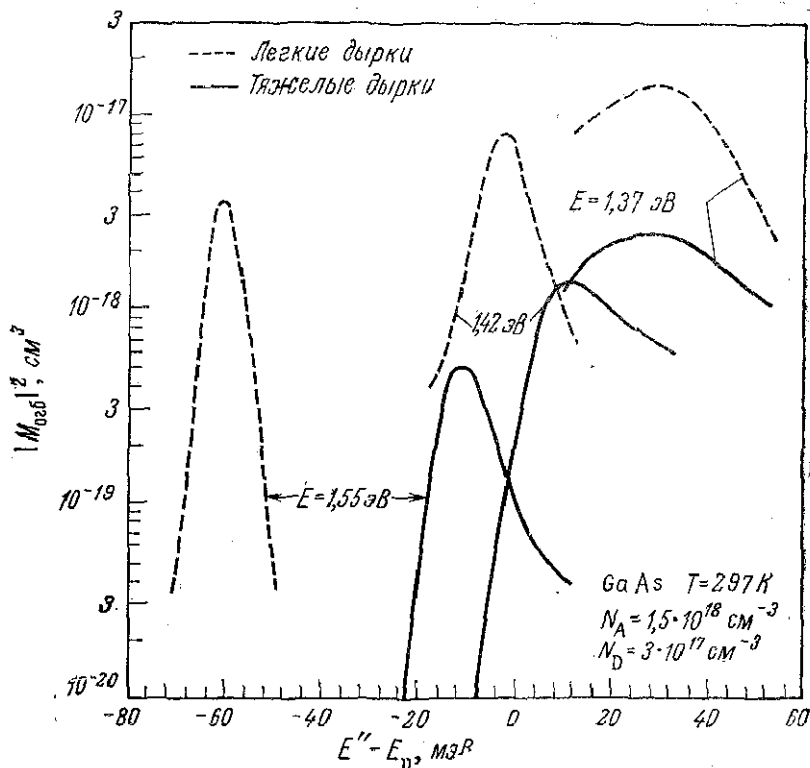


Рис. 3.6.2. Зависимость квадрата модуля матричного элемента $|M_{00rs}|$, соответствующего оптическим переходам между валентной зоной и зоной проводимости GaAs с концентрацией дырок $1,2 \cdot 10^{18} \text{ см}^{-3}$, от энергии начального состояния в валентной зоне. Кривые построены для трех значений энергии фотонов, указанных на рисунке. Переходам из зон тяжелых и легких дырок соответствуют разные семейства кривых [51].

При больших энергиях, таких, что $E^* \approx E'$, выражение (3.6.20) дает обычный результат типа плоских волн. При низких энергиях, где $E^* \approx E_c$, состояние локализовано и имеет соответствующую энергию локализации. При промежуточных энергиях волновая функция (3.6.20) может значительно отличаться от точной огибающей функции.

В выражении (3.6.8), определяющем $M_{огб}$, под знаком интеграла теперь будет стоять произведение функции, комплексно сопряженной с волновой функцией (3.6.20), описывающей состояния в зоне проводимости, на соответствующую волновую функцию, описывающую состояние в валентной зоне, причем это произведение усреднено по всем направлениям волновых векторов и по всем положениям центров локализации. В этом случае [11, 51]

$$|M_{огб}|^2 = (64\pi b/3)(t^4 - q^4)^{-5} [(b^4 - 5b^2B^2 + 5B^4)(3t^4 + q^4)(t^4 - q^4)^2 + 8b^2B^2t^2(3b^2 - 10B^2)(t^8 - q^8) + 16b^4B^4(5t^8 + 10t^4q^4 + q^8)], \quad (3.6.23)$$

где $B = \beta_c\beta_v$, $b = \beta_c + \beta_v$, $t^2 = b^2 + k_c^2 + k_v^2$ и $q^2 = 2k_ck_v$.

На рис. 3.6.2 для образца p -типа с концентрацией дырок $1,2 \cdot 10^{18} \text{ см}^{-3}$ построены для трех разных энергий фотонов зависимости $|M_{огб}|^2$ от энергии состояний в валентной зоне, отсчитанной от края номинальной параболической зоны. Положительные значения $E'' - E_v$ соответствуют состояниям внутри валентной зоны. На рисунке изображены два семейства кривых, одно из которых представляет переходы из зоны тяжелых дырок, а другое — из зоны легких дырок. Следует отметить, что для фотона с наименьшей энергией кривые довольно плавные, в то время как кривые, соответствующие фотону с максимальной энергией, имеют довольно острый пик.

Определенный выражением (3.6.23) матричный элемент был использован в вычислениях оптического поглощения и усиления в GaAs с примесями [50, 59—61], а также с некоторыми изменениями, учитывающими другое строение зонной структуры, в аморфном Si [11]. Детальное сравнение экспериментальных и вычисленных значений коэффициента поглощения в сильно легированном GaAs проведено в работе [51]. Результаты этого сравнения приводятся в следующем параграфе.

§ 7. РАССЧИТАННЫЕ ОПТИЧЕСКИЕ СПЕКТРЫ

Сравнение экспериментальных и рассчитанных спектров поглощения и излучения

Модель [51, 61] для расчета спектров поглощения и излучения в GaAs с концентрациями примесей порядка $1 \cdot 10^{18} \text{ см}^{-3}$ и выше может быть построена на основе результатов, приведен-

ных в предыдущих двух параграфах. В этой модели используется представление зависящей от концентрации плотности состояний в гауссовой форме, причем параметры гауссиана выбираются так, чтобы, как показано на рис. 3.5.8, плотность состояний в хвосте зоны получалась такая же, как в модели Гальперина — Лэкса. Матричный элемент, зависящий от энергии, получается при умножении $|M_b|^2$, определяемого соотношением (3.6.14), на полученное Стерном для $|M_{отб}|^2$ выражение (3.6.23). Это произведение $|M_b|^2|M_{отб}|^2$ в дальнейшем будем называть матричным элементом Стерна (МЭС). Поэтому модель, на основе которой рассчитываются спектры поглощения и излучения, здесь называется моделью ГЛГ — МЭС. Мы сравним вычисленные по формуле (3.4.6) значения $\alpha(E)$ с экспериментальными данными работы [62].

Количественное сравнение рассчитанного по модели ГЛГ — МЭС коэффициента поглощения с экспериментальными значениями демонстрирует, насколько хорошо эта модель описывает реальный прямозонный полупроводник, например такой, как GaAs. Сравнение проводится на примере коэффициента поглощения $\alpha(E)$, так как измерить его значительно легче, чем получить спектры спонтанного или вынужденного излучений. Как следует из полученных выше соотношений (3.2.41), (3.2.45), (3.2.47) и (3.4.6), (3.4.8), (3.4.9), поглощение, спонтанное излучение и вынужденное излучение связаны друг с другом, так что знание плотности состояний и матричного элемента позволяет вычислить каждый из этих оптических спектров. Кроме того, из сравнения экспериментальных и рассчитанных значений коэффициента поглощения может быть получена зависимость ширины запрещенной зоны от концентрации носителей. Эта зависимость может затем быть использована при вычислении спектров излучения.

В табл. 3.7.1 приведены характеристики выращенных из расплава легированных Zn образцов GaAs, для которых проводилось сравнение экспериментальных и рассчитанных спектров поглощения и излучения. В таблице даны также параметры хвостов зон, рассчитанных на основе модели ГЛГ, рассмотренной в § 5 настоящей главы. Значение ширины запрещенной зоны E_g выбиралось исходя из лучшего совпадения вычисленного и измеренного значений края поглощения. Равновесная дырочная концентрация p_0 и подвижность при температуре 297 К определялись из холловских измерений. Для того чтобы гарантировать применимость модели хвостов зон, рассматривались только образцы с $p_0 \geq 1 \cdot 10^{18}$ см⁻³. Как уже говорилось в § 5 этой главы, холловские измерения показывают [31], что энергия ионизации акцепторов становится равной нулю при концентрации дырок в пределах от $1 \cdot 10^{18}$ до $5 \cdot 10^{18}$ см⁻³.

Таблица 3.7.1

Параметры энергетических зон GaAs p -типа при 297 K¹⁾

| p_0 , см ⁻³ | Примесь | Подвижность, см ² (В·с) | N_A , см ⁻³ | N_D , см ⁻³ | E_g , эВ | η_c , мЭВ | η_p , мЭВ | Длина экранирования, Å | $V_{ср.}$ кв. мЭВ | F'_p , мЭВ |
|--------------------------|---------|------------------------------------|--------------------------|--------------------------|------------|----------------|----------------|------------------------|-------------------|--------------|
| $1,2 \cdot 10^{18}$ | Zn | 162 | $1,5 \cdot 10^{18}$ | $3 \cdot 10^{17}$ | 1,408 | 10 | 20 | 39 | 23 | 59 |
| $2,4 \cdot 10^{18}$ | Zn | 102 | $3 \cdot 10^{18}$ | $6 \cdot 10^{17}$ | 1,403 | 9 | 22 | 28 | 28 | 40 |
| $1,6 \cdot 10^{19}$ | Zn | 67 | $2 \cdot 10^{19}$ | $4 \cdot 10^{18}$ | 1,381 | 5,5 | 29 | 13 | 49 | -20 |

1) Здесь $F'_p = F_p - E_p$ — квазиуровень Ферми для дырок, отсчитанный от номинального края валентной зоны; η_c и η_p — параметры хвостов зоны проводимости и валентной зоны, соответственно [51]

Зависимость плотности состояний от концентраций примесей определяется выражениями (3.5.17) — (3.5.19). Выбор значений концентраций акцепторов N_A и доноров N_D при данной концентрации дырок затруднителен. Как видно из табл. 3.7.1, для образцов, легированных Zn, отношение N_A/N_D было взято равным 5,0. Такое значение выбрано на основе результатов Эрманиса и Вольфштира [31]. В области концентраций, рассматриваемых здесь, они дали значения концентраций дырок, определенные по холловским измерениям, и соответствующие значения концентраций Zn, определенные нейтронной активацией. По их измерениям [31], $N_A/N_D \approx 5,0$ для $p_0 = 1,1 \cdot 10^{18}$ и $3,7 \cdot 10^{18}$ см⁻³, и акцепторы (Zn) полностью ионизованы, так что $N_A - N_D = p_0$.

На рис. 3.7.1, а представлена полученная в работе [62] экспериментальная кривая зависимости коэффициента поглощения от энергии фотонов для образца из GaAs, характеристики которого указаны в первой строчке табл. 3.7.1 ($p_0 = 1,2 \cdot 10^{18}$ см⁻³) [51]. Вклад от поглощения на свободных носителях исключен. На этом же рисунке приведена для сравнения кривая коэффициента поглощения образца из слаболегированного GaAs. Две рассчитанные кривые показаны штриховыми линиями. При расчете одной из этих кривых использовались плотность состояний обычной параболической зоны, вычисленный с учетом правила k -отбора и с волновыми функциями в виде плоских волн матричный элемент для прямого перехода, а также значение E_g , равное 1,424 эВ [63]. Эта кривая помечена на рисунке надписью «параболические зоны». Определяемый ею коэффициент поглощения плохо совпадает с экспериментальными значениями $\alpha(E)$ при энергиях меньше 1,45 эВ.

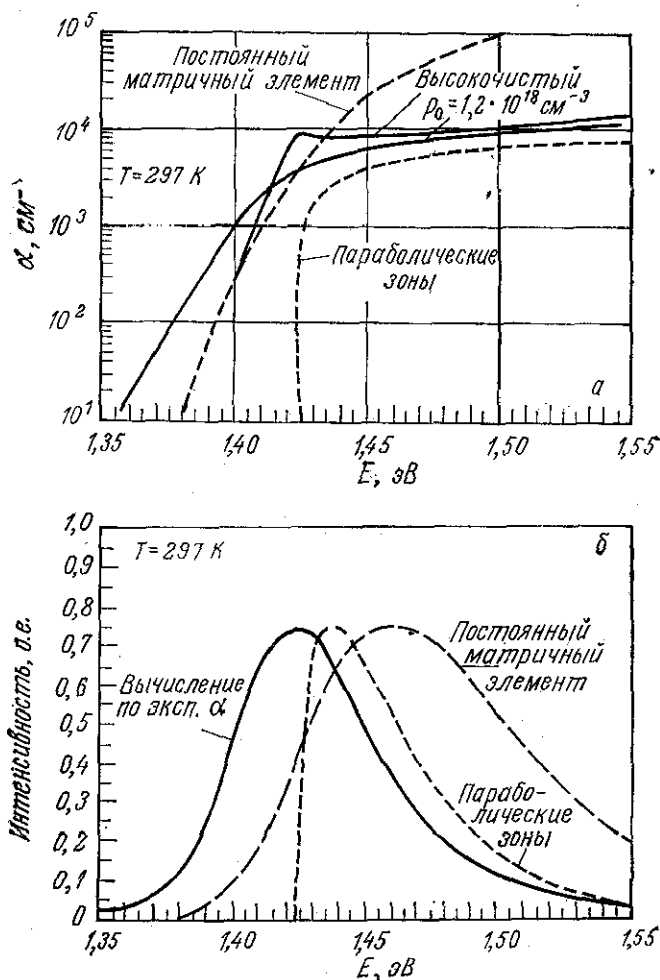


Рис. 3.7.1. *a* — сравнение экспериментальных значений коэффициента поглощения при $\rho_0 = 1,2 \cdot 10^{18}$ см⁻³ [62] и вычисленных с использованием параболической плотности состояний и плотности состояний с хвостами зон, но с постоянным матричным элементом; *б* — спектры спонтанного излучения, рассчитанные с использованием принципа детального равновесия по кривым поглощения рис. *a*. Кривые нормированы к одному и тому же значению интенсивности в максимуме [51].

При получении другой кривой использовались постоянный матричный элемент и плотность состояний, рассчитанная по модели ГЛГ с параметрами хвостов зон, указанными в табл. 3.7.1. Это часто встречающееся предположение о постоянстве матричного

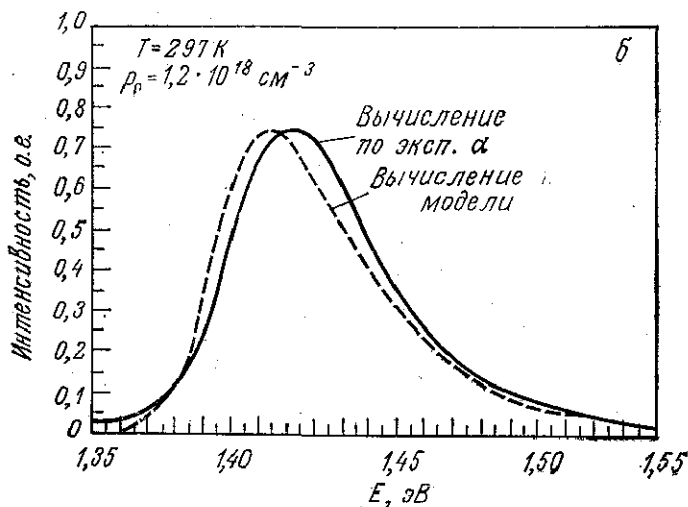
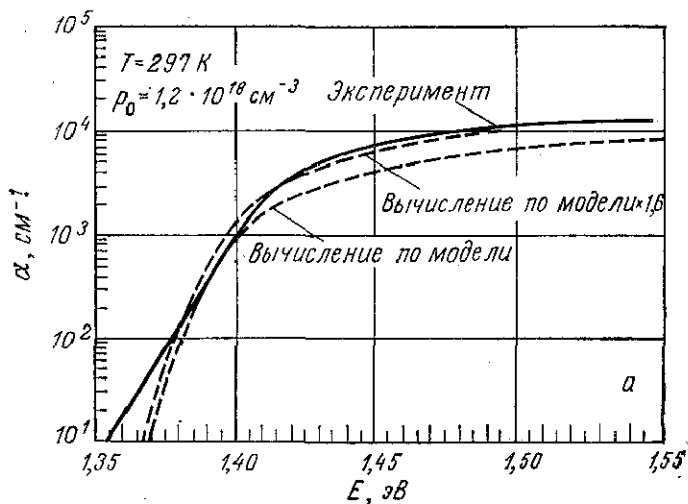


Рис. 3.7.2. а — сравнение экспериментальных [62] и вычисленных значений коэффициента поглощения для $\rho_0 = 1,2 \cdot 10^{18}$ см⁻³; б — сравнение спектров излучения, рассчитанных по формуле (3.7.1) из экспериментальных значений $\alpha(E)$ и по модели [51].

элемента приводит к вычисленным значениям $\alpha(E)$, которые при $E > E_g$ намного превосходят экспериментальные значения.

На рис. 3.7.1, б сравниваются спектры спонтанного излучения [51]. Они рассчитаны по известной зависимости $\alpha(E)$ с помощью основанной на принципе детального равновесия теории

ван Русбрека — Шокли [64]. Согласно этой теории, интенсивность равновесного излучения $\mathcal{J}(E)$ фотонов с энергией E определяется выражением

$$\mathcal{J}(E) = \frac{8\pi\bar{n}^2 E^2 \alpha(E)}{h^3 c^2 [\exp(E/kT) - 1]}, \quad (3.7.1)$$

где $\alpha(E)$ — коэффициент поглощения при энергии фотонов E . Для удобства сравнения форм спектров максимумы спектральных кривых на рис. 3.7.1, б нормированы к одному значению. Согласия со спектром, рассчитанным по экспериментальным значениям $\alpha(E)$ с помощью формулы (3.7.1), ни параболическая зона, ни постоянный матричный элемент не дают. Если $\alpha(E)$ рассчитан по формуле (3.4.6), то для вычисления $r_{\text{спонт}}(E)$ более удобно использовать выражение (3.7.1), чем выражение (3.4.8). Для малых уровней накачки, когда $F_2 - F_1 \ll E$, выражение (3.2.45) для $r_{\text{спонт}}$ имеет тот же вид, что и полученное с использованием принципа детального равновесия выражение (3.7.1).

На рис. 3.7.2, а представлены экспериментальные и рассчитанные по модели ГЛГ — МЭС значения коэффициента поглощения. При расчетах использовались параметры хвостов зон, приведенные в табл. 3.7.1, а матричный элемент в формуле (3.4.6) определялся выражениями (3.6.14), (3.6.23) и (3.6.10). Значение ширины запрещенной зоны выбиралось исходя из наилучшего совпадения расчетной и экспериментальной зависимостей $\alpha(E)$, причем величина E_g влияла только на положение расчетной спектральной кривой поглощения, но не на ее форму. Экспериментальные значения $\alpha(E)$ и концентрация дырок те же, что и на рис. 3.7.1. Расчетные значения $\alpha(E)$ хорошо воспроизводят форму спектральной кривой поглощения, но по абсолютной величине расчет дает интенсивность поглощения примерно в 1,6 раза меньше экспериментальной при энергиях фотонов, превышающих ширину запрещенной зоны. При малых значениях $\alpha(E)$ кривые расходятся. Это говорит о том, что плотность состояний глубоко в хвостах зон или матричный элемент при низких энергиях слишком малы или же что присутствуют другие механизмы уширения. Сравнение спектров излучения приводится на рис. 3.7.2, б. Эти спектры рассчитаны по формуле (3.7.1) из экспериментальных и вычисленных значений коэффициента поглощения.

На рис. 3.7.3, а представлены расчетные значения коэффициентов поглощения в трех образцах, легированных Zn, а на рис. 3.7.3, б для сравнения приведены экспериментальные значения коэффициента поглощения для тех же концентраций дырок [62]. Во всех случаях при низких энергиях расчетные значения $\alpha(E)$ уменьшаются несколько быстрее эксперименталь-

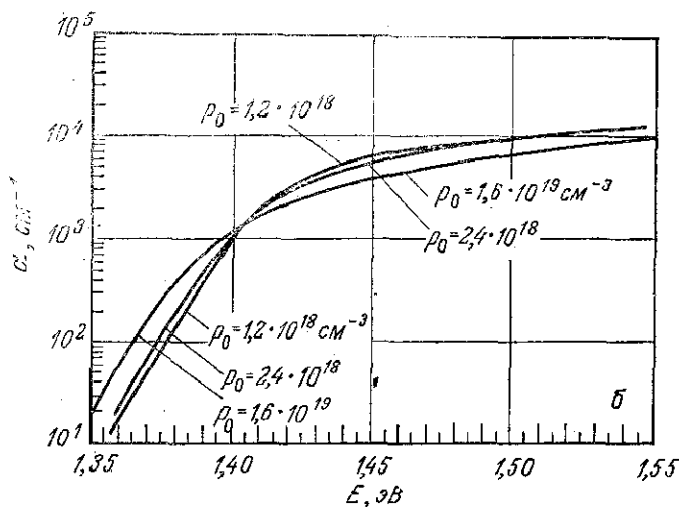
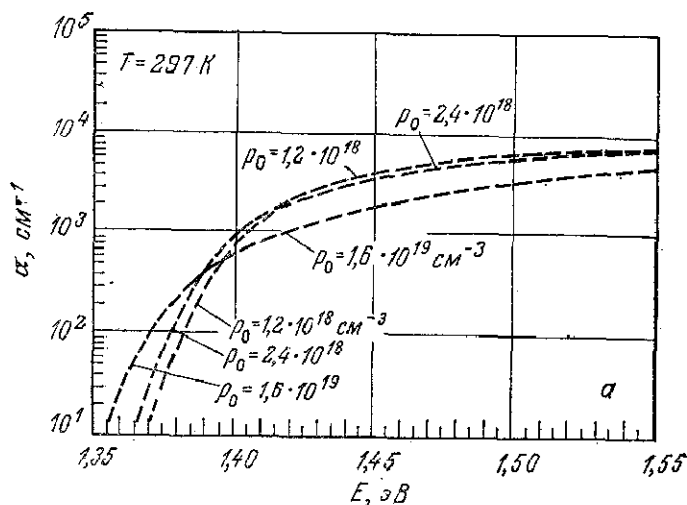


Рис. 3.7.3. Рассчитанные (а) [51] и экспериментальные (б) значения коэффициента поглощения (указаны концентрации дырок) [62].

ных, а вычисленная интенсивность поглощения ($\alpha(E)$ при $E > E_g$) примерно в 1,5—2,1 раза меньше. Тем не менее расчетные кривые коэффициента поглощения имеют те же характерные особенности, что и экспериментальные кривые.

В табл. 3.7.2 представлены данные, позволяющие сравнить положение пика и полуширину спектральных кривых излучения, полученных по спектрам фотолюминесценции (без поправ-

Таблица 3.7.2

Некоторые характеристики спектров излучения GaAs *p*-типа при 297 К [51]

| $\rho_0, \text{см}^{-3}$ | Положение максимума, эВ | | | Полуширина, эВ | | |
|--------------------------|-------------------------|----------------------|-------------------|-----------------------|----------------------|-------------------|
| | для фотолюминесценции | из значений α | по модели ГЛГ—МЭС | для фотолюминесценции | из значений α | по модели ГЛГ—МЭС |
| $1,2 \cdot 10^{18}$ | 1,415 | 1,425 | 1,414 | 0,059 | 0,063 | 0,063 |
| $2,4 \cdot 10^{18}$ | 1,415 | 1,420 | 1,411 | 0,061 | 0,065 | 0,064 |
| $1,6 \cdot 10^{19}$ | 1,400 | 1,405 | 1,398 | 0,070 | 0,073 | 0,074 |

ки на самопоглощение), найденных из экспериментальных значений $\alpha(E)$ и вычисленных по модели ГЛГ—МЭС (§ 5, 6 настоящей главы). Для образцов, легированных Zn, рассчитанные из экспериментальных значений $\alpha(E)$ и по модели ГЛГ—МЭС спектры согласуются в пределах 0,01 эВ по положению пика излучения и 0,002 эВ по полуширине. В случае фотолюминесценции пик смещен в сторону низких энергий, как это и должно происходить вследствие самопоглощения. При низких энергиях спектральные кривые имеют несколько различную форму, как это показано на рис. 3.7.2, б.

В высокочистом GaAs ширина запрещенной зоны при 297 К составляет 1,424 эВ [63]. Сужение запрещенной зоны ΔE_g определяется как разность между 1,424 эВ и значением E_g из табл. 3.7.1. Полученные значения ΔE_g показаны на рис. 3.7.4,

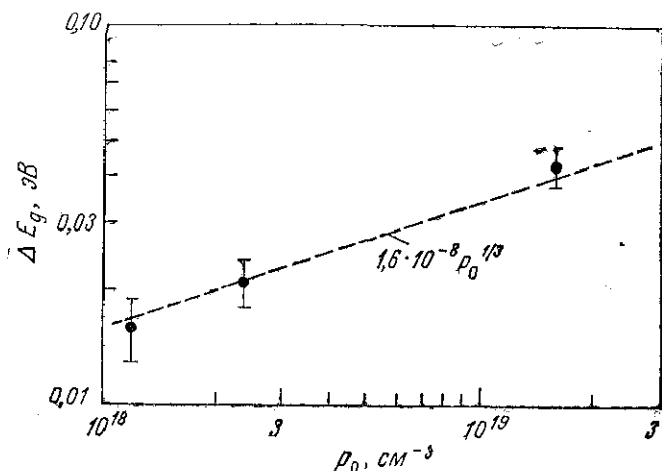


Рис. 3.7.4. Зависимость эффективного сужения запрещенной зоны от концентрации дырок [51].

где для двух низших уровней легирования неопределенность ΔE_g составляет $\pm 0,003$ эВ, а для более высокого уровня легирования $\pm 0,005$ эВ. Полученные данные проявляют зависимость вида $\rho_0^{1/3}$ и могут быть описаны формулой

$$E_g \text{ (эВ)} = 1,424 - 1,6 \cdot 10^{-8} \rho_0^{1/3}. \quad (3.7.2)$$

где ρ_0 в единицах см^{-3} . Многие авторы рассмотрели теоретически вопрос об уменьшении ширины запрещенной зоны в сильно легированных полупроводниках. Краткий обзор предложенных теорий дан в работе [51].

Сравнение показывает, что для GaAs *n*-типа согласие между экспериментальными значениями коэффициента поглощения и значениями, рассчитанными по модели ГЛГ — МЭС, не такое хорошее, как для материала *p*-типа [51]. Такие общие черты поведения экспериментально полученных зависимостей $\alpha(E)$ при изменении концентрации свободных электронов n_0 , как сдвиг Бурштейна [65] и уменьшение наклона кривой $\alpha(E)$ при увеличении *n*, присущи также и расчетным зависимостям $\alpha(E)$. Трудность может быть связана с тем, что на значения коэффициента поглощения в сильно легированном теллуром GaAs *n*-типа могут влиять преципитаты [66, 67].

Из проведенного выше сравнения экспериментальных и вычисленных значений коэффициента поглощения следует, что в модели для расчета оптических спектров необходимо использовать как плотность состояний, учитывающую хвосты зон, так и матричный элемент, зависящий от энергии. Для *p*-GaAs эта модель дает правильную форму кривой коэффициента поглощения и предсказывает ее изменение с увеличением концентрации дырок, согласующееся с наблюдаемым экспериментально. Однако при низких энергиях спад расчетной кривой происходит значительно быстрее, чем экспериментальной, а в области высоких энергий вычисленная интенсивность поглощения примерно в 1,5—2,1 раза меньше истинной. Дин [68] отметил, что увеличение $|M_b|^2$ в 1,2 раза, полученное Чейди и др. [69] при учете вклада в $|M_b|^2$ от более высоких зон, приводит к уменьшению указанной расходимости до 1,2—1,7 раза. Считают, что остающееся расхождение истинной и расчетной интенсивностей поглощения обусловлено не учитываемым в предложенной выше модели электрон-дырочным взаимодействием, которое приводит к возрастанию оптического матричного элемента [70, 71].

Параметры излучательной рекомбинации

Одной из интересующих нас величин является излучательное время жизни τ_r , для определения которого нужно вычислить полную скорость термической генерации пар и произведение

равновесных электронной и дырочной плотностей $n_0 p_0$. При тепловом равновесии полная скорость термической генерации пар в единице объема

$$G = \int_0^{\infty} \mathcal{G}(E) dE \quad (3.7.3)$$

представляет собой интеграл от определяемого выражением (3.7.1) спектра спонтанного равновесного излучения и может быть вычислена по известной зависимости коэффициента поглощения $\alpha(E)$. При внешней накачке скорость излучательной рекомбинации определяется хорошо известным выражением [72, 73]

$$R_r = G n p / n_0 p_0 \equiv B n p, \quad (3.7.4)$$

где n и p — электронная и дырочная концентрации в возбужденном полупроводнике, а постоянная излучения B равна $G/n_0 p_0$. В материале p -типа при достаточно слабом возбуждении, таком, что $p \approx p_0$, излучательное время жизни равно [72, 73]

$$\tau_r = (n - n_0) / R_r = 1 / B p. \quad (3.7.5)$$

Здесь R_r определяется формулой (3.7.4) и $n \gg n_0$. В GaAs невозможно экспериментально определить величину $n_0 p_0$, поэтому в вычислениях должно быть использовано расчетное значение $n_0 p_0 = n_i^2 / \gamma_n \gamma_p$, где n_i — собственная концентрация носителей в чистом материале, а параметры γ представляют собой коэффициенты активности электронов и дырок [47, 74, 75].

В табл. 3.7.3 приведены значения $n_0 p_0$, вычисленные с использованием плотности состояний в гауссовой форме с пара-

Таблица 3.7.3

Параметры излучательной рекомбинации GaAs p -типа при 297 К [51]

| $p_0, \text{см}^{-3}$ | $n_0 p_0, \text{см}^{-6}$ | $G, \text{см}^{-3}/\text{с}$ | | $B, 10^{-10} \text{см}^3/\text{с}$ | | $\tau_r, \text{нс}$ | |
|-----------------------|---------------------------|-----------------------------------------|-------------------|-----------------------------------------|-------------------|-----------------------------------------|-------------------|
| | | по экспериментальным значениям α | по модели ГЛГ—МЭС | по экспериментальным значениям α | по модели ГЛГ—МЭС | по экспериментальным значениям α | по модели ГЛГ—МЭС |
| $1,2 \cdot 10^{18}$ | $6,8 \cdot 10^{12}$ | 2200 | 1300 | 3,2 | 1,9 | 2,6 | 4,4 |
| $2,4 \cdot 10^{18}$ | $8,1 \cdot 10^{12}$ | 2250 | 1400 | 2,8 | 1,7 | 1,5 | 2,4 |
| $1,6 \cdot 10^{19}$ | $1,2 \cdot 10^{13}$ | 2080 | 1100 | 1,7 | 0,9 | 0,37 | 0,7 |

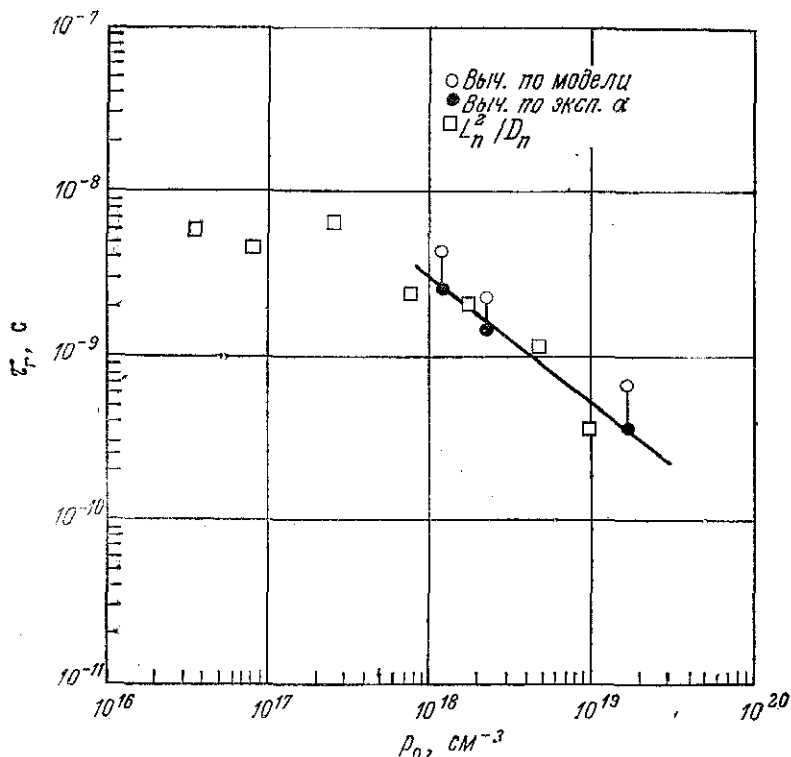


Рис. 3.7.5. Зависимость излучательного времени жизни от концентрации дырок при температуре 297 К. Квадратиками обозначены значения времени жизни, рассчитанные по ранее измеренным длинам диффузии [51, 76].

метром гауссиана η_v , выбранным так, чтобы плотность состояний в хвосте зоны совпадала с результатами модели Гальперина — Лэкса, как это описано в § 5 настоящей главы и продемонстрировано на рис. 3.5.8. Параметры хвостов зон η_v и η_c , а также значения ширины запрещенной зоны, зависящей от концентрации, даны в табл. 3.7.1. В GaAs при температуре 297 К, $m_n = 0,067 m_0$, $m_p = 0,55 m_0$ имеем $E_g = 1,424$ эВ и $n_i^2 = 3,2 \cdot 10^{12} \text{ см}^{-6}$. Следует отметить, что в сильно легированном GaAs $n_0 p_0 > n_i^2$, поэтому $\gamma_n \gamma_p < 1,0$. Этот результат находится в согласии с ранее определенными значениями γ_p при высоких температурах [47, 74]. Полученные значения $n_0 p_0$, а также значения коэффициента поглощения необходимы для вычисления B и τ_r .

В табл. 3.7.3 приведены значения полной скорости термической генерации G , вычисленные по экспериментальным значе-

ниям коэффициента поглощения с помощью формул (3.7.1) и (3.7.3). В таблице приведены также значения G , рассчитанные по описанной выше модели ГЛГ — МЭС, в которой учитываются хвосты зон и используется зависящий от энергии матричный элемент. Значения G , рассчитанные по экспериментальным значениям коэффициента поглощения и по модели, отличаются друг от друга так же (в 1,5—2,1 раза), как и экспериментальные и рассчитанные значения коэффициента поглощения. Значения постоянной B , приведенные в табл. 3.7.3 для трех образцов p -GaAs, рассчитаны по формуле $B = G/n_0\rho_0$ с использованием указанных в этой же таблице значений $n_0\rho_0$ и G . При высоких уровнях возбуждения значения постоянной B представлены на рис. 6 работы [61]. Приведенные в табл. 3.7.3 значения τ_r рассчитаны по формуле (3.7.5); на рис. 3.7.5 они представлены как функция концентрации дырок.

Теперь рассмотрим приведенные ранее результаты, касающиеся времени жизни неосновных носителей τ и длины диффузии L_n . Полное время жизни определяется формулой

$$\tau = \tau_r \tau_{nr} / (\tau_{nr} + \tau_r), \quad (3.7.6)$$

где τ_{nr} — безызлучательное время жизни. Поэтому время жизни неосновных носителей не может быть больше, чем τ_r , и стремится к τ_r только при $\tau_r \ll \tau_{nr}$. Длину диффузии неосновных носителей L_n , определяемую выражением $L_n = (D_n \tau)^{1/2}$, в котором коэффициент диффузии D_n связан с подвижностью μ соотношением $D_n = \mu kT/q$, можно использовать для определения τ . Значения времени жизни, полученные таким способом по приведенным в работе [76] значениям длины диффузии, представлены на рис. 3.7.5. Подвижность электронов в случае, когда они являются неосновными носителями, для каждой концентрации дырок может считаться равной подвижности электронов в материале n -типа с той же концентрацией носителей. Подвижность электронов в случае, когда они являются неосновными носителями заряда, может быть меньше, чем в случае, когда они являются основными носителями, но, как полагают, это отличие не слишком велико [77]. Поэтому следует ожидать, что времена жизни, рассчитанные с использованием выбранных указанным выше способом значений подвижности, несколько завышены по сравнению с действительными.

Времена жизни, рассчитанные как по модели ГЛГ — МЭС, так и по экспериментальным значениям $\alpha(E)$, очень хорошо согласуются с временами жизни, рассчитанными по значениям L_n из работы [76]. Исходя из представленных на рис. 3.7.5 данных, разумно предположить, что при $\rho_0 \leq 1 \cdot 10^{18} \text{ см}^{-3}$ время жизни главным образом определяется безызлучательной рекомбинацией. Приведенные в работе [76] результаты измерения длин

диффузии в основном согласуются с данными, полученными другими исследователями [78—82]. Однако в случаях, когда для определения L_n используется или сканирование световым пятном [83], или сканирование электронным лучом [84], получающиеся значения времени жизни превышают рассчитанные значения τ_r . Это расхождение при измерении L_n в прямозонных полупроводниках ван Олдорп и др. [85] объяснили возникновением при измерении излучательной рекомбинации и затем последующим поглощением фотонов, т. е. процессом переизлучения фотонов [85а]. Учет переизлучения фотонов в уравнении непрерывности [см. уравнение (4.5.26)] показывает, что метод сканирования световым пятном или электронным лучом дает наблюдаемое значение L_n [86]. Это значение больше действительной длины диффузии L_n и возрастает при увеличении квантовой эффективности [86]. Поэтому если сделать правильный выбор среди имеющихся экспериментальных данных, то окажется, что приведенные в табл. 3.7.3 и на рис. 3.7.5 численные значения характерны для рекомбинационных постоянных p -GaAs.

Расчетные спектры усиления

В § 2 настоящей главы коэффициент поглощения определен как разность между скоростями переходов вверх и вниз, деленная на поток фотонов. Поэтому положительные значения $\alpha(E)$ означают, что в целом происходит поглощение излучения, в то время как отрицательные значения $\alpha(E)$ говорят о том, что скорость вынужденного излучения превосходит скорость поглощения. Отрицательный коэффициент поглощения $\alpha(E)$ называется коэффициентом усиления $g(E)$. Условие, необходимое для того, чтобы скорость вынужденного излучения превышала скорость поглощения, выражается соотношением (3.2.31). Оно требует, чтобы энергетический зазор между квазиуровнями Ферми был больше энергии излучаемых фотонов. Для достижения лазерной генерации нужно, чтобы усиление превосходило потери в приборе, такие, как поглощение на свободных носителях, рассеяние, излучение через зеркальные грани. Условия, определяющие порог генерации, рассматриваются в следующем параграфе.

Коэффициент усиления может быть рассчитан по модели ГЛГ—МЭС так же, как коэффициент поглощения. Определяемый выражением (3.4.6) коэффициент поглощения становится отрицательным при таких высоких уровнях накачки, при которых $(F_c - F_v) < \hbar\omega$. Квазиуровни Ферми F_c и F_v для электронов и дырок определяются по соответствующим концентрациям носителей и связаны с ними через интегралы, аналогичные

интегралу в формуле (3.5.2) [21]:

$$n = \int_{-\infty}^{\infty} \rho_c(E - E_c) f_c dE, \quad (3.7.7)$$

$$p = \int_{-\infty}^{\infty} \rho_v(E_v - E) [1 - f_v] dE, \quad (3.7.8)$$

где f_c и f_v определяются выражениями (3.2.18) и (3.2.19). В формулах (3.7.7) и (3.7.8) величины $E - E_c$ и $E_v - E$ описывают края номинальных (параболических) зон, и разность $E_c - E_v$ равна зависящей от концентрации ширине запрещенной зоны E_g . Хвосты зон не оказывают значительного влияния на F_c и F_v [52], поэтому для получения типичных численных значений этих величин можно использовать удобную параболическую аппроксимацию. Более детальное обсуждение вопроса об оценке величин F_c и F_v проводится в § 3 гл. 4. В вычислениях, рассматриваемых здесь, использовалась модель ГЛГ. При высоких концентрациях неосновных носителей концентрация основных носителей должна увеличиться, чтобы выполнялось условие электронейтральности:

$$n + N_A^- = p + N_D^+. \quad (3.7.9)$$

На рис. 3.7.6 показаны графики распределения носителей при высоком уровне накачки. На этом рисунке схематически изображены плотности состояний и распределения электронов и дырок в образце GaAs с $\rho_0 = 1,2 \cdot 10^{18} \text{ см}^{-3}$, различные параметры которого приведены в табл. 3.7.1—3.7.3. Концентрация инжектированных электронов равна $1 \cdot 10^{18} \text{ см}^{-3}$. Заштрихованные участки изображают распределения электронов и дырок. При тепловом равновесии $p = p_0 = 1,2 \cdot 10^{18} \text{ см}^{-3}$. При накачке до концентрации электронов $n = 1 \cdot 10^{18} \text{ см}^{-3}$ p становится равной $2,2 \cdot 10^{18} \text{ см}^{-3}$, согласно уравнению (3.7.9). Как следует из соотношения (3.2.31), только электроны, находящиеся ниже или вблизи F_c , и дырки, расположенные выше или вблизи F_v , могут участвовать в процессе вынужденного излучения. Из рис. 3.7.6 видно, что при комнатной температуре лишь малая доля носителей дает вклад в усиление и вынужденное излучение. При понижении температуры все большая часть носителей участвует в усилении.

Для того же образца с $\rho_0 = 1,2 \cdot 10^{18} \text{ см}^{-3}$ Стерн [87] рассчитал, как меняется определяемая выражением (3.4.6) зависимость $\alpha(E)$ при изменении концентрации неосновных носителей n . Полученное им семейство кривых $\alpha(E)$ показано на рис. 3.7.7. Плотность состояний и матричный элемент брались

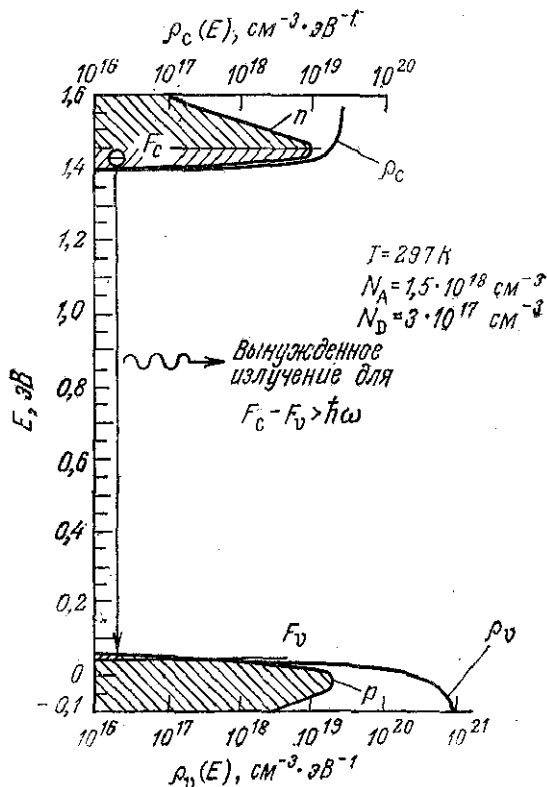


Рис. 3.7.6. Схематическое представление плотностей состояний и распределений электронов и дырок при накачке до концентраций $n = 1 \cdot 10^{18} \text{ см}^{-3}$ и $p = 2,2 \cdot 10^{18} \text{ см}^{-3}$. В процессе вынужденного излучения могут участвовать только те электроны и дырки, разность энергий которых меньше энергетического зазора между квазиуровнями Ферми.

те же, что и при расчетах $\alpha(E)$, результаты которых представлены на рис. 3.7.2. Зависимость ширины запрещенной зоны от концентрации определялась выражением, аналогичным выражению (3.7.2):

$$E_g (\text{эВ}) = 1,424 - 1,6 \cdot 10^{-8} (p^{1/3} + n^{1/3}). \quad (3.7.10)$$

При увеличении концентрации инжектированных носителей до значения $n \approx 1 \cdot 10^{18} \text{ см}^{-3}$ $\alpha(E)$ становится отрицательным, и тем самым достигается усиление. Дальнейшее небольшое увеличение n приводит к сильному возрастанию коэффициента усиления. Полученные кривые для коэффициента усиления спадают до нуля в области низких энергий, а их максимум при увеличении концентрации неосновных носителей сдвигается в сторону более высоких энергий. Большинство расчетов дает

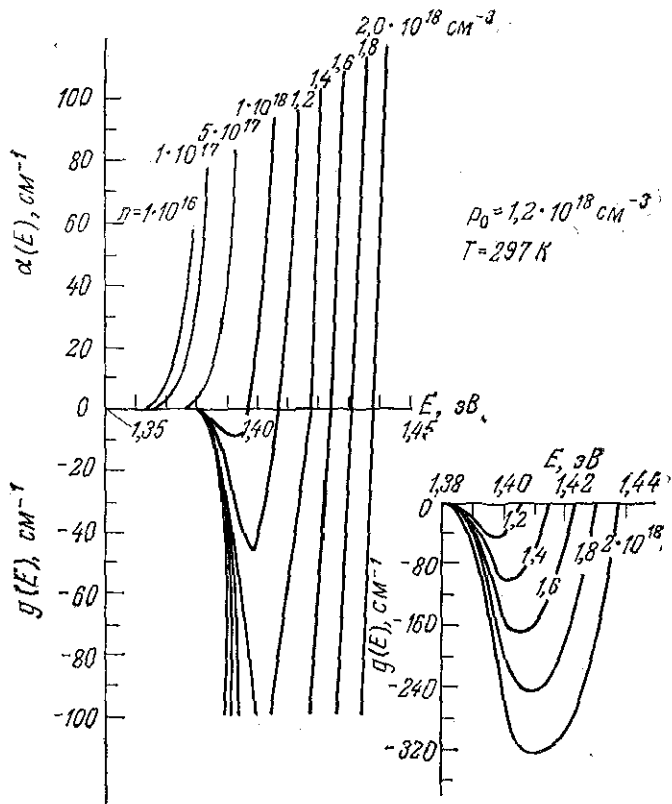


Рис. 3.7.7. Коэффициент поглощения (усиления), рассчитанный для GaAs по модели ГЛГ — МЭС. Уровни накачки характеризуются указанными на рисунке концентрациями инжектированных неосновных носителей [87].

только максимальное значение g_{\max} при заданных накачке или плотности тока. В следующем параграфе после того, как понятие накачки вводится на основе номинальной плотности тока, представлены результаты еще одного расчета спектров усиления. Теория оптического усиления в сильно легированных полупроводниках была предложена также Александяном и др. [88], но здесь она не рассматривается.

§ 8. ПОРОГОВАЯ ПЛОТНОСТЬ ТОКА

Условие генерации

В предыдущем параграфе показано, что когда уровень накачки удовлетворяет необходимому условию вынужденного излучения, коэффициент поглощения $\alpha(E)$, связанный с $g_{\text{вынужд}}(E)$

соотношением (3.2.41), становится отрицательным и определяет коэффициент усиления $g(E)$ в виде

$$g(E) = -\alpha(E) = (h^3 c^2 / 8 \pi \hbar^2 E^2) r_{\text{вынужд}}(E). \quad (3.8.1)$$

В этом выражении $\alpha(E)$ берется со знаком минус, поскольку $r_{\text{вынужд}}(E)$ и $g(E)$ определены таким образом, что они положительны, когда излучение испускается, в то время как $\alpha(E)$ положительно, когда излучение поглощается. Коэффициент усиления $g(E)$ принято также называть просто усилением.

Лазер обычно изготавливают путем создания в среде с усилением параллельных отражающих поверхностей, которые образуют эталон или интерферометр Фабри — Перо [89]. Условие генерации может быть получено при рассмотрении отражения плоских волн на параллельных частично отражающих поверхностях, как это показано на рис. 3.8.1. Такая модель соответствует описанию процесса генерации в лазере, предложенному Яривом [90].

Пусть плоская волна с комплексной постоянной распространения Γ , определяемой выражением (2.2.54), падает на левое зеркало резонатора (рис. 3.8.1). Длина резонатора равна L , а отношения амплитуды прошедшего поля к амплитуде падающего поля на левом и правом зеркалах равны t_1 и t_2 соответственно. Отношения амплитуд отраженного и падающего полей внутри оптического резонатора равны $r_1 \exp(j\theta_1)$ на левом зеркале и $r_2 \exp(j\theta_2)$ на правом. Эти комплексные коэффициенты отражения для амплитуд поля связаны с соответствующими коэффициентами отражения по мощности соотношениями $R_1 = r_1 r_1^*$ и $R_2 = r_2 r_2^*$. В среде с низкими потерями фазовые сдвиги

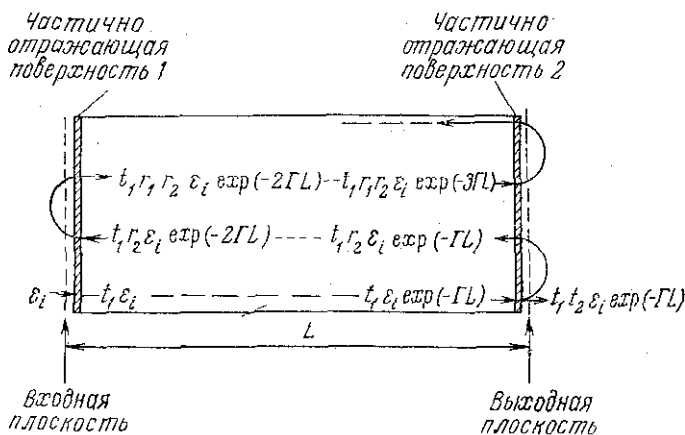


Рис. 3.8.1. К выводу условия генерации в среде с усилением и параллельными отражающими поверхностями [90].

θ_1 и θ_2 малы, и ими обычно пренебрегают, так что $r_1 = R_1^{1/2}$ и $r_2 = R_2^{1/2}$.

Если не учитывать временную зависимость, то выражение (2.2.46) определяет электрическое поле в виде плоской волны $\mathcal{E}_i \exp(-\Gamma z)$, так что \mathcal{E}_x равно $t_1 \mathcal{E}_i$ внутри резонатора на левой границе и $t_1 \mathcal{E}_i \exp(-\Gamma L)$ внутри резонатора непосредственно перед правой границей. При первом падении на правую границу через нее проходит часть поля, равная $t_1 t_2 \mathcal{E}_i \exp(-\Gamma L)$, а отраженная часть поля равна $t_1 r_2 \mathcal{E}_i \exp(-\Gamma L)$. При следующем падении на эту границу через нее проходит часть поля, равная уже $t_1 t_2 r_2 \mathcal{E}_i \exp(-3\Gamma L)$, и т. д. Сложение амплитуд прошедших через правую границу полей дает

$$\mathcal{E}_t = t_1 t_2 \mathcal{E}_i \exp(-\Gamma L) [1 + r_1 r_2 \exp(-2\Gamma L) + r_1^2 r_2^2 \exp(-4\Gamma L) + \dots]. \quad (3.8.2)$$

Это выражение представляет собой сумму геометрической прогрессии, которая равна

$$\mathcal{E}_t = \mathcal{E}_i \left[\frac{t_1 t_2 \exp(-\Gamma L)}{1 - r_1 r_2 \exp(-2\Gamma L)} \right]. \quad (3.8.3)$$

Обращение в нуль знаменателя в формуле (3.8.3) есть условие того, что прошедшее поле может иметь конечную амплитуду \mathcal{E}_t при нулевом значении \mathcal{E}_i , что в свою очередь является условием генерации. Следовательно, генерация возникает в том случае, когда

$$r_1 r_2 \exp(-2\Gamma L) = 1. \quad (3.8.4)$$

Используя выражения (2.2.54), (2.2.38) и (2.2.61), уравнение (3.8.4) можно переписать в виде

$$r_1 r_2 \exp[(g - \alpha_i)L] \exp[-2j(2\pi\bar{n}/\lambda_0)L] = 1. \quad (3.8.5)$$

В этом уравнении член, описывающий поглощение, записан в виде разности усиления и всех потерь α_i . Потери более подробно рассматриваются ниже в этом параграфе. До тех пор пока не выполнены необходимые условия вынужденного излучения, g также обозначает потери и должно быть заменено на α из выражения (3.8.1).

Условие генерации, определяемое уравнением (3.8.5), означает, что волна при возвратно-поступательном движении внутри резонатора длиной L должна, проходя путь $2L$, возвращаться к исходной плоскости с той же амплитудой и фазой, измененной на величину, кратную 2π . Амплитудное условие генерации, таким образом, имеет вид

$$r_1 r_2 \exp[(g - \alpha_i)L] = 1, \quad (3.8.6)$$

или

$$g = \alpha_i + (1/L) \ln (1/r_1 r_2). \quad (3.8.7)$$

Фазовое условие определяется соотношением

$$4\pi \bar{n} L / \lambda_0 = 2m\pi \quad (3.8.8)$$

с $m = 1, 2, 3 \dots$, которое преобразуется к виду

$$m (\lambda_0 / \bar{n}) = 2L. \quad (3.8.9)$$

Расстояние между продольными модами может быть получено при дифференцировании равенства $m\lambda_0 = 2\bar{n}L$, вытекающего из формулы (3.8.9):

$$\lambda_0 dm + m d\lambda_0 = 2L d\bar{n}. \quad (3.8.10)$$

Для соседних мод $dm = -1$, и, выражая m из формулы (3.8.9), получаем расстояние между ними

$$d\lambda_0 = \frac{\lambda_0^2}{2\bar{n}L [1 - (\lambda_0/\bar{n}) (d\bar{n}/d\lambda_0)]}. \quad (3.8.11)$$

Выражение (3.8.7) обычно записывают через коэффициенты отражения по мощности:

$$g = \alpha_i + (1/2L) \ln (1/R_1 R_2), \quad (3.8.12)$$

и при $R_1 = R_2 = R$ требуемое для возникновения генерации усиление определяется формулой

$$g = \alpha_i + (1/L) \ln (1/R). \quad (3.8.13)$$

Численные значения коэффициентов отражения на зеркальных торцах двусторонних гетероструктур в виде резонатора Фабри — Перо даны в § 8 гл. 2. Условия порогового усиления в лазерах с РОС рассматривались в § 10 гл. 2.

Коэффициент усиления как функция номинальной плотности тока

Как показано на рис. 3.7.7, спектральная зависимость коэффициента усиления рассчитывается для данных концентраций доноров и акцепторов при указанных концентрациях неосновных носителей. В примере, который иллюстрирует этот рисунок, коэффициент усиления был получен как функция концентрации неосновных носителей — электронов, в то время как концентрация основных носителей определялась условием электронейтральности (3.7.9). Непосредственно перед порогом генерации и при квантовой эффективности, равной единице, все инжектированные носители излучательно рекомбинируют, давая вклад

в спонтанное излучение. Число излучаемых фотонов в единице объема в одну секунду определяется полной скоростью излучательной рекомбинации

$$\mathcal{R} = \int_0^{\infty} r_{\text{спонт}}(E) dE, \quad (3.8.14)$$

где $r_{\text{спонт}}(E)$ рассчитывается по формуле (3.4.8) с теми же плотностью состояний и матричным элементом в модели ГЛГ — МЭС, которые использовались при вычислении $\alpha(E)$. Число электронов, протекающих по внешней цепи в одну секунду, равно току I , деленному на заряд электрона q . При квантовой эффективности, равной единице, на каждый из этих электронов излучается один фотон. Для пластины толщиной d и площадью поперечного сечения a число излучаемых в одну секунду в единице объема фотонов равно $\mathcal{R} = I/qad$. Номинальная плотность тока $J_{\text{ном}}$ при равной единице квантовой эффективности определяется выражением

$$J_{\text{ном}} (\text{А/см}^2) = I/a = q\mathcal{R}d. \quad (3.8.15)$$

Стало общепринятым брать значение d , равное 1 мкм (10^{-4} см), так что $J_{\text{ном}}$ представляет собой плотность тока, необходимую для поддержания существующей скорости накачки однородно возбужденного слоя толщиной 1 мкм. В этом случае

$$J_{\text{ном}} (\text{А/см}^2 \cdot \text{мкм}) = 1,602 \cdot 10^{23} \mathcal{R}. \quad (3.8.16)$$

Ряд дополнительных факторов необходимо учесть, чтобы связать $J_{\text{ном}}$ с истинной плотностью тока J , текущего через активный слой гетеролазера. Поскольку квантовая эффективность η определяется выражением

$$\eta = \frac{\mathcal{R}ad}{I/q}, \quad (3.8.17)$$

плотность тока $J = I/a$ в случае, когда поле моды целиком заключено в активном слое, равна $q\mathcal{R}d/\eta$. Используя $J_{\text{ном}}$, нормированную на толщину слоя, равную 1 мкм, получаем

$$J (\text{А/см}^2) = J_{\text{ном}}d/\eta, \quad (3.8.18)$$

где d — толщина активного слоя в микрометрах. Как описано в § 5 гл. 2, оптическое поле не ограничивается полностью областью активного слоя, особенно при малых d . Поле снаружи активного слоя не может взаимодействовать с инжектированными в активный слой носителями и поэтому не дает вклада в усиление. Отношение энергии, переносимой модой внутри активного слоя, к полной переносимой модой энергии называется коэффициентом оптического ограничения Γ . Поэтому коэффициент усиления равен $\Gamma g(E)$.

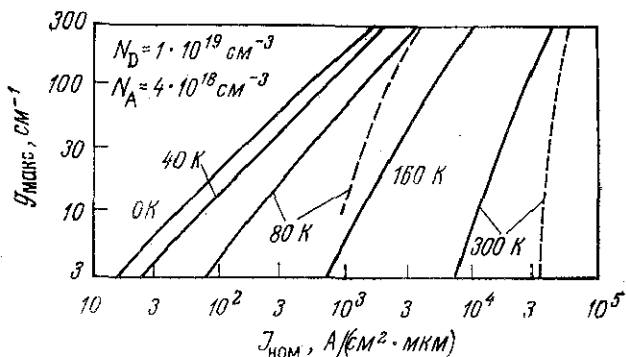


Рис. 3.8.2. Зависимость коэффициента усиления от номинальной плотности тока при рекомбинации в GaAs с указанными на рисунке концентрациями примесей. Для хвостов зон использовалась модель Кейна, а матричный элемент брался для перехода между параболической зоной и примесным уровнем. Штриховые кривые представляют результаты, полученные в работе [8] без учета хвостов зон [9].

Представляется полезным связать $J_{\text{ном}}$ с максимальным коэффициентом усиления $g_{\text{макс}}$. Как видно из рис. 3.7.7, $g_{\text{макс}}$ сдвигается в область более высоких энергий при увеличении уровня накачки, так что максимальное значение коэффициента усиления не соответствует постоянной энергии фотона. При вычислении как $g_{\text{макс}}$, так и $J_{\text{ном}}$ независимым параметром является концентрация неосновных носителей, а концентрация основных носителей определяется из условия электронейтральности (3.7.9). Квазиуровни Ферми F_v и F_c связаны с концентрациями носителей соотношениями (3.7.7) и (3.7.8). Для расчета коэффициента усиления используется выражение (3.4.6), а номинальная плотность тока дается выражением (3.8.16), в котором \mathcal{R} вычисляется по формуле (3.8.14). В этой формуле $r_{\text{спонт}}(E)$ определяется выражением (3.4.8).

Большинство из опубликованных расчетов численных значений коэффициента усиления основаны на моделях, являющихся модификациями модели Лэшера и Стерна [8]. В этой модели используются параболические зона проводимости и валентная зона, а также постоянный матричный элемент. Впоследствии Стерн [9] рассмотрел влияние хвостов зон на вынужденное излучение, используя для вычисления коэффициента усиления представление хвостов зон по модели Кейна [4]. В этом случае матричный элемент описывает рекомбинацию между параболической зоной проводимости и водородоподобным акцепторным уровнем и определяется произведением выражений (3.6.14) и (3.6.19). На рис. 3.8.2 для пяти температур между 0 и 300 К представлены рассчитанные в работе [9] зависимости $g_{\text{макс}}$ от

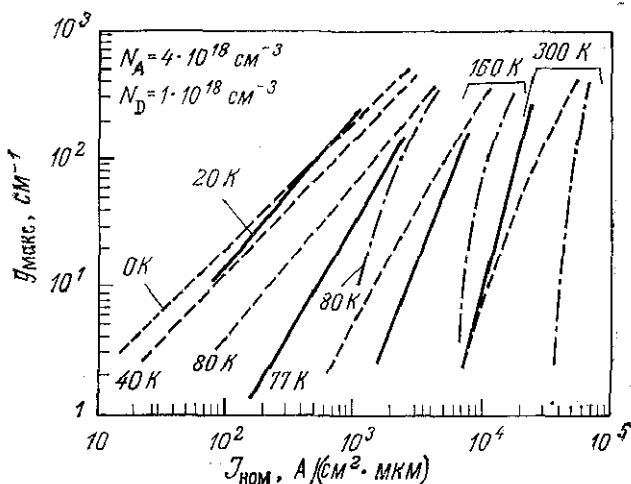


Рис. 3.8.3. Зависимость коэффициента усиления от номинальной плотности тока при рекомбинации в GaAs с указанными на рисунке концентрациями примесей. Сплошными линиями показаны результаты для модели хвостов зон Гальперина — Лэкса [10], штрихпунктирными — результаты для параболических зон [8] и штриховыми — результаты для модели хвостов зон Кейна, полученные при использовании матричного элемента для параболической зоны и примесного уровня [9].

$J_{\text{ном}}$. Для сравнения на этом же рисунке штриховыми кривыми показаны результаты Лэшера и Стерна [8], полученные для параболических зон. Из рисунка хорошо видно, что при заданном $J_{\text{ном}}$ учет хвостов зон ведет к значительному увеличению $g_{\text{макс}}$. Модель хвостов зон Кейна была использована также Унгером [91].

Хуань [10] проделал более точные расчеты коэффициента усиления, используя для хвостов зон модель Гальперина — Лэкса [5] вместо модели Кейна [4]. Как отмечалось в § 5 настоящей главы, модель Кейна дает завышенную степень размытости распределения состояний. Хвосты зон в модели Кейна и в модели Гальперина — Лэкса сравниваются на рис. 3.5.7. На рис. 3.8.3 для четырех температур между 20 и 300 К показаны рассчитанные Хуанем зависимости $g_{\text{макс}}$ от $J_{\text{ном}}$. Для сравнения на этом рисунке приведены также результаты, представленные на рис. 3.8.2.

Стерн [60] получил затем более точное выражение для зависящего от энергии матричного элемента [выражение (3.6.23)], которое описано в § 6 этой главы. Он предложил также модель ГЛГ, дающую плотность состояний, представленную на рис. 3.5.8. Эта модель с матричным элементом Стерна (ГЛГ — МЭС) использована при расчете данных рис. 3.7.7, а также коэффици-

ентов усиления, приведенных в работе [60]. Однако в этой работе не учитывались эффекты сужения запрещенной зоны вследствие сильного легирования. На рис. 3.8.4 показаны полученные в работе [60] зависимости $g_{\text{макс}}$ от $J_{\text{ном}}$ для компенсированного n -GaAs. Как описано в предыдущем параграфе, в работе [87] Стерн использовал модель ГЛГ — МЭС и учел сужение запрещенной зоны, определяемое формулой (3.7.10), при вычислении $g(E)$ в образце, характеристики которого приведены в первых строчках табл. 3.7.1—3.7.3. Полученные результаты соответствуют температуре 297 К и показаны на рис. 3.7.7. Вычисленные значения $g_{\text{макс}}$ как функция номинальной плотности тока представлены на рис. 3.8.5. Для вычисленного $J_{\text{ном}}$ внутренняя квантовая эффективность взята равной единице. Определение значений η обсуждается в § 4 гл. 7. Зависимость $g_{\text{макс}}$ от $J_{\text{ном}}$, полученная Хуанем для температуры 300 К и представленная на рис. 3.8.2, приводится также на рис. 3.8.5, на котором она показана штриховой линией. Концентрации примесей в этих двух случаях были немного различ-

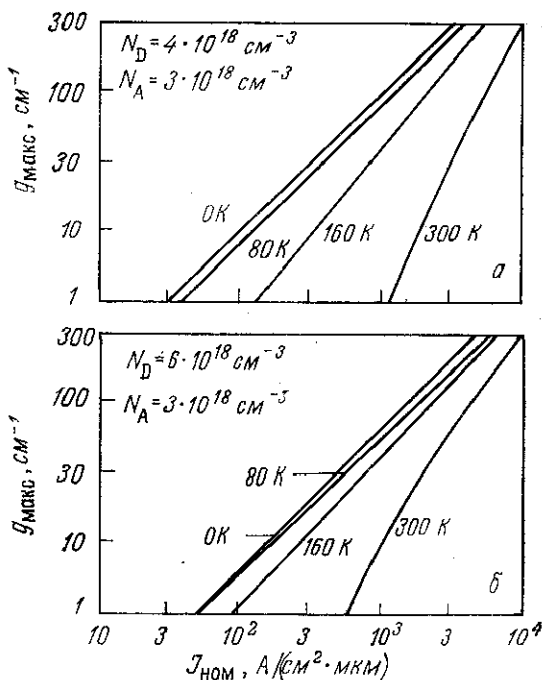


Рис. 3.8.4. Зависимость коэффициента усиления от номинальной плотности тока при рекомбинации в GaAs с указанными на рисунке концентрациями примесей. При расчетах использовалась модель ГЛГ — МЭС и не учитывалось сужение запрещенной зоны [60].

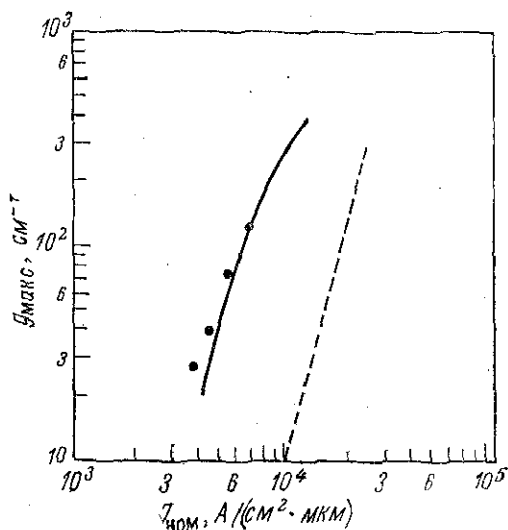


Рис. 3.8.5. Зависимость коэффициента усиления от номинальной плотности тока в GaAs. Точками показаны экспериментальные данные для $p_0 = 1,3 \cdot 10^{18} \text{ см}^{-3}$ [92], сплошной линией — значения g_{\max} , вычисленные Стерном по модели ГЛГ — МЭС с учетом сужения запрещенной зоны [87], и штриховой линией — представленные на рис. 3.8.3 значения g_{\max} , вычисленные по модели хвостов зон Гальперина — Лэкса с использованием матричного элемента для параболической зоны и примесного уровня.

ны, но тем не менее сравнение приведенных на рис. 3.8.5 кривых показывает, что использование более точного выражения для матричного элемента [выражение (3.6.23)] оказывает значительное влияние на расчетные значения коэффициента усиления.

На рис. 3.8.5 кружками показаны экспериментально полученные Пинкасом и др. [92] значения коэффициента усиления для GaAs ДГС-лазера с $p_0 = 1,3 \cdot 10^{18} \text{ см}^{-3}$. Эти значения определены по формуле (3.8.13) при $\alpha_i = 12 \text{ см}^{-1}$ и $R = 0,30$ для лазеров различной длины. Используемый для определения α_i метод обсуждается в настоящем параграфе в разделе, посвященном рассмотрению дифференциальной квантовой эффективности. Толщины активных слоев в ДГС были достаточно велики, чтобы с уверенностью считать $\Gamma \approx 1,0$.

Стерн [61] рассчитал также коэффициент усиления в GaAs с $(N_A - N_D) = 4 \cdot 10^{17} \text{ см}^{-3}$, что дало возможность сравнить вычисленные им значения с экспериментальными результатами Хакки и Паоли [93]. Считалось, что образец имел $N_A = 5 \cdot 10^{17} \text{ см}^{-3}$ и $N_D = 1 \cdot 10^{17} \text{ см}^{-3}$. На рис. 3.8.6 для $T = 297$ и 77 К приведены расчетные значения коэффициента усиления (поглощения) как функции энергии фотона при разных уровнях

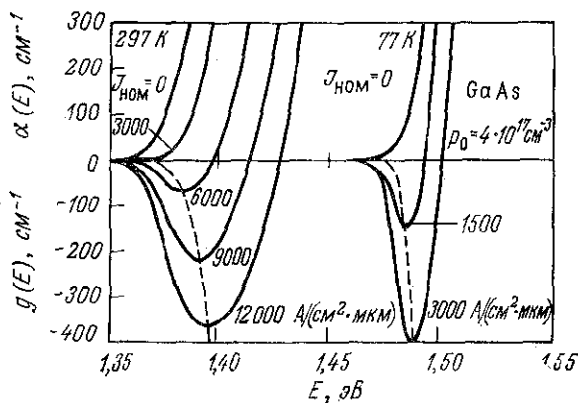


Рис. 3.8.6. Рассчитанный по модели ГЛГ — МЭС коэффициент поглощения (усиления) в GaAs. При расчетах учитывалось сужение запрещенной зоны вследствие накачки. Уровень накачки характеризуется величиной $J_{\text{ном}}$ [61].

возбуждения, характеризуемых величиной $J_{\text{ном}}$. Полезно также представить коэффициент усиления (поглощения) как функцию $J_{\text{ном}}$ при фиксированной энергии фотона. Семейство таких кривых приведено на рис. 3.8.7, на котором показано также максимальное значение коэффициента усиления $g_{\text{макс}}$ для каждого $J_{\text{ном}}$.

Для зависимости коэффициента усиления от тока существует несколько представлений, которые использовались многими авторами. Поскольку ранние вычисления коэффициента усиления, например те, результаты которых приведены на рис. 3.8.2, 3.8.3 и 3.8.4, давали суперлинейную зависимость коэффициента усиления от тока, для представления этой зависимости использовали выражение

$$g_{\text{макс}} = \beta J_{\text{ном}}^b, \quad (3.8.19)$$

где β называют фактором усиления или постоянной усиления. Однако вычисление коэффициентов усиления по улучшенной модели ГЛГ — МЭС (рис. 3.8.7) показывает, что зависимость $g_{\text{макс}}$ от $J_{\text{ном}}$ суперлинейна только при низких значениях $g_{\text{макс}}$, а затем в области $50 \text{ см}^{-1} \leq g_{\text{макс}} \leq 400 \text{ см}^{-1}$ она становится линейной. В обоих случаях нуль $g_{\text{макс}}$ не совпадает с нулем $J_{\text{ном}}$, поэтому выражение (3.8.19) неверно представляет зависимость $g_{\text{макс}}$ от $J_{\text{ном}}$. В области, где усиление мало, определяемый формулой (3.8.17) $g_{\text{макс}}$ может быть записан в виде

$$g_{\text{макс}} = \beta (J_{\text{ном}} - J_0)^2, \quad (3.8.20)$$

а в области значительного усиления

$$g_{\text{макс}} = \beta (J_{\text{ном}} - J_0), \quad (3.8.21)$$

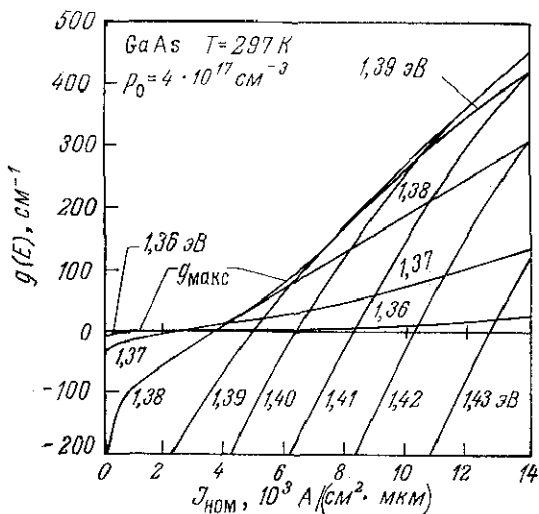


Рис. 3.8.7. Зависимость рассчитанного по модели ГЛГ—МЭС коэффициента усиления (поглощения) от номинальной плотности тока, характеризующей уровень накачки. Учитывалось сужение запрещенной зоны вследствие накачки. У каждой кривой указана соответствующая энергия фотона. Исключение составляет верхняя огибающая кривая, представляющая максимальные значения коэффициента усиления $g_{\text{макс}}$ [61].

где J_0 равно значению $J_{\text{ном}}$, при котором $g_{\text{макс}}$ становится равным нулю при квадратичной зависимости (3.8.20), или определяется линейной экстраполяцией кривой, определяемой выражением (3.8.21), до нулевого значения $g_{\text{макс}}$. Величины β в формулах (3.8.20) и (3.8.21) различны. Хотя использование β удобно для представления зависимости коэффициента усиления от тока, эта величина не имеет большого физического смысла и не играет значительной роли в полупроводниковых лазерах. К тому же β имеет разные значения в зависимости от характеристик материала, таких, как концентрация примесей.

Температурная зависимость усиления

На рис. 3.8.2, 3.8.3 и 3.8.4 для ряда температур представлены зависимости $g_{\text{макс}}$ от $J_{\text{ном}}$. Температурная зависимость может быть получена при нанесении на график значений номинальной плотности тока, необходимых для получения усиления 50 cm^{-1} , как это показано на рис. 3.8.8. Этот рисунок взят из работы [60], и на нем представлены результаты для образцов как n -, так и p -типа. В обоих случаях $J_{\text{ном}}$ значительно возрастает, когда температура становится больше чем $\sim 100 \text{ K}$. Это возрастание $J_{\text{ном}}$

при увеличении температуры главным образом является результатом возбуждения все большего числа носителей на уровни, лежащие выше квазиуровней Ферми.

Трудно связать температурные зависимости вычисленной $J_{\text{ном}}$ и экспериментальной пороговой плотности тока $J_{\text{пор}}$. Температурная зависимость $J_{\text{ном}}$ содержит только температурную зависимость коэффициента усиления. В GaAs—Al_xGa_{1-x}As ДГС-лазерах, а также в гетеролазерах на других системах важный вклад в температурную зависимость $J_{\text{пор}}$ может давать утечка электронов из активного слоя в P-Al_xGa_{1-x}As-слой. Значение температуры, при котором утечка носителей приводит к возрастанию $J_{\text{пор}}$, зависит от молярной доли AlAs [94, 95]. Дополнительный ток, вызванный утечкой носителей, рассматривается в § 6 гл. 4. Он пренебрежимо мал вплоть до комнатной температуры при $x = 0,3$. Хаяси и др. [96] получили приближенное экспериментальное соотношение $J_{\text{пор}} \sim \exp(T/T_0)$ с T_0 , лежащей в области от 120 до 165 К. Экспериментально измеренное возрастание $J_{\text{пор}}$ в 3—5 раз при изменении температуры в пределах от 100 до 300 К находится в согласии с расчетной зависимостью $J_{\text{ном}}$, приведенной на рис. 3.8.8 (см. § 4 гл. 7). Проведение расчетов зависимости коэффициента усиления от $J_{\text{ном}}$ для температур, превышающих 300 К, представляет определенный интерес, поскольку исследования деградации лазеров часто проводят при повышенных температурах, и ДГС-лазеры, работающие в непрерывном режиме при комнатной температуре теплопровода, имеют температуры $p-n$ -переходов выше примерно на 5—20 °С.

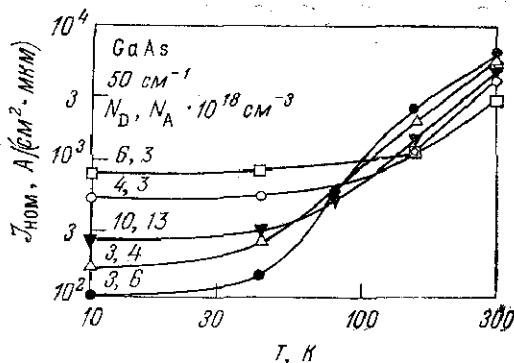


Рис. 3.8.8. Зависимость от температуры номинальной плотности тока, необходимой для получения коэффициента усиления 50 см^{-1} в GaAs для двух образцов n -типа и трех образцов p -типа. Целые числа на каждой кривой обозначают в единицах 10^{18} см^{-3} концентрации доноров и акцепторов соответственно [60].

Потери в гетеролазерах

Как уже говорилось в первой части этого параграфа, для возникновения генерации необходимо, чтобы усиление в лазере превышало потери. Потери, обусловленные излучением с торцов лазера, равны $(1/L)\ln(1/R)$. В § 8 гл. 2 приведены численные значения этой величины в зависимости от толщины активного слоя и скачка показателя преломления. Все другие потери обозначаются через α_i . Наиболее важными из них являются неизбежные потери из-за поглощения свободными носителями $\alpha_{\text{св. нос}}$ и потери вследствие рассеяния α_s . Последние обусловлены нерегулярностями гетерограниц или внутренних областей волновода. Потери для волноводной моды могут также возникать из-за ее связи со слоями структуры, в которых существенны потери.

Поглощение на свободных носителях имеет решающее значение для полупроводниковых лазеров, так как оно является основным неустраняемым механизмом потерь и, как было показано Думке [12], делает невозможным лазерную генерацию в непрямозонных полупроводниках. Это поглощение происходит при рассеянии движущихся носителей, поэтому оно зависит от тех же механизмов рассеяния, что и подвижность носителей. Фэн [97] дал краткую сводку формул, описывающих поглощение на свободных носителях при различных механизмах рассеяния. Удобно выразить $\alpha_{\text{св. нос}}$ через сечения захвата фотона дырками

$$\sigma_p = \alpha_p / p \quad (3.8.22)$$

и электронами

$$\sigma_n = \alpha_n / n. \quad (3.8.23)$$

Из полученных Хиллом [98] данных по поглощению следует, что $\alpha_{\text{св. нос}}$ в GaAs уменьшается при понижении температуры. Спитцер и Уэлан [99] показали, что α_n линейно зависит от концентрации носителей. Различные опубликованные данные [62, 98—100] позволяют представить коэффициент поглощения на свободных носителях при энергиях фотонов, близких к ширине запрещенной зоны GaAs при комнатной температуре, в виде

$$\alpha_{\text{св. нос}} (\text{см}^{-1}) \approx 3 \cdot 10^{-18} n + 7 \cdot 10^{-18} p. \quad (3.8.24)$$

Вследствие проникновения части волноводной моды из активного в прилегающие слои происходит уменьшение $\alpha_{\text{св. нос}}$ в Γ^{-1} раз (Γ — коэффициент оптического ограничения), так же как это происходит с коэффициентом усиления в активном слое. Таким образом, поглощение на свободных носителях в активном слое характеризуется величиной $\Gamma \alpha_{\text{св. нос}}$. В окружающих

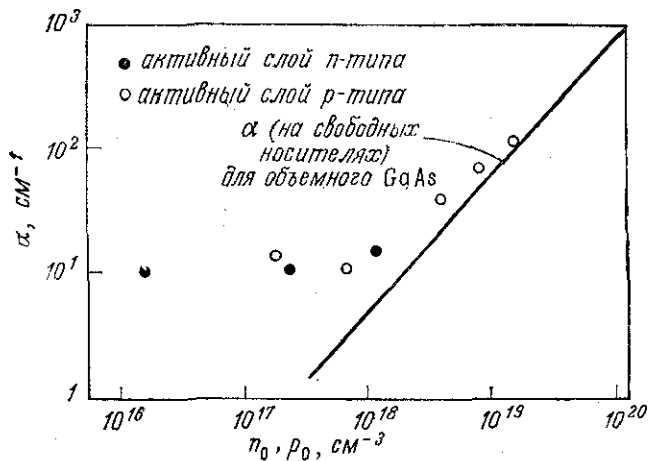


Рис. 3.8.9. Сравнение объемного поглощения на свободных носителях в GaAs и потерь в GaAs—Al_xGa_{1-x}As ДГС-лазере в зависимости от равновесных концентраций носителей в активном слое при комнатной температуре [101].

активный слой слоях Al_xGa_{1-x}As возникают обычно только потери из-за поглощения на свободных носителях $\alpha_{\text{св. нос. } x}$, но может также происходить и межзонное поглощение, если спектр поглощения имеет низкоэнергетический хвост. Поскольку на слой Al_xGa_{1-x}As приходится доля волноводной моды, равная $(1 - \Gamma)$, они дают вклад в α_i , равный $(1 - \Gamma)\alpha_{\text{св. нос. } x}$.

На большое поглощение на свободных носителях указывает трудность достижения вынужденного излучения в непрямозонных полупроводниках. В них коэффициент поглощения при энергии вблизи E_g обычно порядка 10 см^{-1} . Поскольку $g(E)$ связано с $\alpha(E)$, ожидаемые значения коэффициента усиления малы. Думке [12] отметил, что $\alpha_{\text{св. нос}}$ будет превышать $g_{\text{макс}}$ в Ge и Si. Однако каждый полупроводник должен рассматриваться в отдельности, чтобы выяснить, может ли $g_{\text{макс}}$ превосходить $\alpha_{\text{св. нос}}$.

Результаты Пинкаса и др. [101], представленные на рис. 3.8.9, иллюстрируют влияние, которое оказывает наличие свободных носителей на потери в GaAs—Al_xGa_{1-x}As ДГС-лазерах. Предполагалось, что потери на свободных носителях $\alpha_{\text{св. нос}}$ были единственными потерями в этих ДГС-лазерах и активные слои имели достаточно большие толщины, чтобы с большой точностью считать $\Gamma \approx 1,0$ [101]. Как показано на рис. 3.7.7, концентрация инжектированных носителей вблизи $1 \cdot 10^{18} \text{ см}^{-3}$ обеспечивает усиление. Концентрация дырок находится из условия электронейтральности (3.7.9): $p = (N_A^- - N_B^+) + n$. Тогда

при $N_A^- - N_D^+ = 1 \cdot 10^{16} \text{ см}^{-3}$ и $n = 1 \cdot 10^{18} \text{ см}^{-3}$ выражение (3.8.24) дает $\alpha_{\text{св. нос}} = 10 \text{ см}^{-1}$. Как видно из рис. 3.8.9, это значение является нижним пределом изменения $\alpha_{\text{св. нос}}$ при изменении концентрации носителей. В соответствии с выражениями (3.7.9) и (3.8.24) $\alpha_{\text{св. нос}}$ возрастает при увеличении $(N_A^- - N_D^+)$.

Потери на рассеяние α_s обусловлены рассеянием излучения из оптического волновода вследствие непланарности гетерограниц, или из-за несовершенств диэлектрических слоев. Анализ оптического рассеяния в GaAs—Al_xGa_{1-x}As ДГС-лазерах выполнен Томпсоном и др. [102] и Нэшом и др. [103]. Результаты Томпсона и др. [102] показали, что измеренные потери до 12 см^{-1} могут быть объяснены наличием неоднородностей с амплитудой всего 0,01 мкм. Нэш и др. [103] смогли объяснить наблюдаемые высокие значения $J_{\text{пор}}$ деформациями волновода. Рассеяние излучения из волновода может быть вызвано также умышленно созданием периодического гофра, как это обсуждалось в § 10 гл. 2.

Потери за счет связи α_c могут возникать всякий раз, когда оптическое поле проникает за пределы широкозонных ограничивающих слоев [104—106]. Стрейфер и др. [106] исследовали ДГС-лазеры с тонким слоем Al_xGa_{1-x}As, находящимся между активным слоем GaAs и подложкой или верхним слоем из GaAs. В такой структуре существует поток мощности через слой Al_xGa_{1-x}As в слой GaAs. В GaAs—Al_{0,3}Ga_{0,7}As ДГС-лазерах с толщиной активного слоя 0,10—0,15 мкм уменьшение толщины P-Al_{0,3}Ga_{0,7}As-слоя от 0,8 до 0,4 мкм приводило к увеличению $J_{\text{пор}}$ от $1 \cdot 10^3 \text{ А/см}^2$ до $1 \cdot 10^4 \text{ А/см}^2$ (см. рис. 7.4.13).

Все рассмотренные потери объединяют в α_i , стоящее в формуле (3.8.13):

$$\alpha_i = \Gamma \alpha_{\text{св. нос}} + (1 - \Gamma) \alpha_{\text{св. нос, } x} + \alpha_s + \alpha_c. \quad (3.8.25)$$

Для всех мод, кроме основной, Γ обозначает Γ_m — коэффициент ограничения m -й моды. Эти коэффициенты рассматриваются в § 5 гл. 2.

Экспериментальное измерение плотности порогового тока

Полезно рассмотреть поведение ДГС-лазеров при увеличении тока от низких значений, соответствующих спонтанному излучению, до значений, превышающих пороговый ток. На рис. 3.8.10 показана схема опыта для измерения мощности излучения лазера. Выходной ток смещенного в обратном направлении Si-детектора определяется по напряжению V на сопротивлении 50 Ом. Если этот ток разделить на эффективность преобразо-

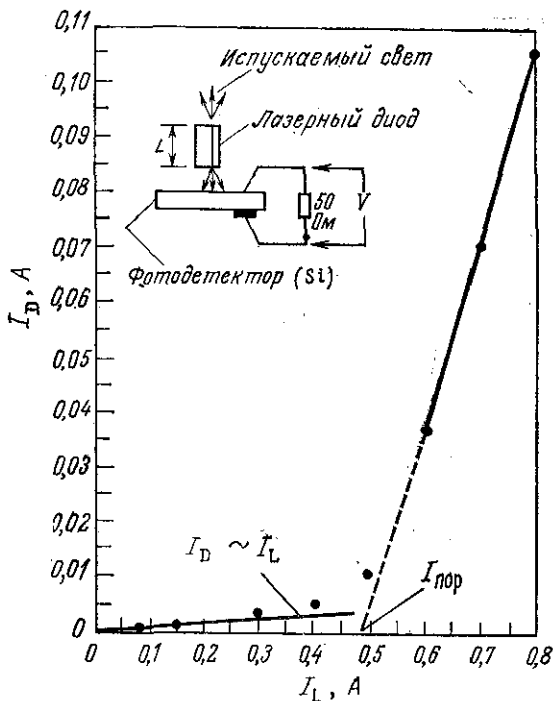


Рис. 3.8.10. Зависимость мощности излучения от тока накачки GaAs— $Al_xGa_{1-x}As$ ДГС-лазера при комнатной температуре. Показана схема измерения. Участок кривой, обозначенный $I_D \sim I_L$, соответствует спонтанному излучению. Пороговое значение $I_{пор}$ определяется, как показано на рисунке, экстраполяцией I_D до нуля.

вания детектора η_c (для большинства Si-детекторов $\eta_c \sim 0,70$ на длинах волн от 0,8 до 0,9 мкм), то получим ток

$$I_D = 2V/\eta_c \cdot 50. \quad (3.8.26)$$

Деление этого выражения на заряд электрона q дает число фотонов, испускаемых с двух торцевых граней кристалла. Коэффициент 2 в выражении (3.8.26) учитывает излучение с обеих граней. На рис. 3.8.10 для GaAs— $Al_xGa_{1-x}As$ ДГС-лазера построена зависимость выходного тока I_D от тока через диод I_L . Этот лазер с широким контактом с выпиленными боковыми сторонами имел ширину 130, длину 360 мкм и толщину активного слоя 0,1 мкм. В рассматриваемом примере лазер не укреплялся на теплоотводе. Накачка велась импульсами длительностью 0,1 мкс с малой частотой повторения, что предотвращало нагревание лазера.

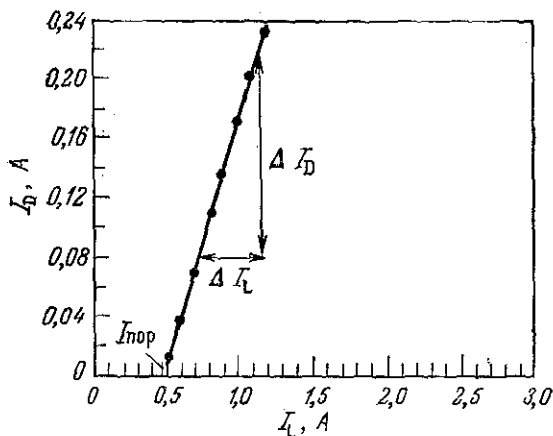


Рис. 3.8.11. Зависимость интенсивности излучения от тока накачки GaAs — $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ДГС-лазера при комнатной температуре. Связь дифференциальной квантовой эффективности η_D с величиной $\Delta I_D/\Delta I_L$ описана в тексте.

При возрастании I_L до $\sim 0,15$ А мощность спонтанного излучения растет линейно, а затем I_D резко возрастает при значении тока около 0,5 А, когда начинается генерация. Экстраполяция этой кривой до нулевого значения I_D , как это выполнено на рис. 3.8.10, дает пороговое значение $I_{\text{пор}}$. Плотность порогового тока равна $I_{\text{пор}} = I_{\text{пор}}/a$. В рассматриваемом случае $a = 4,7 \cdot 10^{-4} \text{ см}^2$ и $I_{\text{пор}} = 1 \cdot 10^3 \text{ А/см}^2$.

Обычно для определения $I_{\text{пор}}$ измеряют ток фотодиода I_D только при токах накачки лазера, превышающих пороговое значение (рис. 3.8.11). Другой полезной величиной, которая может быть определена по графику зависимости I_D от I_L , является дифференциальная квантовая эффективность η_D :

$$\eta_D = \Delta I_D / \Delta I_L. \quad (3.8.27)$$

Смысл стоящих в этом выражении величин ΔI_D и ΔI_L ясен из рис. 3.8.11. Лазер, к которому относятся приведенные на этом рисунке данные, имеет $\eta_D = 0,32$. Обширный обзор данных по $I_{\text{пор}}$ и η_D для разнообразных гетеролазеров дан в гл. 7. Кроме того, в § 7 гл. 7 для лазеров с полосковой геометрией описан специальный метод, позволяющий строить график величины $I dV/dI$. Этот метод показывает, что при возникновении генерации происходит насыщение напряжения, и дает более строгое определение $I_{\text{пор}}$.

На рис. 3.8.12, а показан спектр спонтанного излучения того лазера, к которому относится рис. 3.8.11, при низком уровне накачки. При возрастании тока до значений, лежащих непосредственно перед $I_{\text{пор}}$, возникает суперлюминесценция с $g < \alpha_1$.

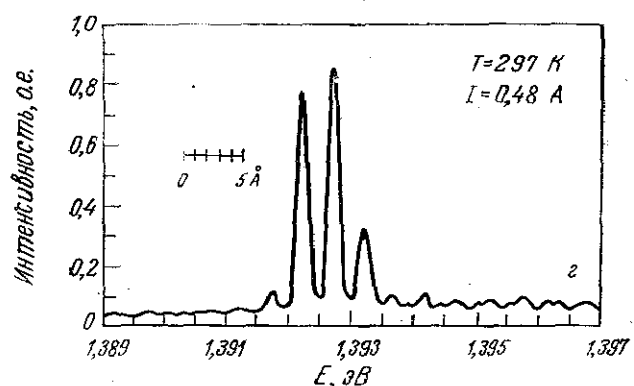
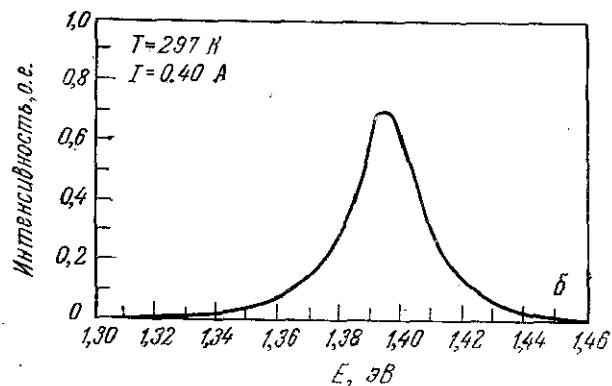
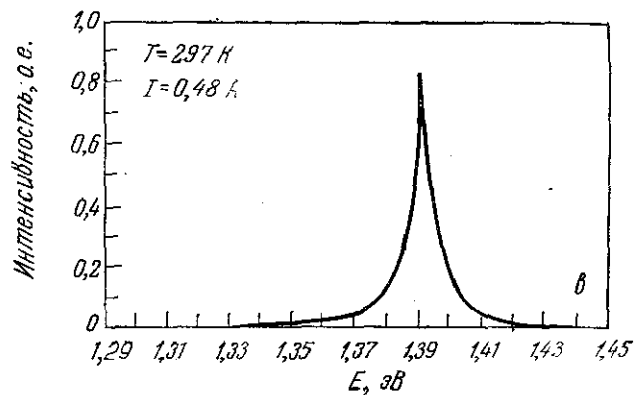
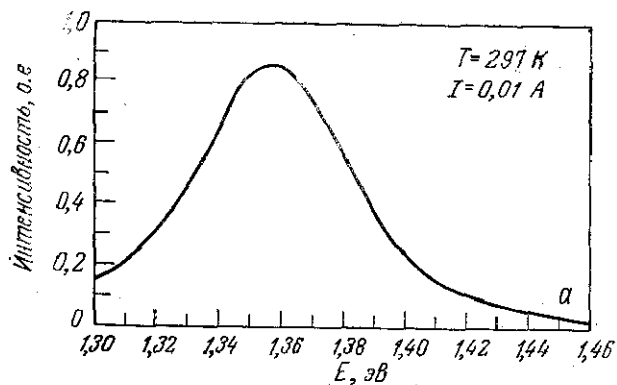


Рис. 3.8.12. Спектры излучения GaAs—Al_xGa_{1-x}As ДГС-лазера. а — спонтанное излучение; б — суперлюминесценция; в — вынужденное излучение; г — приведенный на рис. в спектр в увеличенном масштабе по шкале энергий.

и спектр сужается, как показано на рис. 3.8.12, б. В легированном Si сильно компенсированном активном слое инжектированные электроны продолжают заполнять состояния в хвосте зоны проводимости и тем самым сдвигают пик интенсивности излучения в сторону более высоких энергий [28], что хорошо видно из сравнения рис. 3.8.12, а и б. На рис. 3.8.12, б показано, как непосредственно за порогом $I_{\text{пор}}$ в спектре суперлюминесценции возникает очень узкая линия излучения. На рис. 3.8.12, г приведен спектр лазерной генерации в увеличенном масштабе по энергии. Показанные на нем линии излучения являются продольными модами, удовлетворяющими фазовому условию (3.8.9), с межмодовым расстоянием, определяемым выражением (3.8.11). По данным работы [63] величина $d\bar{n}/d\lambda_0$ равна $-1,0 \cdot 10^4 \text{ см}^{-1}$, так что $1 - (\lambda_0/\bar{n}) d\bar{n}/d\lambda_0 = 1,25$. Это дает межмодовое расстояние 2,5 Å, что и наблюдается на рис. 3.8.12, г. Межмодовое расстояние для продольных мод и дисперсия в широком диапазоне длин волн проиллюстрированы для гомолазеров на рис. 2 и 3 работы [18].

Дифференциальная квантовая эффективность

Измерения плотности порогового тока и дифференциальной квантовой эффективности, определяемой выражением (3.8.27), являются двумя наиболее часто встречающимися измерениями при исследовании полупроводниковых лазеров. Полная внешняя квантовая эффективность равна отношению I_D/I_L (рис. 3.8.11). Дифференциальная квантовая эффективность определяется производной интенсивности излучения $\mathcal{I}_{\text{внешн}}$ по току накачки:

$$\eta_D = q (d\mathcal{I}_{\text{внешн}}/dI_L). \quad (3.8.28)$$

При достижении порога интенсивность спонтанного излучения насыщается, и выше порога полная интенсивность излучения может быть представлена в виде [18, 107]

$$q\mathcal{I}_{\text{внешн}} = F_3\eta_{\text{спонт}}I_{\text{пор}} + F_1\eta_{\text{вынужд}}(I - I_{\text{пор}}), \quad (3.8.29)$$

где F_3 и F_1 — вероятности выхода фотонов спонтанного и вынужденного излучений соответственно, $\eta_{\text{спонт}}$ и $\eta_{\text{вынужд}}$ — внутренние квантовые эффективности для спонтанного и вынужденного излучений, считающиеся постоянными при $I > I_{\text{пор}}$.

Выражения (3.8.28) и (3.8.29) показывают, что

$$\eta_D = \eta_{\text{вынужд}}F_1, \quad (3.8.30)$$

причем вероятность выхода F_1 равна отношению оптической мощности, выходящей из лазера наружу, к полной оптической мощности, генерируемой внутри лазера:

$$F_1 = P_{\text{внешн}}/(P_{\text{внешн}} + P_{\text{погл}}). \quad (3.8.31)$$

Индекс 1 подчеркивает «одномерность» поглощения лазерных мод в отличие от изотропного «трехмерного» случая для спонтанного излучения.

Байард и др. [108] получили выражения для F_1 на пороге, где удобно представить мощность, выходящую наружу, через потери на зеркалах α_m , в то время как полная мощность представляется через сумму α_m и внутренних потерь α_i . Усиление на пороге является суммой внутренних и внешних потерь, как это следует из формулы (3.8.13), в которой величина $(1/L)\ln(1/R)$ представляет собой потери на зеркалах. При этих условиях выражение (3.8.31) принимает вид

$$F_1 = \frac{(1/L)\ln(1/R)}{\alpha_i + (1/L)\ln(1/R)}, \quad (3.8.32)$$

а выражение (3.8.30) дает

$$\eta_D = \eta_{\text{вынужд}} \left/ \left[1 + \frac{\alpha_i L}{\ln(1/R)} \right] \right. . \quad (3.8.33)$$

Строя график зависимости $1/\eta_D$ от длины лазерного диода по экспериментальным значениям η_D , измеренным с учетом излучения с обоих зеркал, можно получить $1/\eta_{\text{вынужд}}$ как значение $1/\eta_D$ при $L=0$, получаемое экстраполяцией. По формуле (3.8.33) с этим значением $\eta_{\text{вынужд}}$ можно затем определить потери α_i . Данные по поглощению на свободных носителях, представленные на рис. 3.8.9, получены именно таким способом (рис. 7.4.2). Измеренные с учетом излучения с двух зеркал значения η_D обычно считают превосходящими в 2 раза значения η_D , полученные с учетом излучения только с одного зеркала.

Выражения для пороговой плотности тока

Вывод полученных вначале для гомолазеров выражений, описывающих пороговую плотность тока, был основан на использовании коэффициента усиления для двухуровневой системы [109, 110]. В эти выражения входила такая величина, как $\bar{n}^2 E^2 \Delta E / c^2$, в которой ΔE обозначает полуширину спектра спонтанного излучения. Однако получены они были для $T=0$ К и не содержали температурной зависимости, имеющей важное значение при комнатной температуре. Основная форма этих выражений была сохранена Лэшером и Стерном [8] при выводе температурной зависимости плотности порогового тока. Аналогичные выражения, описывающие температурную зависимость $J_{\text{пор}}$, выводятся ниже.

Из выражений (3.8.14) и (3.8.15) следует

$$J_{\text{ном}} \text{ (A/cm}^2\text{)} = qd \int_0^\infty r_{\text{спонт}}(E) dE. \quad (3.8.34)$$

Поскольку $g = -\alpha$, выражение (3.8.34) с учетом выражения (3.2.41) может быть переписано в виде

$$J_{\text{ном}} (\text{А/см}^2) = \frac{qg(E_{m,g}) d \int_0^{\infty} r_{\text{спонт}}(E) dE}{(h^3 c^2 / 8\pi \bar{n}^2 E_{m,g}^2) r_{\text{вынужд}}(E_{m,g})}, \quad (3.8.35)$$

где $E_{m,g}$ обозначает энергию фотона, при которой усиление максимально. Используя формулу (3.2.47), связывающую $r_{\text{вынужд}}(E)$ и $r_{\text{спонт}}(E)$, получаем

$$J_{\text{ном}} (\text{А/см}^2) = \frac{q \bar{n}^2 E_{m,g}^2 g(E_{m,g}) \Delta E d \int_0^{\infty} r_{\text{спонт}}(E) dE}{(\pi^2 c^2 \hbar^3) \Delta E r_{\text{спонт}}(E_{m,g}) \{1 - \exp[(E_{m,g} - \Delta F)/kT]\}}. \quad (3.8.36)$$

Здесь ΔE обозначает полуширину спектра спонтанного излучения. Стерн [50] переписал формулу (3.8.36) следующим образом:

$$J_{\text{ном}} (\text{А/см}^2) = (dq \bar{n}^2 E_{m,g}^2 \gamma \Delta E / \pi^2 c^2 \hbar^3) g(E_{m,g}), \quad (3.8.37)$$

где температурная зависимость и нелинейность коэффициента усиления включены в γ , которое определяется выражением

$$\gamma = \int_0^{\infty} \frac{r_{\text{спонт}}(E) dE}{r_{\text{спонт}}(E_{m,\text{сп}}) \Delta E} \cdot \frac{r_{\text{спонт}}(E_{m,\text{сп}})}{r_{\text{спонт}}(E_{m,g})} \cdot \frac{1}{\{1 - \exp[(E_{m,g} - \Delta F)/kT]\}}, \quad (3.8.38)$$

где $E_{m,\text{сп}}$ — энергия фотона, соответствующая максимуму спектра спонтанного излучения. Выражение (3.8.38) было представлено [50] в виде произведения $\gamma = \gamma_1 \gamma_2 \gamma_3$, где

$$\gamma_1 = \int_0^{\infty} \frac{r_{\text{спонт}}(E) dE}{r_{\text{спонт}}(E_{m,\text{сп}}) \Delta E}, \quad (3.8.39)$$

$$\gamma_2 = r_{\text{спонт}}(E_{m,\text{сп}}) / r_{\text{спонт}}(E_{m,g}), \quad (3.8.40)$$

$$\gamma_3 = 1 / \{1 - \exp[(E_{m,g} - \Delta F)/kT]\}. \quad (3.8.41)$$

Первый множитель γ_1 является форм-фактором и имеет порядок единицы. Как температурная зависимость γ , так и связь коэффициента усиления с током содержатся в величинах γ_2 и γ_3 . Обе они при $T = 0$ К равны единице. Для образца с $N_A = 6 \cdot 10^{18} \text{ см}^{-3}$ и $N_D = 3 \cdot 10^{18} \text{ см}^{-3}$ Стерн [50] получил $\gamma = 9,2$ при $T = 300$ К и коэффициенте усиления 50 см^{-1} .

Если волноводная мода не полностью заключена в активном слое, то вместо коэффициента усиления надо использовать

величину $\Gamma g(E)$ (Γ — коэффициент оптического ограничения).
Условие генерации (3.8.13) переписывается тогда в виде

$$\Gamma g(E_{m, g}) = \alpha_i + (1/L) \ln(1/R). \quad (3.8.42)$$

Из этого уравнения и выражения (3.8.25) следует

$$g(E_{m, g}) = (1/\Gamma) [\Gamma \alpha_{\text{св. нос}} + (1 - \Gamma) \alpha_{\text{св. нос, x}} + \alpha_s + \alpha_c + (1/L) \ln(1/R)]. \quad (3.8.43)$$

Выражение для пороговой плотности тока можно получить из соотношения (3.8.37), если разделить его на внутреннюю квантовую эффективность η и использовать для коэффициента усиления на пороге выражение (3.8.43):

$$J_{\text{пор}} (\text{А/см}^2) = (d/\eta\Gamma) (10^{-4} q \bar{n}^2 E_{m, g}^2 \gamma \Delta E / \pi^2 c^2 \hbar^3) \times \\ \times [\Gamma \alpha_{\text{св. нос}} + (1 - \Gamma) \alpha_{\text{св. нос, x}} + \alpha_s + \alpha_c + (1/L) \ln(1/R)], \quad (3.8.44)$$

где d выражено в микрометрах. Следует подчеркнуть, что в $J_{\text{пор}}$ содержится и температурная зависимость, и зависимость коэффициента усиления от тока.

Интересно численно оценить значение $J_{\text{пор}}$ (300 К) для однородно накачанного слоя p -GaAs с $p_0 \approx 1 \cdot 10^{18} \text{ см}^{-3}$, $d = 1 \text{ мкм}$, $L = 400 \text{ мкм}$, $\eta = 1$ и $\Gamma = 1$. Для GaAs $\bar{n} = 3,6$, $R = 0,3$ для границы воздух — GaAs, $\gamma = 9,2$ [50], $E_{m, g} = 1,40 \text{ эВ}$ и $\Delta E = 0,06 \text{ эВ}$ [51]. Из данных рис. 3.8.9 находим $\alpha_{\text{св. нос}} \approx 10 \text{ см}^{-1}$, так что $[\alpha_{\text{св. нос}} + (1/L) \ln(1/R)] = 40 \text{ см}^{-1}$. Другими потерями пренебрегаем. Тогда по формуле (3.8.44) получаем $J_{\text{пор}} (300 \text{ К})/d = 3,5 \cdot 10^3 \text{ А/см}^2 \cdot \text{мкм}$. Пинкас и др. [92] экспериментально получили $J_{\text{пор}} (300 \text{ К})/d = 5 \cdot 10^3 \text{ А/см}^2 \cdot \text{мкм}$.

Обычно из приведенных выше формул для объяснения экспериментальных результатов получают другие выражения. Выражения для порогового тока могут быть получены для квадратичной и линейной зависимостей коэффициента усиления от тока, определяемых формулами (3.8.20) и (3.8.21). Это легко сделать, подставляя в выражение (3.8.42) значения $g(E_{m, g})$, полученные по формулам (3.8.20) и (3.8.21), и используя соотношение (3.8.18), связывающее $J_{\text{пор}}$ с $J_{\text{ном}}$. При этом для квадратичной зависимости коэффициента усиления от тока будем иметь

$$J_{\text{пор}} (\text{А/см}^2) = \frac{J_0 d}{\eta} + \frac{d}{\eta} \left[\frac{1}{\Gamma \beta} \left(\alpha_i + \frac{1}{L} \ln \frac{1}{R} \right) \right]^{1/2}, \quad (3.8.45)$$

а для линейной зависимости

$$J_{\text{пор}} (\text{А/см}^2) = \frac{J_0 d}{\eta} + \frac{d}{\eta \Gamma \beta} \left(\alpha_i + \frac{1}{L} \ln \frac{1}{R} \right), \quad (3.8.46)$$

где α_i определяется формулой (3.8.25). В § 4 гл. 7 показано, что выражение (3.8.46) наиболее пригодно и удобно для описания зависимости $I_{\text{пор}}$ от d и x в $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ ДГС-лазерах. В § 5 и 4 гл. 7 проводится сравнение экспериментальных значений $I_{\text{пор}}$ с расчетными, полученными при использовании для $g_{\text{макс}}$ выражения (3.8.7), и согласие между ними найдено очень хорошим.

§ 9. ЗАКЛЮЧЕНИЕ

В этой главе изложены основы теории вынужденного излучения в полупроводниках и даны выражения для численного расчета плотности порогового тока. Много внимания уделено соотношениям между поглощением, вынужденным излучением и спонтанным излучением. Показано, что необходимым условием вынужденного излучения является превышение энергетическим интервалом между квазиуровнями Ферми для электронов и дырок энергии испускаемых фотонов. Для достижения порога генерации нужно, чтобы усиление превышало потери на внешнее излучение плюс потери внутри резонатора, такие, как поглощение на свободных носителях и рассеяние.

Сравнение экспериментальных и расчетных спектров поглощения демонстрирует необходимость учета образования хвостов зон и зависимости матричного элемента от энергии в теоретической модели, описывающей оптические спектры. Представленная в этой главе теоретическая модель позволяет рассчитать коэффициент усиления как функцию тока накачки и температуры. Эти вычисления дают представление о поведении усиления в прямозонных полупроводниках. Удивительно, насколько хорошо такой сложный предмет, как вынужденное излучение в полупроводниках, может быть описан выражениями, выведенными из основных физических законов. Численные оценки по этим выражениям сделаны без привлечения свободных параметров.

ЛИТЕРАТУРА

1. Einstein A., Phys. Zs., 18, 121 (1917).
2. Bernard M. G., Duraffourg G., Phys. Stat. Solidi, 1, 699 (1961).
3. White R. L., Basic Quantum Mechanics, McGraw-Hill, New York, 1966, p. 223.
4. Kane E. O., Phys. Rev., 131, 79 (1963).
5. Halperin B. I., Lax M., Phys. Rev., 148, 722 (1966).
6. Eagles D. M., Journ. Phys. Chem. Solids, 16, 76 (1960).
7. Dumke W. P., Phys. Rev., 132, 1998 (1963).
8. Lasher G., Stern F., Phys. Rev., 133, A553 (1964).
9. Stern F., Phys. Rev., 148, 186 (1966).
10. Hwang C. J., Phys. Rev., B2, 4126 (1970).

- 11) Stern F., Phys. Rev., **B3**, 2636 (1971).
12. Dumke W. P., Phys. Rev., **127**, 1559 (1962).
13. McCumber D. E., Phys. Rev., **136**, A954 (1964).
14. Kestin J., Dorfman J. R., A Course in Statistical Thermodynamics, Academic Press, New York, 1971, p. 426.
15. Joyce W. B., Phys. Rev., **D9**, 3234 (1974).
16. Mandl F., Statistical Physics, Wiley, London, 1971, p. 325.
17. Mandl F., Statistical Physics, Wiley, London, 1971, p. 264.
18. Stern F., Semiconductors and Semimetals (Willardson R. K., Beer A. C., eds.), Academic Press, New York, 1966, vol. 2, p. 371.
19. Stern F., Solid State Physics (Seitz F., Turnbull D., eds), Academic Press, New York, 1963, vol. 15, p. 300.
20. Bebb H. B., Williams E. W., Semiconductors and Semimetals (Willardson R. K., Beer A. C., eds), Academic Press, New York, 1972, vol. 8, p. 181.
21. Blakemore J. S., Semiconductors Statistics, Pergamon Press, New York, 1962, Chapters 1 and 2. (Имеется перевод: Блэкмор Дж. Статистика электронов в полупроводниках. — М.: Мир, 1964.)
22. McKelvey J. P., Solid State and Semiconductor Physics, Harper, New York, 1966.
23. Bube R. H., Electronic Properties of Crystalline Solids. An Introduction to Fundamentals, Academic Press, New York, 1974.
24. Nathan M. I., Morgan T. N., Physics of Quantum Electronics (Kelley P. L., Lax V., Tannenwald P. E., eds), McGraw-Hill, New York, 1966, p. 478.
25. Southgate P. D., Journ. Appl. Phys., **40**, 5333 (1969).
26. Redfield D., Wittke J. P., Pankove J. I., Phys. Rev., **B2**, 1830 (1970).
27. Ocunov B. B., ФТП, **7**, 2106 (1973).
28. Casey H. C., Jr., Bachrach R. L., Journ. Appl. Phys., **44**, 2795 (1973).
29. Pearson G. L., Bardeen J., Phys. Rev., **75**, 865 (1949).
30. Емельяненко О. В. и др., ФТТ, **7**, 1315 (1965).
31. Ermanis F., Wolfstirn K., Journ. Appl. Phys., **37**, 1963 (1966).
32. Bonch-Bruевич V. L., Semiconductors and Semimetals (Willardson R. K., Beer A. C., eds), Academic Press, New York, 1966, vol. 1, p. 101.
33. Summers C. J., Dingle R., Hill D. E., Phys. Rev., **B1**, 1603 (1970).
34. White A. M. et al., Journ. Phys., **C6**, L243 (1973).
35. Mott N. F., Twose W. D., Advan. Phys., **10**, 107 (1961).
36. Strzalkowski I., Joshi S., Crowell C. R., Appl. Phys. Lett., **28**, 350 (1976).
37. Vrehen J. H. F., Journ. Chem. Solids, **29**, 129 (1968).
38. Brews J. R., Hwang C. J., Journ. Chem. Phys., **54**, 3263 (1971).
39. Stern F., Solid State Physics (Seitz F., Turnbull D., eds), Academic Press, New York, 1963, Vol. 15, p. 398.
40. Zittartz J., Langer J. S., Phys. Rev., **148**, 741 (1966).
41. Шкловский Б. И., Эфрос А. Л., ФТП, **4**, 305 (1970).
42. Saitoh M., Edwards S. F., Journ. Phys., **C7**, 3937 (1974).
43. Morgan T. N., Phys. Rev., **139**, A343 (1965).
44. Келдыш Л. В., Прошко Г. П., ФТТ, **5**, 3378 (1963).
45. Бонч-Бруевич В. Л., ДАН СССР, **189**, 505 (1969).
46. Stern F., Conduction in Low-Mobility Materials, (Klein N., Pollak M., eds), Taylor and Francis, London, 1971, p. 211.
47. Panish M. B., Casey H. C., Jr., Journ. Phys. Chem. Solids, **28**, 1673 (1967).
48. Shockley W., Electrons and Holes in Semiconductors, Van Nostrand, Princeton, New Jersey, 1950, p. 242. (Имеется перевод: Шокли В. Теория электронных полупроводников. — М.: ИЛ, 1953.)
49. Hwang C. J., Phys. Rev., **B2**, 4117 (1970).
50. Stern F., Laser Handbook (Arecchi F. T., Schulz-DuBois E. O. eds), North-Holland Publ., Amsterdam, 1972, p. 425.
- 51) Casey H. C., Jr., Stern F., Journ. Appl. Phys., **47**, 631 (1976).
52. Hwang C. J., Journ. Appl. Phys., **41**, 2668 (1970).
53. Eymard R., Duraffourg G., Journ. Phys., **D6**, 66 (1973).

54. *Smith R. A.*, Wave Mechanics of Crystalline Solids, Chapman Hall, London, 1961, p. 406.
55. *Callaway J.*, Energy Band Theory, Academic Press, New York, 1964, 240 (Имеется перевод: *Каллуэй Дж.* Теория энергетической зонной структуры — М.: Мир, 1969.)
56. *Bube R. H.*, Electronic Properties of Crystalline Solids. An Introduction to Fundamentals, Academic Press, New York, 1974, p. 122.
57. *Kane E. O.*, Journ. Phys. Chem. Solids, **1**, 249 (1957).
58. *Leighton R. B.*, Principles of Modern Physics, McGraw-Hill, New York, 1959, p. 183.
59. *Stern F.*, Journ. Non-Cryst. Solids, **4**, 256 (1970).
60. *Stern F.*, IEEE Journ. Quantum Electron., **QE-9**, 290 (1973).
61. *Stern F.*, Journ. Appl. Phys., **47**, 5382 (1976).
62. *Casey H. C., Jr., Sell D. D., Wecht K. W.*, Journ. Appl. Phys., **46**, 250 (1975).
63. *Sell D. D., Casey H. C., Jr., Wecht K. W.*, Journ. Appl. Phys., **45**, 2650 (1974).
64. *van Roosbroeck W., Shockley W.*, Phys. Rev., **94**, 1558 (1954).
65. *Burstein E.*, Phys. Rev., **93**, 632 (1954).
66. *Vieland L. J., Kudman I.*, Journ. Phys. Chem. Solids, **24**, 437 (1963).
67. *Abrahams M. S., Buiocchi S. J., Tietjen J. J.*, Journ. Appl. Phys., **38**, 760 (1967).
68. *Dean P. J.*, частное сообщение.
69. *Chadi D. J., Clark A. H., Burnham R. D.*, Phys. Rev., **B13**, 4466 (1976).
70. *Dow J. D., Smith D. L., Lederman F. L.*, Phys. Rev., **B8**, 4612 (1973).
71. *Brinkman W. F., Lee P. A.*, Phys. Rev. Lett., **31**, 237 (1973).
72. *Varshni Y. P.*, Phys. Status Solidi, **19**, 459 (1967).
73. *Hall R. N.*, Proc. Inst. Elec. Eng., **B106**, 923 (1959).
74. *Casey H. C., Jr., Panish M. B., Chang L. L.*, Phys. Rev., **162**, 660 (1967).
75. *Hwang C. J., Brews J. R.*, Journ. Phys. Chem. Solids, **32**, 837 (1971).
76. *Casey H. C., Jr., Miller B. I., Pinkas E.*, Journ. Appl. Phys., **44**, 1281 (1973).
77. *Prince M. B.*, Phys. Rev., **92**, 681 (1953).
78. *James L. W., Moll J. L., Spicer W. E.*, Gallium Arsenide: 1968 Symposium Proceedings, Inst. Phys. and Phys. Soc., London, 1969, p. 230.
79. *Schade H., Neison H., Kressel H.*, Appl. Phys. Lett., **18**, 121 (1971).
80. *Abrahams M. S., Buiocchi C. J., Williams B. F.*, Appl. Phys. Lett., **18**, 220 (1971).
81. *Kawakami T., Sugiyama K.*, Jpn. Journ. Appl. Phys., **12**, 151 (1973).
82. *Acket G. A., Nijman W., Lam H.*, Journ. Appl. Phys., **45**, 3033 (1974).
83. *Ashley K. L., Carr D. L., Romano-Moran R.*, Appl. Phys. Lett., **22**, 23 (1973).
84. *Ettenberg M., Kressel H., Gilbert S. L.*, Journ. Appl. Phys., **44**, 827 (1973).
85. *van Opdorp C., Peters R. C., Klerk M.*, Appl. Phys. Lett., **24**, 125 (1974).
- 85a. *Asbeck P.*, Journ. Appl. Phys., **48**, 820 (1977).
86. *Kuriyama I., Kamiya T., Yanai H.*, Jpn. Journ. Appl. Phys., **16**, 465 (1977).
87. *Stern F.*, частное сообщение.
88. *Aleksanian A. G., Poluektov I. A., Popov Yu. M.*, IEEE Journ. Quantum Electron., **QE-10**, 297 (1974).
89. *Fabry C., Perot A.*, Ann. Chim. Phys., **16**, 115 (1899).
90. *Yariv A.*, Introduction to Optical Electronics, Holt, New York, 1976, 2-nd ed., p. 114.
91. *Unger K.*, Zs. Phys., **207**, 322 (1967).
92. *Pinkas E. et al.*, Journ. Appl. Phys., **43**, 2827 (1972).
93. *Hakki B. W., Paoli T. L.*, Journ. Appl. Phys., **46**, 1299 (1975).
94. *Goodwin A. R. et al.*, Journ. Appl. Phys., **46**, 3126 (1975).
95. *Rode D. L.*, Journ. Appl. Phys., **45**, 3887 (1974).
96. *Hayashi I., Panish M. B., Reinhart F. K.*, Journ. Appl. Phys., **42**, 1929 (1971).

97. *Fan H. Y.*, Semiconductors and Semimetals (*Willardson R. K.*, *Beer A. C.*, eds), Academic Press, New York, 1967, vol 3, p. 405. (Имеется перевод в кн.: Оптические свойства полупроводников/Под ред. Р. Уиллардсона и А. Бира. — М.: Мир, 1970, с. 385.)
98. *Hill D. E.*, Phys. Rev., **133**, A866 (1964).
99. *Spitzer W. G.*, *Whelan J. M.*, Phys. Rev., **114**, 59 (1959).
100. *Merz J. L.*, *Logan R. A.*, *Sergent A. M.*, Journ. Appl. Phys., **47**, 1436 (1976).
101. *Pinkas E. et al.*, IEEE Journ. Quantum Electron., **QE-9**, 281 (1973).
102. *Thompson G. H. B.*, *Kirkby P. A.*, *Whiteaway I. E. A.*, IEEE Journ. Quantum Electron., **QE-11**, 481 (1975).
103. *Nash F. R.*, *Wagner W. R.*, *Brown R. L.*, Journ. Appl. Phys., **47**, 3992 (1976).
104. *Casey H. C., Jr.*, *Panish M. B.*, Journ. Appl. Phys., **46**, 1393 (1975).
105. *Butler J. K.*, *Kressel H.*, *Ladany J.*, IEEE Journ. Quantum Electron., **QE-11**, 402 (1975).
106. *Streifer W.*, *Burnham R. D.*, *Seifres D. R.*, IEEE Journ. Quantum Electron., **QE-12**, 177 (1976).
107. *Cheroff G.*, *Stern F.*, *Triebwasser S.*, Appl. Phys. Lett., **2**, 173 (1963).
108. *Biard J. R.*, *Carr W. N.*, *Reed B. S.*, Trans. AIME, **230**, 286 (1964).
109. *Lasher G. J.*, IBM Journ., **7**, 58 (1963).
110. *Nathan M. I.*, Appl. Opt., **5**, 1514 (1966).

ГЕТЕРОПЕРЕХОДЫ

§ 1. ВВЕДЕНИЕ

Гетеропереходом называется созданный в одном кристалле переход между двумя различными полупроводниками. Различие между полупроводниками обычно заключается в их зонной структуре, наиболее важной характеристикой которой является ширина запрещенной зоны. Гетеропереход называется изотипным, если он образован между полупроводниками с одинаковым типом проводимости, и анизотипным, если образующие его полупроводники обладают разными типами проводимости. При описании гетеропереходов удобно обозначать тип проводимости полупроводника с узкой запрещенной зоной строчными буквами n или p , а полупроводника с широкой запрещенной зоной — прописными буквами N или P . По используемому здесь соглашению при обозначении гетероперехода тип проводимости полупроводника с более узкой запрещенной зоной мы будем писать первым, так что изотипный гетеропереход будет записываться как $n-N$ или $p-P$, а анизотипный гетеропереход — как $n-P$ или $p-N$.

Первыми публикациями по гетеропереходам были теоретические работы Губанова [1—3], появившиеся в 1951 г. В том же году в патенте Шокли [4] было предложено использовать два различных полупроводника для изготовления эмиттера и базы транзистора. В 1957 г. Кремер [4] опубликовал статью по теории широкозонного эмиттера для транзисторов. Гетеропереходы хорошего качества не были созданы вплоть до 1960 г., когда Андерсон [6] изготовил туннельные диоды на переходе Ge-GaAs . Германий и арсенид галлия имеют почти одинаковую постоянную решетки при комнатной температуре. Исследуя Ge-GaAs -гетеропереходы, Андерсон предложил для их описания модель [6, 7], в которой разрыв зоны проводимости на границе раздела полупроводников полагался равным разности их значений электронного сродства. Электронное сродство определяется как разность между энергиями дна зоны проводимости и уровня вакуума над свободной поверхностью. Кремер [8] отметил, что желательно рассчитать взаимное расположение зон в резком гетеропереходе, исходя из объемных свойств обоих полупроводников, а не привлекая характеристики свобод-

ной поверхности, например, такие, как электронное сродство. В теоретических расчетах взаимного расположения зон в гетеропереходе, использующих объемные зонные структуры полупроводников, был достигнут определенный прогресс [9]. Значительные экспериментальные и теоретические усилия направлены в настоящее время на создание более совершенных моделей, описывающих гетеропереходы. Можно ожидать, что полученные результаты заставят пересмотреть и усовершенствовать «модель Андерсона», которая используется в этой главе. Однако изложенные здесь концепции позволяют ввести в рассмотрение основные понятия, которые будут полезны при дальнейшем развитии наших представлений о гетеропереходах, происходящем в результате продолжающихся экспериментальных и теоретических исследований.

В этой главе рассматриваются те свойства гетеропереходов, которые имеют отношение к гетеролазерам. Гетеропереходы между полупроводниками с разными постоянными решетки характеризуются такими плотностями центров безызлучательной рекомбинации, находящиеся на границе раздела полупроводников, которые сделают недостижимо высокими плотности носителей и фотонов, необходимые для получения вынужденного излучения. Поэтому мы рассматриваем только изопериодические гетеропереходы без поверхностных состояний на границе раздела. Поскольку в настоящее время наиболее широко используются GaAs—Al_xGa_{1-x}As-гетеролазеры и их свойства лучше всего изучены, приведенные в этой главе примеры относятся к этой комбинации полупроводников. Идеи, иллюстрируемые здесь на основе структуры GaAs—Al_xGa_{1-x}As, легко могут быть распространены на другие изопериодические пары, такие, как InP—Ga_xIn_{1-x}P_yAs_{1-y}, GaAs_ySb_{1-y}—Al_xGa_{1-x}As_ySb_{1-y} и GaAs—Ga_xIn_{1-x}P. Составы с совпадающими периодами решеток обсуждаются в гл. 5. Полные обзоры имеющейся литературы по гетеропереходам даны в работах Милнса и Фойхта [10] и Шарма и Пурохита [11]. Эти работы можно использовать для более подробного ознакомления с отдельными вопросами; в них даны также ссылки на соответствующую литературу. Полезны для этой цели и материалы конференции по гетеропереходам 1970 г. [12].

Для построения энергетических зонных диаграмм гетеропереходов необходимо знать зонные структуры образующих их полупроводников. Для тройных или четверных полупроводниковых твердых растворов необходимо также знать зависимость их зонной структуры от состава. В следующем параграфе мы вначале определим параметры, необходимые для описания GaAs и AlAs, а затем установим зависимость этих параметров от состава x для Al_xGa_{1-x}As.

В § 3 выводятся выражения, описывающие гетеропереход в модели Андерсона, и они иллюстрируются на примере гетеропереходов $p-N$, $n-P$, $p-P$ и $n-N$. Влияние градиента состава на энергетическую зонную диаграмму рассматривается в § 4. Выражения, описывающие поведение вольт-емкостных и вольт-амперных характеристик, выводятся в § 5. Одним из наиболее важных свойств гетеропереходов является ограничение для носителей тока, создаваемое более широкозонным полупроводником. В § 6 рассматриваются потенциальные барьеры, создающие ограничение для электронов и дырок, а также токи утечки через них. Экспериментальные результаты представлены в гл. 7.

§ 2. ЗОННАЯ СТРУКТУРА ГЕТЕРОПЕРЕХОДА

$\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$

GaAs

Основы расчета полупроводниковой зонной структуры могут быть найдены в таких широкоизвестных книгах, как книги Лонга [13] и Харрисона [14]. На рис. 4.2.1 показаны главные особенности энергетической зонной структуры GaAs [15]. Кривые на этом рисунке представляют электронную энергию как функцию волнового вектора k . Зонная структура зависит от направления волнового вектора в кристалле. Верхние зоны соответствуют зоне проводимости, нижние — валентной зоне. В GaAs, являющемся прямозонным материалом, минимум энергии в зоне проводимости и максимум энергии в валентной зоне отвечают одному и тому же значению k ($k=0$). В непрямозонном материале минимум зоны проводимости и максимум валентной зоны соответствуют разным значениям k . В прямозонном материале электрон может совершить переход ($\Gamma_8 \rightarrow \Gamma_6$) из валентной зоны в зону проводимости (или обратный) непосредственно при поглощении (или испускании) фотона, в то время как в непрямозонном материале, таком, например, у которого минимум X_6 или L_6 расположен ниже минимума Γ_6 , переход с минимальной энергией должен сопровождаться для сохранения импульса испусканием или поглощением фонона. Следовательно, не прямой переход по своей природе является менее вероятным процессом, чем прямой переход.

Ширина прямой запрещенной зоны $E_g^\Gamma(\Gamma_8 \rightarrow \Gamma_6)$ арсенида галлия высокой чистоты равна 1,424 эВ при комнатной температуре [16]. Ее зависимость от температуры может быть выражена формулой [17]

$$E_g(T) = 1,519 - 5,405 \cdot 10^{-4} T^2 / (204 + T). \quad (4.2.1)$$

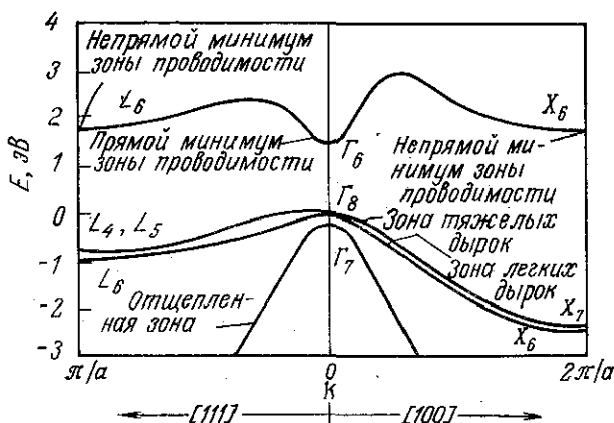


Рис. 4.2.1. Зонная структура GaAs. Энергия E представлена как функция волнового вектора k в направлениях $[100]$ и $[111]$. Обозначения Γ , X и L соответствуют разным межзонным энергетическим зазорам.

Как говорилось в § 7 гл. 3, запрещенная зона сужается при увеличении концентраций носителей. Это сужение составляет примерно $0,013$ эВ для концентрации дырок $p = 1 \cdot 10^{18}$ см $^{-3}$. В настоящей главе мы не будем усложнять анализ вопросов, касающихся теории гетеропереходов, учетом сужения запрещенной зоны. Относительно положения непрямых минимумов X и L зоны проводимости существовали значительные разногласия. До экспериментов Аспнеса [18] по электроотражению синхротронного излучения от барьера Шоттки полагали, что низшим непрямым минимумом зоны проводимости является минимум X , отделенный от прямого минимума энергией $0,38$ эВ при комнатной температуре [19]. Измерения Аспнеса [18] показали, что энергетические зазоры между максимумом валентной зоны и непрямыми минимумами L и X зоны проводимости равны соответственно $E_g^L = 1,708$ эВ и $E_g^X = 1,900$ эВ при комнатной температуре.

Электроны и дырки под действием электрического поля движутся подобно свободным частицам с отрицательными и положительными зарядами и соответствующими эффективными массами. Эффективные массы электрона (m_n) и дырки (m_p) определяются зависимостью E от k и находятся по формулам [20, 21]

$$m_n = \hbar^2 / [\partial^2 E_c(k) / \partial k^2], \quad (4.2.2)$$

$$m_p = -\hbar^2 / [\partial^2 E_v(k) / \partial k^2], \quad (4.2.3)$$

где $E_c(k)$ — край зоны проводимости, а $E_v(k)$ — край валентной зоны. Чем более пологая кривая зависимости E от k в ми-

Таблица 4.2.1

Значения энергетических зазоров и эффективных масс плотности состояний в GaAs, AlAs и $\text{Al}_x\text{Ga}_{1-x}\text{As}^1$

| GaAs | AlAs | $\text{Al}_x\text{Ga}_{1-x}\text{As}$ |
|--------------------------------------------------------------------|----------------------------------------|--------------------------------------------------------------------------------------------|
| E_g^Γ (прямой) = 1,424 эВ при 297 К | E_g^Γ (прямой) = 3,018 эВ | $E_g^\Gamma = (0 < x < 0,45) =$ = 1,424 + 1,247x |
| E_g^L (непрямой) = 1,708 эВ при 297 К | E_g^L (непрямой) = 2,25— —2,35 эВ | $E_g^\Gamma = (0,45 < x < 1,0) =$ = 1,424 + 1,247x + + 1,147 (x - 0,45) ² |
| E_g^X (непрямой) = 1,900 эВ при 297 К | E_g^X (непрямой) = 2,168 эВ | $E_g^L = 1,708 + 0,642x$ $E_g^X = 1,900 + 0,125x +$ + 0,143x ² |
| $m_p = 0,48 m_0$ | $m_p = 0,79 m_0$ | $m_p = (0,48 + 0,31x)m_0$ |
| $m_n^\Gamma = 0,067 m_0$ | $m_n^\Gamma = 0,15 m_0$ | $m_n^\Gamma = (0,067 + 0,083x)m_0$ |
| $m_n^L = 0,55 m_0$ $m_n^X =$ = 0,85 m_0 | $m_n^L = 0,67 m_0$ | $m_n^L = (0,55 + 0,12x) m_0$ |
| $m_0 = 9,11 \cdot 10^{-28}$ г | $m_n^X = 0,78 m_0$ | $m_n^X = (0,85 - 0,07x) m_0$ |
| $E_g(T) = 1,519 -$ $- 5,405 \cdot 10^{-4} T^2 / (204 +$ + T) | | |

¹ Ссылки на литературу см. в тексте.

нимуме зоны проводимости или максимуме валентной зоны, тем больше эффективная масса.

Валентная зона в точке $k = 0$ имеет три ветви. Две верхние ветви называются зонами тяжелых и легких дырок и характеризуются эффективными массами m_{p_h} и m_{p_l} соответственно. Отщепленная валентная зона отделена от них энергией 0,34 эВ [22] и имеет $m_p = 0,15 m_0$ [23], где m_0 — масса свободного электрона; в последующем анализе ее можно не учитывать. Были сделаны многочисленные расчеты величин m_{p_h} и m_{p_l} . Тщательные вычисления Фрэйена [24] дали

$$m_{p_h} = 0,45 m_0, \quad (4.2.4)$$

$$m_{p_l} = 0,082 m_0, \quad (4.2.5)$$

а Мирс и Стрэдлин [25] получили $m_{p_h} = 0,475 m_0$ и $m_{p_l} = 0,087 m_0$. В § 3 этой главы эффективная масса плотности со-

стояний m_p определяется выражением $[(m_{p_h})^{3/2} + (m_{p_l})^{3/2}]^{2/3}$, которое при использовании значений Фрэйена дает $m_p = 0,48 m_0$.

Фрэйен [24] получил следующее значение эффективной массы в прямом минимуме зоны проводимости:

$$m_n^\Gamma = 0,067 m_0. \quad (4.2.6)$$

Для непрямого минимума X зоны проводимости Питт и Лииз [26, 27] получили значение эффективной массы плотности состояний

$$m_n^X = 0,85 m_0, \quad (4.2.7)$$

а Аспнес [18] для эффективной массы плотности состояний в минимуме L зоны проводимости дал значение

$$m_n^L = 0,55 m_0. \quad (4.2.8)$$

Значения энергетических зазоров и эффективных масс в GaAs приведены в первом столбце табл. 4.2.1.

AlAs

В зоне проводимости арсенида алюминия непрямо́й минимум X расположен ниже непрямого минимума L и прямого минимума Γ . Существует значительная неопределенность в значениях E_g^L и E_g^Γ для AlAs. Онтон [28] дал значение $E_g^\Gamma = 2,95$ эВ при комнатной температуре. Монемар [29] получил $E_g^\Gamma = 3,070$ эВ (это значение включает энергию связи экситона 4 мэВ, взяв такую же, как для GaAs) и $E_g^X = 2,163$ эВ (включающее энергию связи экситона 10 мэВ, взяв такую же, как для GaP) при $T = 300$ К. Дингл и др. [30] для арсенида алюминия при комнатной температуре предложили значения $E_g^X = 2,168$ эВ и $E_g^\Gamma = 3,018$ эВ, а также оценили значение E_g^L как лежащее в пределах от 2,25 до 2,35 эВ. Значения энергетических зазоров, полученные Динглом и др. [30], мы будем использовать в дальнейшем.

Лаваецц [23] определил значение эффективной массы дырок в AlAs, равное $m_p = 0,79 m_0$, а Штюкель и Ювима [31] для эффективной массы электрона в прямом минимуме зоны проводимости получили значение $m_n^\Gamma = 0,15 m_0$. Для непрямого минимума X зоны проводимости значение эффективной массы плотности состояний $m_n^X = 0,78 m_0$ было вычислено Роудом [32]. Тем же способом, что и описанный в работе [18] для GaAs, Аспнес [33] получил значение $m_n^L = 0,67 m_0$ для эффективной массы плотности состояний в минимуме L зоны проводимости. Некоторую осторожность надо проявлять при выборе значения

эффективной массы в непрямом минимуме зоны проводимости. Часто приводят продольную m_l и поперечную m_t массы вместе с числом эквивалентных минимумов M . Исходя из этих величин, эффективная масса плотности состояний вычисляется по формуле $m = M^{1/3} (m_l m_t^2)^{1/3}$. Обсуждение вопроса об эффективных массах в непрямых минимумах можно найти в работе Блэйкмора [34], и в дальнейшем этот вопрос рассматриваться не будет. Здесь даны только эффективные массы плотности состояний. Значения энергетических зазоров и эффективных масс в AlAs приведены во втором столбце табл. 4.2.1.

$\text{Al}_x\text{Ga}_{1-x}\text{As}$

Несколько работ было посвящено определению зависимости энергетических зазоров от состава в $\text{Al}_x\text{Ga}_{1-x}\text{As}$. При построении такой зависимости наибольшую ошибку вносило определение состава x . Зависимость E_g^Γ от x при комнатной температуре, полученная на основе данных Кейси и Пейниша [35], Бероло и Вули [36], а также Онтон и др. [37], находится в согласии с результатами Дингла и др. [30] в диапазоне $0 < x < 0,35$, наи-

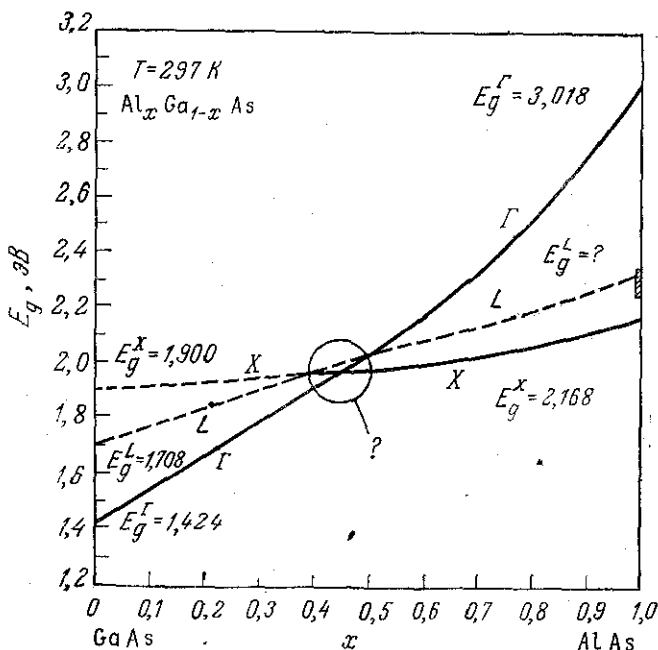


Рис. 4.2.2. Зависимость величин прямого $\Gamma_8 \rightarrow \Gamma_6$ и непрямых $\Gamma_8 \rightarrow X_6$ и $\Gamma_8 \rightarrow L_6$ энергетических зазоров в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ от состава.

более интересном с точки зрения гетеролазеров. Несколько отличные от этих результаты были получены Монемаром и др. [38]. На рис. 4.2.2 представлена зависимость энергетических зазоров от состава в $\text{Al}_x\text{Ga}_{1-x}\text{As}$, построенная по данным Дингла и др. [30]. Составу, при котором происходит переход от прямой запрещенной зоны к непрямой, приписаны значения от $x = 0,37$ [35] до $x = 0,45$ [30]. В настоящее время относительное расположение зон в этой области составов $0,4 \leq x \leq 0,5$ еще недостаточно изучено.

Для вычислений удобно аналитически представить зависимость энергетических зазоров от состава. В области $0 < x < 0,45$ зазор E_g^Γ меньше, чем зазоры E_g^X и E_g^L (рис. 4.2.2), и может быть описан формулой

$$E_g^\Gamma (\text{эВ}) = 1,424 + 1,247x. \quad (4.2.9)$$

При $0,45 < x < 1,0$, когда зазор E_g^X меньше зазоров E_g^Γ и E_g^L , имеем

$$E_g^\Gamma (\text{эВ}) = 1,424 + 1,247x + 1,147(x - 0,45)^2. \quad (4.2.10)$$

Хотя для величины E_g^L в AlAs существует значительная неопределенность, этот энергетический зазор в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ берут в виде линейной функции состава:

$$E_g^L (\text{эВ}) = 1,708 + 0,642x. \quad (4.2.11)$$

И наконец, энергетический зазор E_g^X вычисляют по формуле

$$E_g^X (\text{эВ}) = 1,900 + 0,125x + 0,143x^2. \quad (4.2.12)$$

Выражения (4.2.9)—(4.2.12) написаны для $T = 297$ К.

Линейная интерполяция по значениям эффективных масс плотности состояний в GaAs и AlAs приводит к следующим выражениям для эффективной массы электрона в указанных минимумах зоны проводимости $\text{Al}_x\text{Ga}_{1-x}\text{As}$:

$$m_n^\Gamma = (0,067 + 0,083x) m_0, \quad (4.2.13)$$

$$m_n^L = (0,55 + 0,12x) m_0, \quad (4.2.14)$$

$$m_n^X = (0,85 - 0,07x) m_0. \quad (4.2.15)$$

Дырочная масса плотности состояний в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ рассчитывается по формуле

$$m_p = (0,48 + 0,31x) m_0. \quad (4.2.16)$$

Эти выражения, описывающие композиционную зависимость энергетических зазоров и эффективных масс в $\text{Al}_x\text{Ga}_{1-x}\text{As}$, даны в третьем столбце табл. 4.2.1.

§ 3. ЭНЕРГЕТИЧЕСКИЕ ЗОННЫЕ ДИАГРАММЫ ГЕТЕРОПЕРЕХОДОВ

Разрывы в зоне проводимости и валентной зоне

Энергетические зонные диаграммы гетеропереходов строятся в предположении, что объемные свойства полупроводников, образующих гетеропереход, остаются неизменными вплоть до их границы раздела. На рис. 4.3.1 показан профиль энергетических зон p -GaAs и N -Al_xGa_{1-x}As, взятых в отдельности. Области пространственного заряда у поверхностей пренебрегают, так что профили энергетических зон можно считать не меняющимися с расстоянием. Индексы 1 и 2 относятся к GaAs и Al_xGa_{1-x}As соответственно. E_c и E_v соответствуют зоне проводимости и валентной зоне, F — уровни Ферми, δ — расстояние между краем зоны и уровнем Ферми.

Предполагается, что положение зоны проводимости по отношению к уровню вакуума задается электронным сродством χ . Электронное сродство определяется как энергия, необходимая для переноса электрона со дна зоны проводимости на уровень вакуума. Работа выхода Φ равна энергии, необходимой для переноса электрона с уровня Ферми на уровень вакуума. Более удобно использовать электронное сродство, поскольку работа выхода зависит от положения уровня Ферми, которое изменяется с концентрацией примесей. В настоящее время считают, что

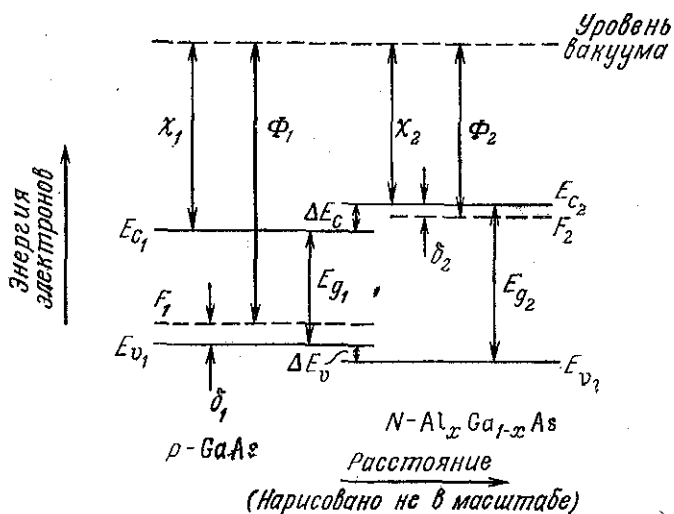


Рис. 4.3.1. Энергетическая зонная диаграмма до образования резкого p - N GaAs-Al_xGa_{1-x}As-гетероперехода. Обозначения поясняются в тексте.

электронное средство от концентрации примесей не зависит. Целесообразность его использования при построении энергетических зонных диаграмм была поставлена под вопрос Кремером [8]. Электронное средство очень трудно определить, и обычно его получают из емкостных измерений, интерпретируя результаты на основе этой модели, или из измерений внешнего фотоэффекта.

На рис. 4.3.1 видно, что разрыв зоны проводимости равен просто разности энергий электронного средства:

$$\Delta E_c = \chi_1 - \chi_2 \quad (4.3.1)$$

и

$$E_{g_1} = \Delta E_c + E_{g_1} + \Delta E_v. \quad (4.3.2)$$

Из уравнения (4.3.2) следует, что разрыв валентной зоны равен

$$\Delta E_v = (E_{g_2} - E_{g_1}) - \Delta E_c \quad (4.3.3)$$

и

$$\Delta E_c + \Delta E_v = E_{g_2} - E_{g_1}. \quad (4.3.4)$$

Хотя уравнение (4.3.2) представляет очевидный результат, оно имеет важное значение, так как ΔE_c и ΔE_v для GaAs—Al_xGa_{1-x}As хорошо известны, в то время как разность между электронными средствами этих материалов с достаточной точностью не установлена.

Из анализа квантовых состояний, возникающих благодаря ограничению носителей в сверхтонких гетероструктурах GaAs—Al_xGa_{1-x}As с $x = 0,20$, было получено [39]

$$\Delta E_c / (E_{g_2} - E_{g_1}) = 0,85 \pm 0,03 \quad (4.3.5)$$

и

$$\Delta E_v / (E_{g_2} - E_{g_1}) = 0,15 \pm 0,03. \quad (4.3.6)$$

Предполагают, что выражения (4.3.5) и (4.3.6) справедливы также и для других составов в области $x \leq 0,45$, где E_{g_2} (прямая) $\leq E_{g_1}$ (непрямая). Никаких данных относительно ΔE_v или ΔE_c не имеется в настоящее время для области $x > 0,45$, в которой E_{g_1} (непрямая) $< E_{g_2}$ (прямая). По вычислениям Френсли и Кремера [9] ΔE_v определяется выражением (4.3.6), в котором E_{g_2} обозначает прямую запрещенную зону E_{g_2} (прямая), рассчитываемую по формуле (4.2.9). Тогда ΔE_c при $x > 0,45$ вычисляется по формуле

$$\Delta E_c = [E_{g_1} (\text{непрямая}) - E_{g_1}] - \Delta E_v, \quad (4.3.7)$$

где E_{g_1} (непрямая) определяется выражением (4.2.12). В любом случае, когда ΔE_c и ΔE_v известны, при построении энергетической зонной диаграммы можно использовать электронные средства.

Определение уровня Ферми

При тепловом равновесии уровень Ферми должен быть постоянным на протяжении всего гетероперехода. Поэтому изгиб зон на границе раздела в гетеропереходе, так же как и в гомопереходе, определяется разностью уровней Ферми с p - и n -сторон. Как показано на рис. 4.3.1, она равна контактной разности потенциалов V_D и определяется формулой

$$V_D(B) = (F_2 - F_1)/q = [E_g + (\Delta E_c - \delta_2) - \delta_1]/q. \quad (4.3.8)$$

Эта формула справедлива также и для других анизотипных и изотипных гетеропереходов. Уровень Ферми определяется из условия электронейтральности:

$$n + \sum_j N_{A_j}^- = p + \sum_k N_{D_k}^+, \quad (4.3.9)$$

где концентрации электронов и дырок в узкозонном n - или p -полупроводнике в гетеропереходе обозначены через n или p , тогда как в более широкозонном N - или P -полупроводнике они обозначаются через N или P . В уравнении (4.3.9) концентрации ионизованных примесей просуммированы по j и k различным типам ионизованных акцепторов $N_{A_j}^-$ и доноров $N_{D_k}^+$ соответственно. Обычно преобладает один тип доноров и акцепторов, и условие электронейтральности записывается в виде (3.7.9)

$$n + N_A^- = p + N_D^+;$$

мы уже встречались с ним в гл. 3.

Другое условие, которое должно быть выполнено, выражается равенством

$$np = n_i^2 / \gamma_n \gamma_p, \quad (4.3.10)$$

где n_i — собственная концентрация носителей, а γ — коэффициенты активности. Хотя в сильно легированном материале коэффициенты γ не равны единице [40, 41], в рассматриваемых здесь случаях мы все же положим их равными единице. Электронная и дырочная концентрации ранее определены выражениями (3.7.7) и (3.7.8):

$$N \text{ или } n = \int_{-\infty}^{\infty} \rho_c(E - E_c) f_c dE,$$

$$P \text{ или } p = \int_{-\infty}^{\infty} \rho_v(E_v - E) [1 - f_v] dE,$$

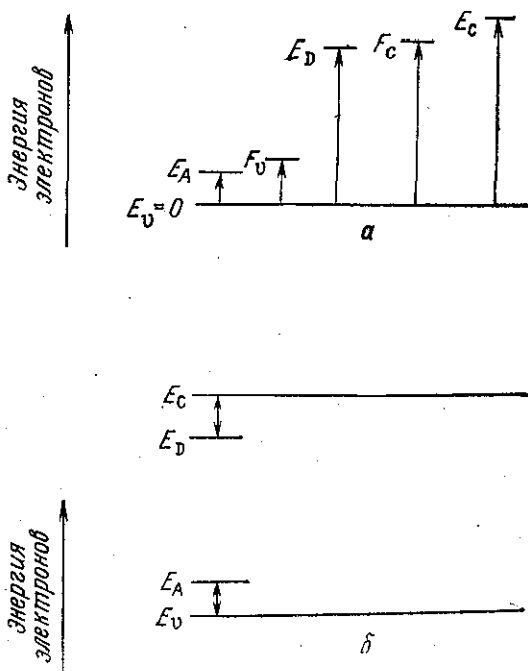


Рис. 4.3.2. *a* — к определению величин E_A , F_v , E_D , F_c и E_C в выражениях (3.2.18)—(3.2.19) и (4.3.12)—(4.3.15). Все эти энергии отмеряются от уровня энергии $E_v = 0$; *б* — к определению энергий ионизации акцепторов (равной $E_A - E_v$) и доноров (равной $E_C - E_D$).

где f_c и f_v определяются выражениями (3.2.19) и (3.2.18):

$$f_c = 1/\{\exp[(E - F_c)/kT] + 1\},$$

$$f_v = 1/\{\exp[(E - F_v)/kT] + 1\}.$$

Для локализованных примесных уровней вероятность $P(E)$ того, что уровень занимает электрон с любым из двух значений спина, равна

$$P(E) = 1/\left\{1 + \frac{1}{2} \exp[(E - F_c)/kT]\right\}. \quad (4.3.11)$$

Коэффициент $1/2$ часто называют фактором спинового вырождения. Если имеется N_D доноров, то число нейтральных доноров (доноров с электронами) определяется формулой

$$N_D^0 = N_D/\left\{1 + \frac{1}{2} \exp[(E_D - F_c)/kT]\right\}, \quad (4.3.12)$$

где E_D — энергия донорного уровня. Тогда число ионизованных доноров N_D^+ есть просто $N_D - N_D^0$ и равно

$$N_D^+ = N_D / \{1 + 2 \exp[(F_c - E_D)/kT]\}. \quad (4.3.13)$$

Аналогично этому для акцепторов имеем

$$N_A^0 = N_A / \left\{ 1 + \frac{1}{2} \exp[(F_v - E_A)/kT] \right\}, \quad (4.3.14)$$

$$N_A^- = N_A / \{1 + 2 \exp[(E_A - F_v)/kT]\}, \quad (4.3.15)$$

где E_A — энергия акцепторного уровня. В выражениях (4.3.12) — (4.3.15) энергии E_A , E_D , F_v и F_c измеряются в такой шкале энергий, в которой положительное направление соответствует возрастанию энергии электронов. Величины E_A , E_D , F_v и F_c принято определять так, как показано на рис. 4.3.2, а. Удобно принять $E_v = 0$, так что эти величины характеризуют положение соответствующих уровней относительно потолка валентной зоны. Значения энергий ионизации доноров и акцепторов в GaAs при низкой концентрации примесей даны в табл. 4.3.1. Стало общепринятым приводить в качестве энергии ионизации акцептора величину $E_A - E_v$, являющуюся также энергией

Таблица 4.3.1

Энергии ионизации примесей в GaAs

| Элемент | Тип примеси | Энергия уровня, эВ | Литература | Элемент | Тип примеси | Энергия уровня, эВ | Литература |
|------------|-------------|----------------------|------------|---------|-------------|--------------------|------------|
| S | донор | 0,00610 | [74] | C | акцептор | 0,026 | [79] |
| Se | донор | 0,00589 | [74] | Si | акцептор | 0,0345 | [79] |
| Te | донор | 0,0058 ¹⁾ | — | Ge | акцептор | 0,0404 | [79] |
| Si | донор | 0,00581 | [74] | Sn | акцептор | 0,171 | [80] |
| Ge | донор | 0,00608 | [74] | Mn | акцептор | 0,113 | [81] |
| Sn | донор | 0,00582 | [75] | Cu | акцептор | 0,15 | [82] |
| C | донор | ~0,006 | [76] | Cu | акцептор | 0,47 | [83] |
| O | донор | 0,75 | [77] | Co | акцептор | 0,16 | [84] |
| неизвестен | донор | 0,17 | [78] | Ni | акцептор | 0,21 | [84] |
| Be | акцептор | 0,028 | [79] | Fe | акцептор | 0,52 | [84] |
| Mg | акцептор | 0,0284 | [79] | Cr | акцептор | 0,79 | [84] |
| Zn | акцептор | 0,0307 | [79] | Au | акцептор | 0,090 | [85] |
| Cd | акцептор | 0,0347 | [79] | Ag | акцептор | 0,238 | [86] |

1) Точное значение энергии уровня, создаваемого Te, неизвестно. Здесь приведено значение, предсказанное Р. Динглом (частное сообщение).

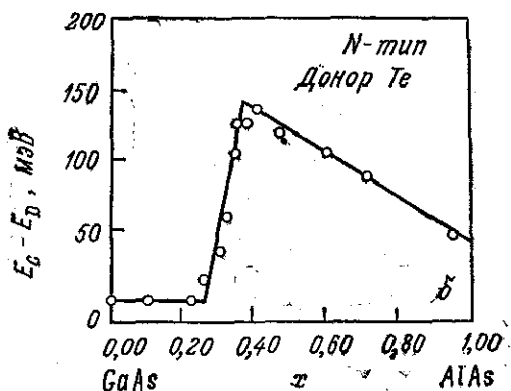
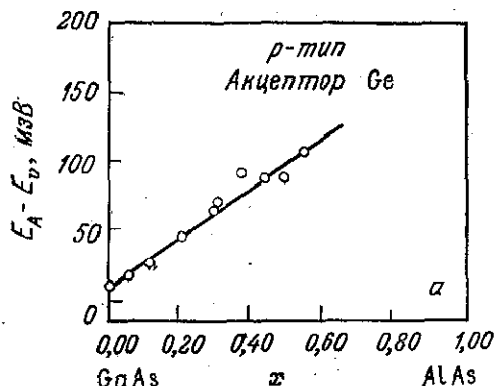


Рис. 4.3.3. Зависимость энергии ионизации примеси от молярной доли x арсенида алюминия [43]. *a* — акцептор Ge; *б* — донор Te.

акцепторного уровня, а в качестве энергии ионизации донора давать величину $E_C - E_D$ (рис. 4.3.2, б). Значения, приведенные в табл. 4.3.1, соответствуют такому определению энергий ионизации примесей. Как уже говорилось в § 5 гл. 3 и иллюстрировалось на рис. 3.5.1, величины $E_A - E_V$ и $E_C - E_D$ зависят от концентрации примесей.

Электрические измерения, проведенные на легированном оловом [42] и германием и теллуром [43] $Al_xGa_{1-x}As$, показали, что $E_A - E_V$ и $E_C - E_D$ возрастают при увеличении x . Результаты Спринг Торпа и др. [43] представлены на рис. 4.3.3. Вследствие зависимости величин $E_A - E_V$ и $E_C - E_D$ от концентрации примесей их значения не стремятся к приведенным в табл. 4.3.1 энергиям ионизации Ge и Te в GaAs при низких концентрациях примесей. Однако влияние молярной доли AlAs

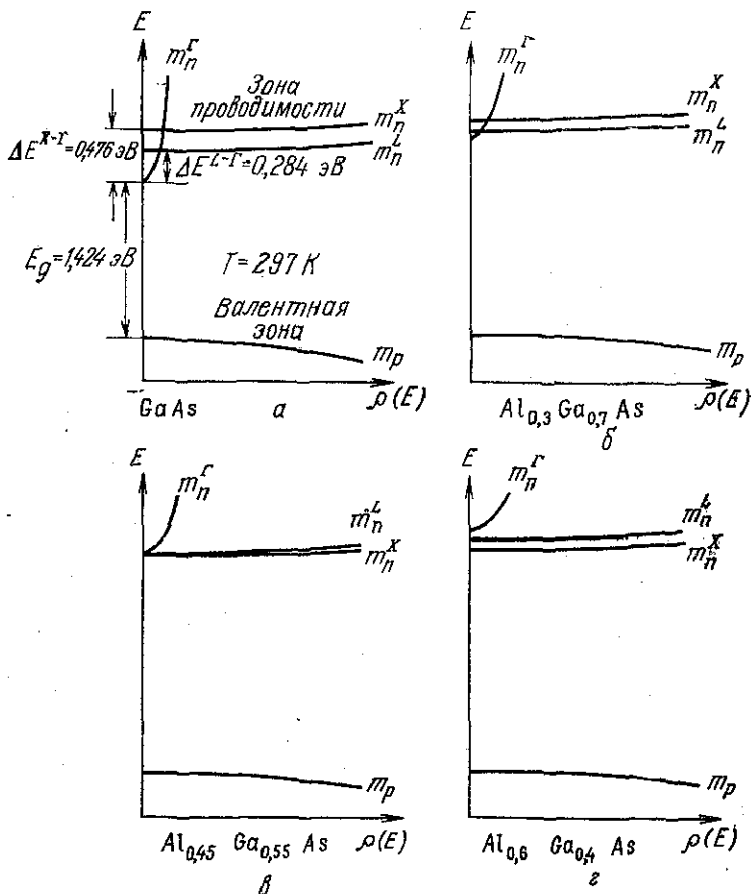


Рис. 4.3.4. Схематическое представление плотности состояний $\rho(E)$ как функции энергии электрона E при комнатной температуре. Прямой минимум зоны проводимости обозначен символом m_n^Γ , а непрямые — символами m_n^L и m_n^X .

а — GaAs; б — $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$; в — $\text{Al}_{0,45}\text{Ga}_{0,55}\text{As}$; г — $\text{Al}_{0,6}\text{Ga}_{0,4}\text{As}$.

на значения $E_A - E_D$ и $E_C - E_D$ отчетливо продемонстрировано на этом рисунке. Энергии ионизации примесей в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ также зависят от концентраций примесей. Определить величину $E_C - E_D$ в $N\text{-Al}_x\text{Ga}_{1-x}\text{As}$ из электрических измерений сложно из-за вклада в коэффициент Холла от не прямых минимумов зоны проводимости [44]. Надлежащий анализ результатов электрических измерений изменит значения $E_C - E_D$, приведенные на рис. 4.3.3, б. Тем не менее зависимость этой величины от x все же останется.

В этом параграфе, посвященном построению энергетических зонных диаграмм гетеропереходов, рассматриваются только параболические плотности состояний, определяемые выражениями (3.4.1) и (3.4.2). Однако для сравнения теоретических и экспериментальных результатов было бы необходимо рассмотреть зависимость величин $E_A - E_v$, $E_c - E_D$, $\rho_c(E - E_c)$ и $\rho_v(E_v - E)$ от концентрации, а также сужение запрещенной зоны, описываемое формулой (3.7.2), при высоких концентрациях носителей. Рассмотрение этих концентрационных зависимостей не включено в последующий анализ, так как основные характерные свойства энергетических зонных диаграмм могут быть проиллюстрированы и без привлечения этих дополнительных усложнений. Параболическая плотность состояний в GaAs представлена на рис. 4.3.4, а, а значения эффективных масс даны в табл. 4.2.1. Плотности состояний в $Al_xGa_{1-x}As$ для $x = 0,3, 0,45$ и $0,6$ представлены на рис. 4.3.4, б—г. Значения эффективных масс и ширина запрещенной зоны вычислялись по приведенным в табл. 4.2.1 формулам. Из рис. 4.3.4 видно, что, во-первых, плотность состояний в зоне проводимости для непрямых минимумов значительно больше, чем для прямого минимума, и, во-вторых, ΔE уменьшается с увеличением x .

Используя выражения (3.4.1), (3.2.19) и (3.7.7), выражение для концентрации электронов n в GaAs при учете вкладов как от прямого Γ -минимума, так и от непрямого L -минимума зоны проводимости можно записать в виде

$$n = \frac{1}{2\pi^2} \left(\frac{2m_n^\Gamma}{\hbar^2} \right)^{3/2} \int_{E_c}^{\infty} \frac{(E - E_c)^{1/2} dE}{1 + \exp[(E - F_c)/kT]} + \frac{1}{2\pi^2} \left(\frac{2m_n^L}{\hbar^2} \right)^{3/2} \int_{E_c + \Delta E}^{\infty} \frac{[E' - (E_c + \Delta E)]^{1/2} dE'}{1 + \exp[(E' - F_c)/kT]}. \quad (4.3.16)$$

В этом выражении прямому Γ -минимуму соответствует первый интеграл, а вклад от непрямого L -минимума, отстоящего от E_c на величину $\Delta E = E_g^L - E_h^\Gamma$, учитывается вторым интегралом. Таким же интегралом может быть учтен вклад и от непрямого X -минимума. Для $E' = E + \Delta E$ второй интеграл в формуле (4.3.16) может быть переписан следующим образом:

$$\int_{E_c + \Delta E}^{\infty} \frac{[E' - (E_c + \Delta E)]^{1/2} dE'}{1 + \exp[(E' - F_c)/kT]} = \int_{E_c}^{\infty} \frac{(E - E_c)^{1/2} dE}{1 + \exp\{[E - (F_c - \Delta E)]/kT\}}. \quad (4.3.17)$$

Следовательно, оба интеграла в формуле (4.3.16) имеют стандартную форму интеграла Ферми — Дирака. Используя безраз-

мерные обозначения Блэйкмора [45]

$$\epsilon = (E - E_c)/kT, \quad (4.3.18)$$

$$\zeta = (F_c - E_c)/kT, \quad (4.3.19)$$

определяющее n выражение (4.3.16) можно переписать в виде

$$n = \frac{1}{2\pi^2} \left(\frac{2m_n^\Gamma kT}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{\epsilon^{1/2} d\epsilon}{1 + \exp(\epsilon - \zeta)} + \\ + \frac{1}{2\pi^2} \left(\frac{2m_n^L kT}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{\epsilon^{1/2} d\epsilon}{1 + \exp[\epsilon - \zeta + (\Delta E/kT)]}. \quad (4.3.20)$$

Интеграл легко вычислить при $(E - F_c)/kT \gg 1$ или $(\epsilon - \zeta) \gg 1$, когда функция Ферми — Дирака может быть представлена экспонентой. В этом случае

$$n = 4\pi \left(\frac{2m_n^\Gamma kT}{\hbar^2} \right)^{3/2} \exp \zeta \int_0^\infty \epsilon^{1/2} \exp(-\epsilon) d\epsilon + \\ + 4\pi \left(\frac{2m_n^L kT}{\hbar^2} \right)^{3/2} \exp \zeta \exp(-\Delta E/kT) \int_0^\infty \epsilon^{1/2} \exp(-\epsilon) d\epsilon, \quad (4.3.21)$$

где интеграл имеет стандартную форму и называется гамма-функцией $\Gamma(3/2) = \pi^{1/2}/2$. Выражение, определяющее концентрацию электронов, теперь принимает вид

$$n = 2 \left[\left(\frac{2\pi m_n^\Gamma kT}{\hbar^2} \right)^{3/2} + \left(\frac{2\pi m_n^L kT}{\hbar^2} \right)^{3/2} \exp(-\Delta E/kT) \right] \exp \left[\frac{F_c - E_c}{kT} \right]. \quad (4.3.22)$$

Величина, служащая в этой формуле предэкспоненциальным множителем, встречается очень часто и называется эффективной плотностью состояний зоны проводимости N_c :

$$N_c = 2 (2\pi kT/\hbar^2)^{3/2} [(m_n^\Gamma)^{3/2} + (m_n^L)^{3/2} \exp(-\Delta E/kT)], \quad (4.3.23)$$

или при учете всех трех минимумов

$$N_c = 2,5 \cdot 10^{19} [(m_n^\Gamma/m_0)^{3/2} + (m_n^L/m_0)^{3/2} \exp(-\Delta E^{L-\Gamma}/kT) + \\ + (m_n^X/m_0)^{3/2} \exp(-\Delta E^{X-\Gamma}/kT)] (T/300)^{3/2}, \quad (4.3.24)$$

где $\Delta E^{L-\Gamma} = E_g^L - E_g^\Gamma$ и $\Delta E^{X-\Gamma} = E_g^X - E_g^\Gamma$. При таком обозначении выражение (4.3.22) можно записать в следующем виде:

$$n = N_c \exp[(F_c - E_c)/kT]. \quad (4.3.25)$$

В GaAs при комнатной температуре имеем $\exp(-\Delta E^{L-\Gamma}/kT) = 1,5 \cdot 10^{-5}$, так что вкладом от непрямых минимумов можно пренебречь:

$$N_c = 2,5 \cdot 10^{19} (m_n^{\Gamma}/m_0)^{3/2} (T/300)^{3/2}.$$

В $\text{Al}_x\text{Ga}_{1-x}\text{As}$ при возрастании x от нуля величина ΔE уменьшается, как показано на рис. 4.2.2 и рис. 4.3.3, поэтому необходимо учитывать распределение электронов между прямым и непрямым минимумами зоны проводимости. Выражение (4.3.24) определяет N_c для области $x \leq 0,45$, где $\text{Al}_x\text{Ga}_{1-x}\text{As}$ имеет прямую запрещенную зону. В области непрямозонных составов N_c равно

$$N_c (\text{непрямозон.}) = 2,5 \cdot 10^{19} [(m_n^{\Gamma}/m_0)^{3/2} \exp(-\Delta E^{\Gamma-X}/kT) + (m_n^L/m_0)^{3/2} \exp(-\Delta E^{L-X}/kT) + (m_n^X/m_0)^{3/2}] (T/300)^{3/2}, \quad (4.3.26)$$

где $\Delta E^{\Gamma-X} = E_g^{\Gamma} - E_g^X$ и $\Delta E^{L-X} = E_g^L - E_g^X$ при $x > 0,45$.

Когда экспоненциальное приближение для функции Ферми — Дирака неприменимо, интеграл можно записать следующим образом:

$$n = N_c \mathcal{F}_{1/2} \left(\frac{F_c - E_c}{kT} \right). \quad (4.3.27)$$

Здесь $\mathcal{F}_{1/2}(\zeta)$ представляет собой интеграл Ферми — Дирака

$$\mathcal{F}_{1/2}(\zeta) = \frac{2}{\pi^{1/2}} \int_0^{\infty} \frac{\epsilon^{1/2} d\epsilon}{1 + \exp(\epsilon - \zeta)}, \quad (4.3.28)$$

значения которого приведены в табл. 4.3.2 [46].

Концентрация дырок p вычисляется таким же способом. Как было отмечено в предыдущем параграфе, валентная зона содержит зоны тяжелых и легких дырок, характеризующиеся эффективными массами m_{p_h} и m_{p_l} соответственно. Концентрация дырок может быть представлена в виде суммы двух членов:

$$p = \frac{1}{2\pi^2} \left(\frac{2m_{p_h}}{\hbar^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{1 + \exp[(F_v - E)/kT]} + \frac{1}{2\pi^2} \left(\frac{2m_{p_l}}{\hbar^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{1 + \exp[(F_v - E)/kT]}. \quad (4.3.29)$$

Для того чтобы упростить это выражение, полезно вновь использовать безразмерные переменные (4.3.18) и (4.3.19) вместе с

$$\epsilon_p = (E_v - E)/kT, \quad (4.3.30)$$

Значения интеграла Ферми—Дирака $\mathcal{F}_{1/2}(\zeta)$

| ζ | $\mathcal{F}_{1/2}$ | ζ | $\mathcal{F}_{1/2}$ | ζ | $\mathcal{F}_{1/2}$ |
|---------|---------------------|---------|---------------------|---------|---------------------|
| -4,0 | 1,8199 (-2) | -0,3 | 6,0022 (-1) | 3,4 | 5,2571 (0) |
| -3,9 | 2,0099 (-2) | -0,2 | 6,5161 (-1) | 3,5 | 5,4580 (0) |
| -3,8 | 2,2195 (-2) | -0,1 | 7,0654 (-1) | 3,6 | 5,6623 (0) |
| -3,7 | 2,4510 (-2) | 0,0 | 7,6515 (-1) | 3,7 | 5,8699 (0) |
| -3,6 | 2,7063 (-2) | 0,1 | 8,2756 (-1) | 3,8 | 6,0806 (0) |
| -3,5 | 2,9880 (-2) | 0,2 | 8,9388 (-1) | 3,9 | 6,29451 (0) |
| -3,4 | 3,2986 (-2) | 0,3 | 9,6422 (-1) | 4,0 | 6,5115 (0) |
| -3,3 | 3,6412 (-2) | 0,4 | 1,0387 (0) | 4,2 | 6,9548 (0) |
| -3,2 | 4,0187 (-2) | 0,5 | 1,1173 (0) | 4,4 | 7,4100 (0) |
| -3,1 | 4,4349 (-2) | 0,6 | 1,2003 (0) | 4,6 | 7,8769 (0) |
| -3,0 | 4,8933 (-2) | 0,7 | 1,2875 (0) | 4,8 | 8,3550 (0) |
| -2,9 | 5,3984 (-2) | 0,8 | 1,3791 (0) | 5,0 | 8,8442 (0) |
| -2,8 | 5,9545 (-2) | 0,9 | 1,4752 (0) | 5,2 | 9,3441 (0) |
| -2,7 | 6,5665 (-2) | 1,0 | 1,5756 (0) | 5,4 | 9,8546 (0) |
| -2,6 | 7,2398 (-2) | 1,1 | 1,6806 (0) | 5,6 | 1,0375 (+1) |
| -2,5 | 7,9804 (-2) | 1,2 | 1,7900 (0) | 5,8 | 1,0906 (+1) |
| -2,4 | 8,7944 (-2) | 1,3 | 1,9038 (0) | 6,0 | 1,1447 (+1) |
| -2,3 | 9,6887 (-2) | 1,4 | 2,0221 (0) | 6,2 | 1,1997 (+1) |
| -2,2 | 1,0671 (-1) | 1,5 | 2,1449 (0) | 6,4 | 1,2556 (+1) |
| -2,1 | 1,1748 (-1) | 1,6 | 2,2720 (0) | 6,6 | 1,3125 (+1) |
| -2,0 | 1,2930 (-1) | 1,7 | 2,4035 (0) | 6,8 | 1,3703 (+1) |
| -1,9 | 1,4225 (-1) | 1,8 | 2,5393 (0) | 7,0 | 1,4290 (+1) |
| -1,8 | 1,5642 (-1) | 1,9 | 2,6794 (0) | 7,2 | 1,4886 (+1) |
| -1,7 | 1,7193 (-1) | 2,0 | 2,8237 (0) | 7,4 | 1,5491 (+1) |
| -1,6 | 1,8889 (-1) | 2,1 | 2,9722 (0) | 7,6 | 1,6104 (+1) |
| -1,5 | 2,0740 (-1) | 2,2 | 3,1249 (0) | 7,8 | 1,6725 (+1) |
| -1,4 | 2,2759 (-1) | 2,3 | 3,2816 (0) | 8,0 | 1,7355 (+1) |
| -1,3 | 2,4959 (-1) | 2,4 | 3,4423 (0) | 8,2 | 1,7993 (+1) |
| -1,2 | 2,7353 (-1) | 2,5 | 3,6070 (0) | 8,4 | 1,8639 (+1) |
| -1,1 | 2,9955 (-1) | 2,6 | 3,7755 (0) | 8,6 | 1,9293 (+1) |
| -1,0 | 3,2780 (-1) | 2,7 | 3,9480 (0) | 8,8 | 1,9954 (+1) |
| -0,9 | 3,5841 (-1) | 2,8 | 4,1241 (0) | 9,0 | 2,0624 (+1) |
| -0,8 | 3,9154 (-1) | 2,9 | 4,3040 (0) | 9,2 | 2,1301 (+1) |
| -0,7 | 4,2733 (-1) | 3,0 | 4,4876 (0) | 9,4 | 2,1986 (+1) |
| -0,6 | 4,6595 (-1) | 3,1 | 4,6747 (0) | 9,6 | 2,2678 (+1) |
| -0,5 | 5,0754 (-1) | 3,2 | 4,8653 (0) | 9,8 | 2,3378 (+1) |
| -0,4 | 5,5224 (-1) | 3,3 | 5,0595 (0) | 10,0 | 2,4085 (+1) |

так что

$$(F_v - E)/kT = \epsilon_p + \epsilon_g + \zeta. \quad (4.3.31)$$

В этой формуле ϵ_g является безразмерной шириной запрещенной зоны

$$\epsilon_g = (E_c - E_v)/kT. \quad (4.3.32)$$

Выражение (4.3.29) при таких обозначениях принимает вид

$$p = \left[4\pi \left(\frac{2m_{p_h} kT}{h^2} \right)^{3/2} + 4\pi \left(\frac{2m_{p_l} kT}{h^2} \right)^{3/2} \right] \int_0^\infty \frac{\epsilon_p^{1/2} d\epsilon_p}{1 + \exp(\epsilon_p + \epsilon_g + \zeta)}. \quad (4.3.33)$$

При экспоненциальном приближении для функции распределения Ферми — Дирака из этой формулы получаем

$$p = \left\{ 4\pi \left(\frac{2kT}{h^2} \right)^{3/2} [(m_{p_h})^{3/2} + (m_{p_l})^{3/2}] \right\} \times \\ \times \exp[-(\epsilon_g + \zeta)] \int_0^\infty \epsilon_p^{1/2} \exp(-\epsilon_p) d\epsilon_p, \quad (4.3.34)$$

или

$$p = 2(2\pi kT/h^2)^{3/2} [(m_{p_h})^{3/2} + (m_{p_l})^{3/2}] \exp[(E_v - F_v)/kT]. \quad (4.3.35)$$

Предэкспоненциальный множитель в выражении (4.3.35) называется эффективной плотностью состояний в валентной зоне и обозначается через N_v :

$$N_v = 2(2\pi kT/h^2)^{3/2} [(m_{p_h})^{3/2} + (m_{p_l})^{3/2}]. \quad (4.3.36)$$

Она обычно выражается через массу m_p эффективной плотности состояний в валентной зоне:

$$N_v = 2(2\pi m_p kT/h^2)^{3/2}. \quad (4.3.37)$$

Из выражений (4.3.36) и (4.3.37) следует

$$m_p = [(m_{p_h})^{3/2} + (m_{p_l})^{3/2}]^{2/3} \quad (4.3.38)$$

и

$$N_v = 2,5 \cdot 10^{19} (m_p/m_0)^{3/2} (T/300)^{3/2} \text{ см}^{-3}. \quad (4.3.39)$$

Выражение (4.3.35) может быть теперь записано в виде

$$p = N_v \exp[(E_v - F_v)/kT]. \quad (4.3.40)$$

Когда экспоненциальное приближение для функции Ферми — Дирака неприменимо, имеем

$$p = N_v \int_0^\infty \frac{\epsilon_p^{1/2} d\epsilon_p}{1 + \exp(\epsilon_p + \epsilon_g + \zeta)} \quad (4.3.41)$$

и, полагая $E_v = 0$, получаем выражение

$$p = N_v \mathcal{F}_{1/2}(-F_v/kT), \quad (4.3.42)$$

имеющее тот же вид, что и выражение (4.3.27) для электронов.

Произведение концентраций электронов и дырок, которое, согласно формуле (4.3.10), должно равняться n_i^2 (при низком уровне легирования γ_n и γ_p равны единице), определяется произведением выражений (4.3.25) и (4.3.40):

$$np = n_i^2 = N_c N_v \exp(-E_g/kT), \quad (4.3.43)$$

где n_i — собственная концентрация носителей. Тщательное изучение зависимости n_i от температуры было проведено Тэрмондом [17].

Простейший случай для нахождения F_c или F_v реализуется, когда n/N_c и p/N_v меньше чем примерно 0,1, так что может быть использовано экспоненциальное приближение. Если все доноры ионизованы, то условие электронейтральности дает $n = N_D$, или, когда величиной N_A нельзя пренебречь, $n = N_D - N_A$. Тогда из выражения (4.3.25) получаем

$$(F_c = E_c + kT \ln(n/N_c)) \quad (4.3.44)$$

При полной ионизации акцепторов в материале p -типа условие электронейтральности дает $p = N_A$ или $p = N_A - N_D$, и уровень Ферми определяется аналогичным выражением, получаемым из формулы (4.3.40).

Когда концентрация основных носителей больше чем $0,1 N_c$ для материала n -типа или $0,1 N_v$ для материала p -типа (так называемый случай вырождения), необходимо использовать выражения (4.3.27) или (4.3.42), определяющие n или p через интеграл Ферми — Дирака. Для нахождения значений $\mathcal{F}_{1/2}[(F_c - E_c)/kT]$ или $\mathcal{F}_{1/2}(F_v/kT)$ надо использовать табл. 4.3.2. Джойс и Диксон [47] распространили больцмановское приближение для уровня Ферми, с помощью которого была получена формула (4.3.44), на вырожденный случай, что дало

$$F_c = E_c + kT [\ln(n/N_c) + 3,53553 \cdot 10^{-1} (n/N_c) - 4,95009 \cdot 10^{-3} (n/N_c)^2 + 1,48386 \cdot 10^{-4} (n/N_c)^3 - 4,42563 \cdot 10^{-6} (n/N_c)^4]. \quad (4.3.45)$$

Для интеграла Ферми — Дирака бралось следующее приближенное выражение [47]:

$$\begin{aligned} \mathcal{F}_{1/2}[(F_c - E_c)/kT] = \\ = \exp[(F_c - E_c)/kT] - 2^{-3/2} \exp[2(F_c - E_c)/kT] + \dots \\ \dots + (-1)^{m+1} m^{-3/2} \exp[m(F_c - E_c)/kT]. \end{aligned} \quad (4.3.46)$$

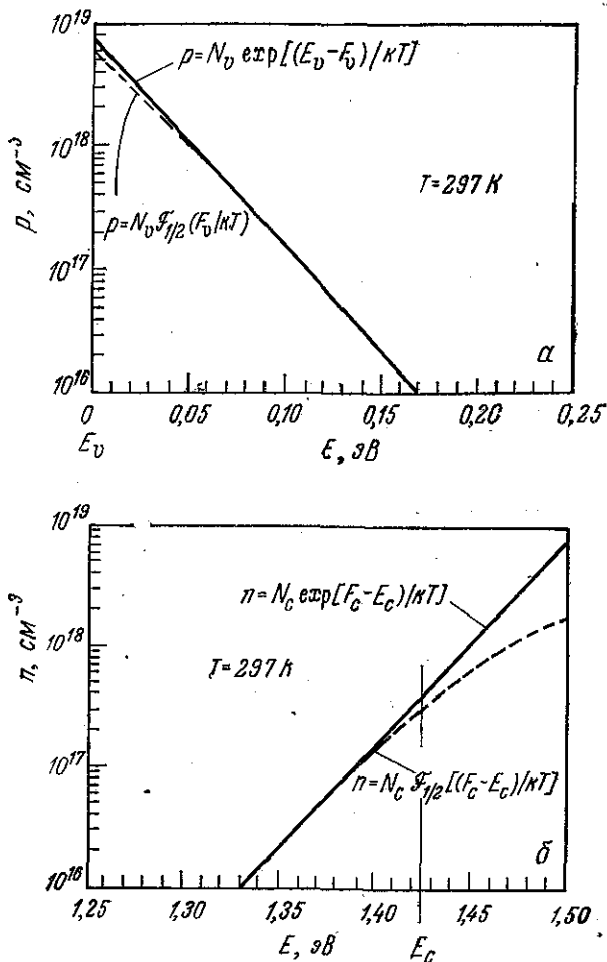


Рис. 4.3.5. Сравнение концентраций носителей в GaAs при $T = 297 \text{ K}$, полученных при экспоненциальном приближении функции распределения и непосредственно при использовании интеграла Ферми — Дирака. а — дырки; б — электроны.

Обычно для оценки выражений (4.3.45) и (4.3.46) необходимо учитывать только несколько первых членов. Аналогичные этим выражения могут быть написаны и для дырок в валентной зоне. Если вырождение возникает, когда значительная часть носителей заселяет более чем один минимум зоны проводимости, например при большом N для $x \approx 0,45$, коэффициенты в выражении (4.3.45) имеют другие значения [47].

На рис. 4.3.5 на примере GaAs при $T = 297$ К продемонстрировано различие между значениями энергии уровней Ферми, рассчитанными с помощью экспоненциального приближения и интеграла Ферми — Дирака. Для значений эффективных масс плотности состояний, приведенных в табл. 4.2.1, выражения (4.3.24) и (4.3.39) дают $N_c = 4,27 \cdot 10^{17}$ см⁻³ и $N_v = 8,19 \cdot 10^{18}$ см⁻³. Из рис. 4.3.5, а видно, что для p -GaAs вплоть до значений концентрации дырок $p \approx 1 \cdot 10^{18}$ см⁻³ экспоненциальное приближение и интеграл Ферми — Дирака дают для уровня Ферми один и тот же результат. В случае n -GaAs эти методы расчета F_c дают разные значения при $n > 1 \cdot 10^{17}$ см⁻³.

Контактная разность потенциалов

Для того чтобы перейти к построению энергетической зонной диаграммы, необходимо рассчитать пространственную зависимость контактной разности потенциалов с каждой стороны гетерограницы. Если два полупроводника приведены в контакт друг с другом, то уровень Ферми при тепловом равновесии должен быть постоянным по всему образовавшемуся твердому телу:

$$\frac{dF}{dx} = 0. \quad (4.3.47)$$

Это требование вытекает из второго закона термодинамики [48]. Энергия уровня Ферми является химическим потенциалом или свободной энергией электронов и дырок [49]. После приведения полупроводников в контакт рис. 4.3.1 надо изменить так (рис. 4.3.6, а), чтобы выполнялось условие (4.3.47).

На рис. 4.3.6, б показано распределение концентрации свободных носителей в момент образования гетероперехода, зонная диаграмма которого изображена на рис. 4.3.6, а. Концентрация дырок с p -стороны гетерограницы, где они являются основными носителями, значительно больше, чем с N -стороны, где основными носителями являются электроны. Такой большой градиент концентрации приводит к возникновению диффузионного дырочного тока. В случае когда градиент есть только вдоль одной из осей, плотность диффузионного тока дырок равна

$$i_p = -qD_p dp/dx. \quad (4.3.48)$$

Здесь D_p — коэффициент диффузии дырок. Он связан с подвижностью соотношением Эйнштейна

$$D_p = \mu_p kT/q, \quad (4.3.49)$$

где μ_p — подвижность дырок. Для диффузионного тока электронов имеем

$$i_n = qD_n dn/dx. \quad (4.3.50)$$

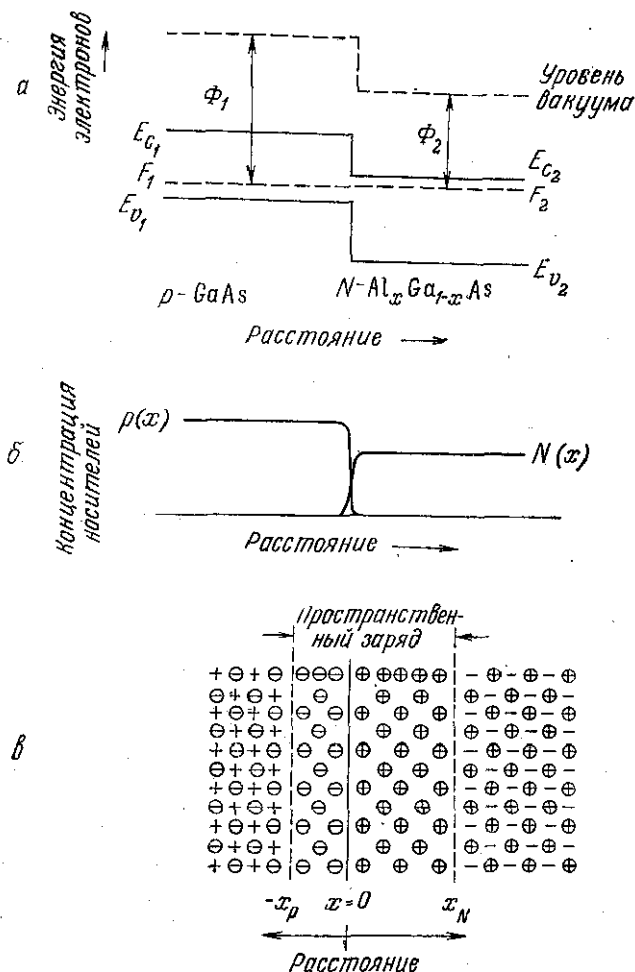


Рис. 4.3.6. Резкий GaAs—Al_xGa_{1-x}As ρ — N -гетеропереход при термическом равновесии. а — энергетическая зонная диаграмма в момент образования гетероперехода; б — распределение носителей в момент образования гетероперехода; в — обедненная подвижными носителями область пространственного заряда после установления поля контактной разности потенциалов. Неподвижные ионизованные доноры обозначены через \oplus , неподвижные акцепторы — через \ominus , подвижные электроны и дырки — через $-$ и $+$ соответственно.

Здесь D_n — коэффициент диффузии электронов. Дырки с ρ -стороны и электроны с N -стороны диффундируют через переход и рекомбинируют, оставляя позади себя области, содержащие неподвижные заряженные акцепторы и доноры, как показано на рис. 4.3.6, в. Эти обедненные носителями области вблизи ге-

терограницы образуют область пространственного заряда. Пространственный заряд создает поле контактной разности потенциалов, препятствующее диффузии носителей. Область около гетерограницы обедняется основными носителями до тех пор, пока дрейфовый ток, обусловленный полем \mathcal{E} контактной разности потенциалов, не становится достаточно большим, чтобы в точности компенсировать диффузионный ток. Плотности дрейфового тока дырок и электронов равны

$$i_p = q\mu_p p \mathcal{E} \quad (4.3.51)$$

и

$$i_n = q\mu_n n \mathcal{E} \quad (4.3.52)$$

соответственно.

При тепловом равновесии полные токи дырок и электронов должны равняться нулю. Плотность полного электронного тока

$$i_n = 0 = q\mu_n n [(D_n/n\mu_n)(dn/dx) + \mathcal{E}], \quad (4.3.53)$$

откуда

$$\mathcal{E} = -(D_n/n\mu_n)(dn/dx) = -(kT/q)(1/n)(dn/dx). \quad (4.3.54)$$

Аналогичное выражение получается для дырок. Если концентрация электронов может быть представлена в экспоненциальной форме (4.3.25), то выражение (4.3.54) принимает вид

$$\mathcal{E} = (1/q)(dE_c/dx). \quad (4.3.55)$$

При получении этой формулы использовано то обстоятельство, что по условию (4.3.47) $dF_c/dx = 0$. Формула (4.3.55) показывает, что наклон края зоны проводимости относительно оси x равен $q\mathcal{E}$. Отсутствие наклона означает, что $\mathcal{E} = 0$, и поскольку в области пространственного заряда $|\mathcal{E}| > 0$, энергетическая зонная диаграмма на рис. 4.3.6, а должна быть изменена с учетом поля контактной разности потенциалов.

Детальная картина поля контактной разности потенциалов наиболее просто может быть получена из уравнения Пуассона. Уравнение (2.2.6) связывает плотность пространственного заряда ρ с плотностью электрического потока $\mathcal{D} = \epsilon\mathcal{E}$, и в одномерном случае с учетом выражения (2.2.3) для \mathcal{D} оно принимает вид

$$d\mathcal{E}/dx = \rho/\epsilon. \quad (4.3.56)$$

В статическом случае потенциал V связан с полем соотношением [см. выражение (3.3.15)]

$$\mathcal{E} = -dV/dx. \quad (4.3.57)$$

Из этого соотношения и уравнения (4.3.56) вытекает уравнение

$$d^2V/dx^2 = -\rho/\epsilon, \quad (4.3.58)$$

которое и является уравнением Пуассона. Плотность заряда определяется суммой зарядов всех типов:

$$\rho = q(p - n + N_D^+ - N_A^-). \quad (4.3.59)$$

Решение уравнения (4.3.58) обычно получают в предположении, что концентрациями подвижных носителей в области пространственного заряда можно пренебречь. Потенциал с p -стороны обозначим через V_p , а с N -стороны — через V_N . Как показано на рис. 4.3.6, в, координаты $-x_p$ и x_N соответствуют границам области пространственного заряда. Уравнение Пуассона теперь запишется в виде

$$d^2V_{p,N}/dx^2 = 0 \quad \text{при } -\infty < x < -x_p, \quad x_N < x < \infty, \quad (4.3.60)$$

$$d^2V_p/dx^2 = (q/\epsilon_1)(N_{A_1}^- - N_{D_1}^+) \quad \text{при } -x_p \leq x < 0, \quad (4.3.61)$$

$$d^2V_N/dx^2 = (-q/\epsilon_2)(N_{D_2}^+ - N_{A_2}^-) \quad \text{при } 0 < x \leq x_N. \quad (4.3.62)$$

Для GaAs $\epsilon_{\text{GaAs}} = \epsilon_1 = 13,1 \epsilon_0$ [50], а для AlAs $\epsilon = 10,06 \epsilon_0$ [51]. Используя линейную интерполяцию, получаем ϵ_2 для $\text{Al}_x\text{Ga}_{1-x}\text{As}$:

$$\epsilon_{\text{Al}_x\text{Ga}_{1-x}\text{As}} = \epsilon_2 = (13,1 - 3,0x) \epsilon_0. \quad (4.3.63)$$

Величина ϵ_0 равна $8,85 \cdot 10^{-14}$ Ф/см.

Для p -стороны гетероперехода из уравнения (4.3.61) с помощью соотношения (4.3.57) получаем

$$d\mathcal{E}/dx = -(q/\epsilon_1)(N_{A_1}^- - N_{D_1}^+) \quad \text{при } -x_p \leq x < 0. \quad (4.3.64)$$

Аналогично для N -стороны

$$d\mathcal{E}/dx = (q/\epsilon_2)(N_{D_2}^+ - N_{A_2}^-) \quad \text{при } 0 < x \leq x_N. \quad (4.3.65)$$

Интегрирование уравнений (4.3.64) и (4.3.65) дает

$$\mathcal{E} = -dV_p/dx = -(q/\epsilon_1)(N_{A_1}^- - N_{D_1}^+)x + A_1 \quad \text{при } -x_p \leq x < 0 \quad (4.3.66)$$

и

$$\mathcal{E} = -dV_N/dx = (q/\epsilon_2)(N_{D_2}^+ - N_{A_2}^-)x + A_2 \quad \text{при } 0 < x \leq x_N, \quad (4.3.67)$$

где A_1 и A_2 — постоянные интегрирования, которые должны быть определены из граничных условий. На границах области пространственного заряда поле должно обращаться в нуль. Из этого условия получаем

$$A_1 = -(q/\epsilon_1)(N_{A_1}^- - N_{D_1}^+)x_p, \quad (4.3.68)$$

$$A_2 = -(q/\epsilon_2)(N_{D_2}^+ - N_{A_2}^-)x_N, \quad (4.3.69)$$

и уравнения (4.3.66) и (4.3.67) принимают вид

$$\mathcal{E} = -dV_p/dx = -(q/\varepsilon_1)(N_{A_1}^- - N_{D_1}^+)(x + x_p) \quad \text{при } -x_p \leq x < 0, \quad (4.3.70)$$

$$\mathcal{E} = -dV_N/dx = -(q/\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)(x_N - x) \quad \text{при } 0 < x \leq x_N. \quad (4.3.71)$$

Интегрирование уравнений (4.3.70) и (4.3.71) дает

$$V_p = (q/\varepsilon_1)(N_{A_1}^- - N_{D_1}^+) [(x^2/2) + x_p x] + B_1 \quad \text{при } -x_p \leq x < 0, \quad (4.3.72)$$

$$V_N = (q/\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)[x_N x - (x^2/2)] + B_2 \quad \text{при } 0 < x \leq x_N. \quad (4.3.73)$$

Положим $V_p = 0$ при $x = -x_p$, тогда

$$B_1 = (q/\varepsilon_1)(N_{A_1}^- - N_{D_1}^+)(x_p^2/2), \quad (4.3.74)$$

$$V_p = (q/2\varepsilon_1)(N_{A_1}^- - N_{D_1}^+)(x_p + x)^2 \quad \text{при } -x_p \leq x < 0. \quad (4.3.75)$$

Контактная разность потенциалов V_{D_p} , приходящаяся на p -сторону гетероперехода, получается теперь из формулы (4.3.75) при $x = 0$:

$$V_{D_p} = (q/2\varepsilon_1)(N_{A_1}^- - N_{D_1}^+)x_p^2. \quad (4.3.76)$$

С N -стороны при $x = x_N$ V_N представляет собой полную контактную разность потенциалов V_D , поэтому

$$B_2 = V_D - (q/\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)(x_N^2/2), \quad (4.3.77)$$

$$V_N = V_D - (q/2\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)(x_N - x)^2 \quad \text{при } 0 < x \leq x_N. \quad (4.3.78)$$

При $x = 0$ имеем $V_N = V_{D_p}$:

$$V_N = V_{D_p} = V_D - (q/2\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)x_N^2. \quad (4.3.79)$$

С другой стороны,

$$V_D = V_{D_p} + V_{D_N}, \quad (4.3.80)$$

где V_{D_N} — контактная разность потенциалов, приходящаяся на N -сторону гетероперехода. Поэтому

$$V_{D_N} = V_D - V_{D_p} = (q/2\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)x_N^2. \quad (4.3.81)$$

Из уравнений (4.3.78), (4.3.80) и (4.3.81) следует

$$V_N = V_{D_p} + (q/2\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)x_N^2 - (q/2\varepsilon_2)(N_{D_2}^+ - N_{A_2}^-)(x_N - x)^2, \quad (4.3.82)$$

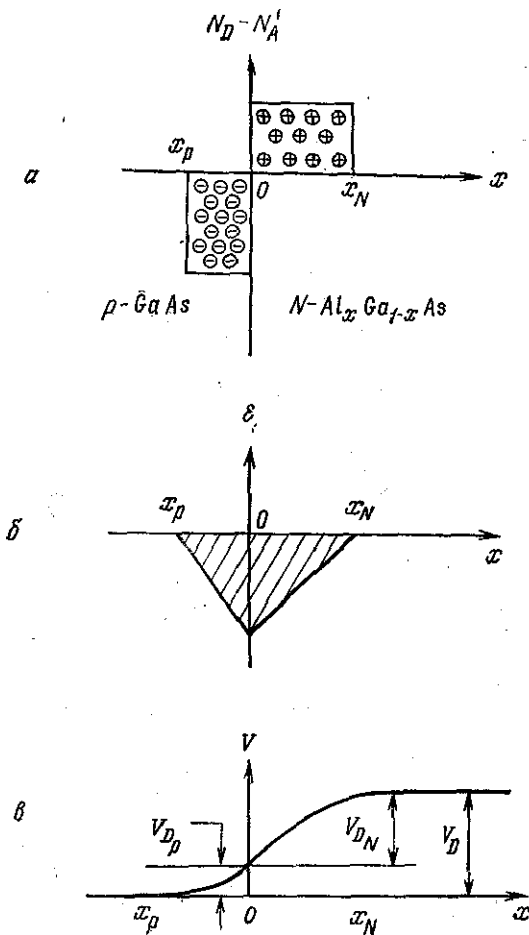


Рис. 4.3.7. Резкий GaAs — Al_xGa_{1-x}As *p* — *n*-гетеропереход при тепловом равновесии. Краю области пространственного заряда на *p*-стороне соответствует координата $-x_p$, а на *n*-стороне — x_n . *a* — распределение примесей, поясняющее образование области пространственного заряда; *б* — распределение электрического поля; *в* — зависимость потенциала от расстояния.

или в более удобной форме

$$V_N = V_{Dp} + (q/2e_2)(N_{D_2}^+ - N_{A_2}^-)[x_N^2 - (x_N - x)^2] \quad \text{при } 0 < x \leq x_N. \quad (4.3.83)$$

На рис. 4.3.7, *a*, *б* и *в* схематически представлены пространственный заряд, электрическое поле и контактная разность потенциалов соответственно. Из выражений (4.3.70) и (4.3.71) видно, что \mathcal{E} отрицательно и линейно меняется с расстоянием,

Таблица 4.3.3

Выражения, позволяющие оценить контактную разность потенциалов V_D

$$V_D (B) = \{E_{g_1} + [\Delta E_c - (E_{c_2} - F_2)] - (E_1 - E_{v_1})\} / q$$

$$E_{c_2} - F_2 = -kT \ln (n/N_c) \text{ при } N/N_c \text{ или } n/N_c < 0,1$$

$$F_1 - E_{v_1} = -kT \ln (p/N_v) \text{ при } P/N_v \text{ или } p/N_v < 0,1$$

$$\Delta E_c = 0,85 (E_{g_2} - E_{g_1}) \text{ при } Al_xGa_{1-x}As$$

как это изображено на рис. 4.3.7, б. Потенциал определяется выражениями (4.3.75) и (4.3.83). Он положителен и меняется с расстоянием квадратично (рис. 4.3.7, в). Проведенный выше анализ области пространственного заряда для $p-N$ -гетероперехода такой же, как и для гомоперехода [52, 53]. Единственное отличие заключается в том, что в энергетическую зонную диаграмму гетероперехода введены разрывы зоны проводимости и валентной зоны.

Из формул (4.3.55) и (4.3.57) следует

$$\mathcal{E} = (1/q) (dE_c/dx) = - (dV/dx), \quad (4.3.84)$$

или

$$dE_c = -q dV. \quad (4.3.85)$$

Эти уравнения показывают, что края зон в области пространственного заряда изгибаются в направлении, противоположном изгибу описывающей потенциал кривой. Изгиб зон в области пространственного заряда теперь можно учесть на рис. 4.3.6, а, для того чтобы получить энергетическую зонную диаграмму $p\text{-GaAs} - N\text{-Al}_x\text{Ga}_{1-x}\text{As}$ -гетероперехода. Детальная схема численного расчета энергетической зонной диаграммы приведена в следующем разделе настоящего параграфа. Сводка выражений, необходимых для расчета V_D , дана в табл. 4.3.3.

GaAs — Al_xGa_{1-x}As $p-N$ -гетеропереходы

В этом разделе на основе полученных выше формул выводятся выражения, необходимые для построения энергетической зонной диаграммы при определенных концентрациях примесей и молярной доле AlAs. Применяемый здесь метод описания энергетической зонной диаграммы гетероперехода, аналогичный тому, который обычно используется для гомоперехода, часто называется моделью Андерсона [6, 7]. Эта модель отличается от случая гомоперехода только наличием ΔE_c и ΔE_v на гетерогранице и тем, что полупроводники с разных сторон перехода имеют разные значения ширины запрещенной зоны и диэлектри-

ческой проницаемости. Для демонстрации влияния концентрации примесей на энергетическую зонную диаграмму приводятся численные примеры.

Если поверхностного заряда на гетерогранице нет, то плотность электрического потока $\mathcal{D} = e\mathcal{E}$ должна быть непрерывной. Из уравнений (4.3.70) и (4.3.71) для $\varepsilon_1 \mathcal{E}_1 = \varepsilon_2 \mathcal{E}_2$ при $x=0$ получаем

$$(N_{A_1}^- - N_{D_1}^+) x_p = (N_{D_2}^+ - N_{A_2}^-) x_N, \quad (4.3.86)$$

а деление контактной разности потенциалов V_{DN} [выражение (4.3.81)] на контактную разность потенциалов V_{Dp} [выражение (4.3.76)] дает

$$V_{DN}/V_{Dp} = [\varepsilon_1 (N_{D_2}^+ - N_{A_2}^-) x_N^2] / [\varepsilon_2 (N_{A_1}^- - N_{D_1}^+) x_p^2]. \quad (4.3.87)$$

Из этих двух формул следует

$$V_{DN}/V_{Dp} = [(N_{A_1}^- - N_{D_1}^+) \varepsilon_1] / [(N_{D_2}^+ - N_{A_2}^-) \varepsilon_2]. \quad (4.3.88)$$

Полная контактная разность потенциалов V_D определяется выражением (4.3.8):

$$V_D(B) = \{E_{g_1} + [\Delta E_c - (E_{c_2} - F_2)] - [F_1 - E_{v_1}]\} / q,$$

где $F_1 - E_{v_1}$ и $E_{c_2} - F_2$ обозначались через δ_1 и δ_2 соответственно.

Первым этапом в численном расчете энергетической зонной диаграммы является определение величин F_1 и F_2 способом, описанным выше в этом параграфе. Когда положение уровней Ферми известно, V_D рассчитывается по формуле (4.3.8). Затем из выражений (4.3.88) и (4.3.80) находят контактную разность потенциалов, приходящуюся на p -сторону:

$$V_{Dp} = V_D / [1 + \varepsilon_1 [N_{A_1}^- - N_{D_1}^+] / \varepsilon_2 (N_{D_2}^+ - N_{A_2}^-)] \quad (4.3.89)$$

и далее

$$V_{DN} = V_D - V_{Dp}. \quad (4.3.90)$$

После нахождения величин V_{DN} и V_{Dp} можно определить толщины обедненных слоев по формулам

$$x_p = [2\varepsilon_1 V_{Dp} / q (N_{A_1}^- - N_{D_1}^+)]^{1/2}, \quad (4.3.91)$$

$$x_N = [2\varepsilon_2 V_{DN} / q (N_{D_2}^+ - N_{A_2}^-)]^{1/2}, \quad (4.3.92)$$

которые получаются из выражений (4.3.76) и (4.3.81).

Для определения зависимости E_c и E_v от расстояния x проще всего положить $E_{v_1} = 0$ при $x > x_N$. Следует подчеркнуть, что, как видно из выражения (4.3.85), края зон изгибаются

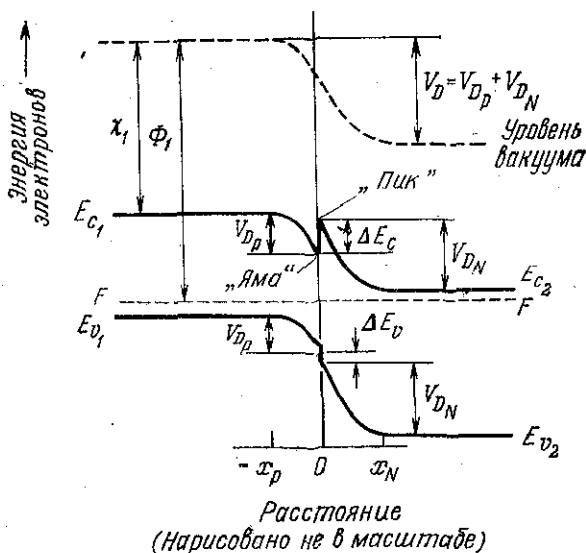


Рис. 4.3.8. Схематическое представление GaAs-Al_xGa_{1-x}As p-N-гетероперехода при тепловом равновесии; χ — электронное сродство, Φ — работа выхода, V_{Dp} и V_{DN} — контактная разность потенциалов, приходящаяся на на p- и N-стороны. Край зоны проводимости и валентной зоны с p-стороны обозначены через E_{c1} и E_{v1} , с N-стороны — через E_{c2} и E_{v2} .

в направлении, противоположном направлению изгиба кривой, описывающей электростатический потенциал. Как показано на рис. 4.3.8, край валентной зоны с p-стороны может быть описан формулами

$$E_{v1} = V_{DN} + V_{Dp} + \Delta E_v \quad \text{при } -\infty < x < -x_p \quad (4.3.93)$$

и с учетом выражения (4.3.75) для V_p

$$E_{v1} = V_{DN} + V_{Dp} + \Delta E_v - [q(N_{A1}^- - N_{D1}^+)/2\epsilon_1](x_p + x)^2 \quad \text{при } -x_p \leq x < 0. \quad (4.3.94)$$

Из формул (4.3.94) и (4.3.75) следует, что при x , стремящемся к нулю, E_{v1} стремится к $V_{DN} + \Delta E_v$. В точке $x = 0$ E_{v1} испытывает скачкообразное уменьшение на величину ΔE_v , так что с N-стороны E_{v2} описывается выражением

$$E_{v2} = V_{DN} - (V_N - V_{Dp}), \quad (4.3.95)$$

которое после подстановки V_N из формулы (4.3.83) принимает вид

$$E_{v2} = V_{DN} - [q(N_{D2}^+ - N_{A2}^-)/2\epsilon_2][x_N^2 - (x_N - x)^2] \quad \text{при } 0 < x \leq x_N \quad (4.3.96)$$

и

$$E_{v_2} = 0 \quad \text{при } x_N < x < \infty. \quad (4.3.97)$$

Поскольку $\Delta E_v = E_{g_2} - E_{g_1} - \Delta E_c$ [см. выражение (4.3.2)], форма края зоны проводимости с p -стороны может быть представлена следующим образом:

$$E_{c_1} = E_{v_1} + E_{g_1} = V_{D_N} + V_{D_p} + E_{g_2} - \Delta E_c \quad \text{при } -\infty < x < -x_p \quad (4.3.98)$$

и

$$E_{c_1} = V_{D_N} + V_{D_p} + E_{g_2} - \Delta E_c - [q(N_{A_1}^- - N_{D_1}^+)/2\epsilon_1](x_p + x)^2 \quad \text{при } -x_p \leq x < 0. \quad (4.3.99)$$

С N -стороны

$$E_{c_2} = E_{v_2} + E_{g_2} = V_{D_N} + E_{g_2} - [q(N_{D_2}^+ - N_{A_2}^-)/2\epsilon_2][x_N^2 - (x_N - x)^2] \quad \text{при } 0 < x \leq x_N \quad (4.3.100)$$

и

$$E_{c_2} = E_{v_2} + E_{g_2} = E_{g_2} \quad \text{при } x_N < x < \infty. \quad (4.3.101)$$

Формулы (4.3.93), (4.3.94), (4.3.96) и (4.3.97) определяют E_v как функцию от x ; E_{c_1} может быть получено просто сложением E_{g_1} и E_{v_1} . Аналогично E_{c_2} получается прибавлением E_{g_2} к E_{v_2} . Различные величины, которые обсуждались при рассмотрении гетероперехода, показаны на рис. 4.3.8. Небольшой опыт в построении таких энергетических зонных диаграмм позволяет рисовать диаграмму для гетероперехода по диаграмме для гомоперехода, добавляя просто соответствующие ΔE_c и ΔE_v , как описано в § 4 настоящей главы.

GaAs — Al_xGa_{1-x}As p — N -гетеропереходы при смещении

Полярность прямого смещения V_a показана на рис. 4.3.9, а. Как видно из рис. 4.3.9, б, V_a имеет полярность, противоположную полярности контактной разности потенциалов, показанной на рис. 4.3.7, в. Приложенное смещение распределяется между p - и N -сторонами:

$$V_a = V_1 + V_2, \quad (4.3.102)$$

где V_1 и V_2 — доли V_a , приходящиеся на p - и N -стороны соответственно. Полученные выше выражения для построения энергетической зонной диаграммы изменяются при наличии смещения. При $x = 0$ изменится граничное условие:

$$V_p |_{x=0} = (V_{D_p} - V_1), \quad (4.3.103)$$

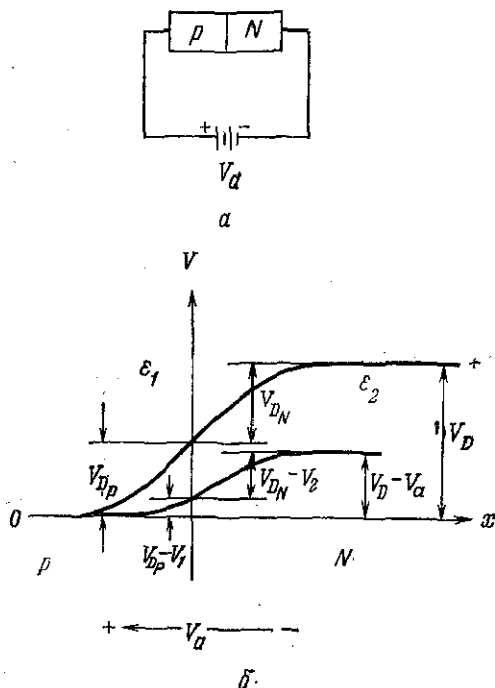


Рис. 4.3.9. Гетеропереход с приложенным внешним напряжением. a — полярность прямого смещения V_a ; b — зависимость потенциала от расстояния при прямом смещении, уменьшающем V_D на величину V_a .

и выражение (4.3.76) заменится на

$$(V_{Dp} - V_1) = (q/2\epsilon_1)(N_{A_1}^- - N_{D_1}^+) x_p^2. \quad (4.3.104)$$

Аналогично

$$(V_{Dn} - V_2) = (q/2\epsilon_2)(N_{D_2}^+ - N_{A_2}^-) x_n^2. \quad (4.3.105)$$

Из соотношений (4.3.86), (4.3.104) и (4.3.105) получаем

$$\frac{(V_{Dn} - V_2)}{(V_{Dp} - V_1)} = \frac{(N_{A_1}^- - N_{D_1}^+) \epsilon_1}{(N_{D_2}^+ - N_{A_2}^-) \epsilon_2}. \quad (4.3.106)$$

Вместо соотношения (4.3.80) будем иметь

$$(V_D - V_a) = (V_{Dp} - V_1) + (V_{Dn} - V_2). \quad (4.3.107)$$

Из выражений (4.3.106) и (4.3.107) следует

$$(V_{Dp} - V_1) = (V_D - V_a) / [1 + \epsilon_1(N_{A_1}^- - N_{D_1}^+) / \epsilon_2(N_{D_2}^+ - N_{A_2}^-)], \quad (4.3.108)$$

откуда можно определить V_1 , и далее

$$(V_{D_N} - V_2) = (V_D - V_a) - (V_{D_p} - V_1). \quad (4.3.109)$$

Толщины обедненных слоев определяются из формул (4.3.104) и (4.3.105):

$$x_p = [2\epsilon_1 (V_{D_p} - V_1) / q (N_{A_1}^- - N_{D_1}^+)]^{1/2}, \quad (4.3.110)$$

$$x_N = [2\epsilon_2 (V_{D_N} - V_2) / q (N_{D_2}^+ - N_{A_2}^-)]^{1/2}. \quad (4.3.111)$$

Эти выражения ясно показывают, что толщины обедненных слоев уменьшаются с увеличением прямого смещения. Обратное смещение дает $(V_D + V_a)$ и приводит к увеличению толщин обедненных слоев.

Положение краев зоны проводимости и валентной зоны в зависимости от расстояния x теперь будет определяться выражениями

$$E_{v_1} = V_{D_N} + V_{D_p} - V_a + \Delta E_v \quad \text{при } -\infty < x < -x_p \quad (4.3.112)$$

и

$$E_{v_1} = V_{D_N} + V_{D_p} - V_a + \Delta E_v - [q(N_{A_1}^- - N_{D_1}^+) / 2\epsilon_1] (x_p + x)^2 \quad \text{при } -x_p \leq x < 0 \quad (4.3.113)$$

с p -стороны и выражениями

$$E_{v_2} = (V_{D_N} - V_2) - [q(N_{D_2}^+ - N_{A_2}^-) / 2\epsilon_2] [x_N^2 - (x_N - x)^2] \quad \text{при } 0 < x \leq x_N \quad (4.3.114)$$

и

$$E_{v_2} = 0 \quad \text{при } x_N < x < \infty \quad (4.3.115)$$

с N -стороны; для зоны проводимости снова имеем $E_{c_1} = E_{v_1} + E_{g_1}$ и $E_{c_2} = E_{v_2} + E_{g_2}$. Ключевым выражением при наличии смещения является выражение (4.3.108), определяющее долю V_a , приходящуюся на p -сторону гетероперехода.

Формулы, необходимые для расчета энергетической зонной диаграммы $p-N$ -гетероперехода, даны в табл. 4.3.4. Обозначение

$$K = \left[1 + \frac{\epsilon_1 (N_{A_1}^- - N_{D_1}^+)}{\epsilon_2 (N_{D_2}^+ - N_{A_2}^-)} \right] \quad (4.3.116)$$

введено для упрощения выражений. Как показано на рис. 4.3.8, $E_c - F$ и $F - E_v$ в области пространственного заряда между $-x_p$ и x_N много больше, чем в области нейтральных полупроводников. Из формул (4.3.13) и (4.3.15) имеем в области пространственного заряда $N_D^+ \approx N_D$ при $F_c \leq E_D - 4kT$ и $N_A^- \approx N_A$ при

Таблица 4.3.4

Формулы для расчета энергетической зонной диаграммы
p-N-гетероперехода

$$\epsilon_x = (13,1 - 3,0x) \epsilon_0 \quad \epsilon_0 = 8,85 \cdot 10^{-14} \text{ Ф/см}$$

$$\text{Обозначение: } K = \left[1 + \frac{\epsilon_1 (N_{A_1}^- - N_{D_1}^+)}{\epsilon_2 (N_{D_2}^+ - N_{A_2}^-)} \right]$$

$$V_{D_p} = V_D / K \quad V_{D_N} = V_D - V_{D_p}$$

$$(V_{D_p} - V_1) = (V_D - V_a) / K \quad (V_{D_N} - V_2) = (V_D - V_a) - (V_{D_p} - V_1)$$

$$x_p = [2\epsilon_1 (V_{D_p} - V_1) / q (N_{A_1}^- - N_{D_1}^+)]^{1/2}$$

$$x_N = [2\epsilon_2 (V_{D_N} - V_2) / q (N_{D_2}^+ - N_{A_2}^-)]^{1/2}$$

С p-стороны:

$$E_{v_1} = V_{D_N} + V_{D_p} - V_a + \Delta E_v - [q (N_{A_1}^- - N_{D_1}^+) / 2\epsilon_1] (x_p + x)^2 \quad \text{при} \\ -x_p \leq x < 0$$

$$E_{c_1} = E_{v_1} + E_{g_1}$$

С N-стороны:

$$E_{v_2} = (V_{D_N} - V_2) - [q (N_{D_2}^+ - N_{A_2}^-) / 2\epsilon_2] [x_N^2 - (x_N - x)^2] \quad \text{при} \\ 0 < x \leq x_N$$

$$E_{c_2} = E_{v_2} + E_{g_2}$$

$F_0 > E_A + 4kT$. Поэтому в выражение (4.3.116) входят концентрации N_D и N_A полностью ионизованных примесей, т. е.

$$(N_{A_1}^- - N_{D_1}^+) = (N_{A_1} - N_{D_1}) \quad \text{и} \quad (N_{D_2}^+ - N_{A_2}^-) = (N_{D_2} - N_{A_2}).$$

Численный пример показывает, какая информация требуется для построения энергетической зонной диаграммы гетероперехода. Необходимые для расчета контактной разности потенциалов величины приведены в табл. 4.3.3. Уровни Ферми в выражении для V_D вычисляются по значениям p и N способом, описанным выше. Если для частного случая заданы концентрации p и N , то легко находятся значения F_1 и F_2 , в то время как концентрациям N_A и N_D с обеих p- и N-сторон гетероперехода в соответствии с величинами $E_A - E_v$ и $E_c - E_D$ должны быть приписаны такие значения, чтобы выполнялось условие электронейтральности в области нейтральных полупроводников. Если, наоборот, для обеих p- и N-сторон указаны значения N_A и N_D , то для нахождения p , N , F_1 и F_2 должны быть известны

$E_A - E_D$ и $E_C - E_D$. Другими словами, величины ρ , $N_{A_1} - N_{D_1}$, F_1 и N , $N_{D_2} - N_{A_2}$, F_2 должны быть связаны между собой самосогласованным образом. Часто информации бывает недостаточно, чтобы определить самосогласованные значения этих величин. Этой дилеммы можно избежать в том случае, когда имеется возможность приготовить экспериментальные образцы. Концентрация носителей в области нейтральных полупроводников и величины $N_{A_1}^- - N_{D_1}^+$ и $N_{D_2}^+ - N_{A_2}^-$ для полностью ионизованных примесей, которые требуются для расчетов в области пространственного заряда, легко могут быть определены из электрических измерений.

Концентрации носителей могут быть найдены из холловских измерений [54] на одиночных слоях, выращенных на полуизолирующих подложках из GaAs, легированного Cr. Наиболее удобной является методика Ван-дер-Пау [55] с образцами в форме клеверного листа. Величины $N_{A_1}^- - N_{D_1}^+$ и $N_{D_2}^+ - N_{A_2}^-$, входящие в выражения для E_{v_1} , E_{v_2} , E_c и E_{c_2} , могут быть определены из измерений емкости барьера Шотки [56] на слоях, которые используются для холловских измерений, или на слоях, выращенных на проводящих подложках из GaAs.

Если для GaAs указана концентрация примесей или носителей, то легко определить значение величин, необходимых для построения энергетической зонной диаграммы. Для обычных доноров $E_C - E_D$ составляет только 0,006 эВ при низкой концентрации доноров и стремится к нулю при $n > 2 \cdot 10^{16}$ см⁻³ (рис. 3.5.1, а). Поэтому доноры полностью ионизованы при концентрациях примесей, встречающихся в гетеролазерах, и $N_D^+ - N_A^-$ в области пространственного заряда равно n . Тщательные измерения Эрманиса и Вольфштинна [57] на легированном цинком GaAs в области концентраций, представляющих интерес для гетеролазеров, дали значение разности концентраций полностью ионизованных примесей $N_A^- - N_D^+$, примерно в 1,1 раза превышающее концентрацию дырок. Акцепторы Ge, Cd и Be должны вести себя так же, как и Zn. Поэтому на энергетической зонной диаграмме в той части, которая отвечает GaAs, в области пространственного заряда величину $N_D^+ - N_A^-$ можно положить равной n , если это слой n -типа, и $-1,1 \rho$, если это слой p -типа.

Более сложной оказывается ситуация для $Al_xGa_{1-x}As$. Для $N-Al_xGa_{1-x}As$ встречается ряд проблем при определении N по значениям N_D и N_A или разности концентраций полностью ионизованных примесей $N_D^+ - N_A^-$ по известному значению N . Как показано на рис. 4.3.3, б, для Te, являющегося типичной донорной примесью, $E_C - E_D$ меняется от 0 до 0,140 эВ при изменении x от $\sim 0,25$ до 0,40. Поэтому неопределенность в значении x

ведет к значительной неопределенности величины $E_c - E_D$. Кроме того, имеющиеся значения $E_c - E_D$ для $Al_xGa_{1-x}As$ выбирались исходя только из обычной ($T^{3/2}$) температурной зависимости N_c [43], в то время как из выражения (4.3.24) видно, что N_c содержит зависящие от температуры члены $\exp(-\Delta E^{L-\Gamma}/kT)$ и $\exp(-\Delta E^{X-\Gamma}/kT)$. Холловские измерения и измерения на барьере Шоттки, выполненные на $N-Al_xGa_{1-x}As$, показали, что для полностью ионизованных примесей $N_D^+ - N_A^-$ приблизительно вдвое превышает концентрацию носителей при N вблизи $1 - 2 \cdot 10^{17} \text{ см}^{-3}$. Это соотношение $N_D^+ - N_A^- = 2N$ будет выбрано для слоев $N-Al_{0,3}Ga_{0,7}As$ в примерах, рассматриваемых в последующей части этой главы.

Эти неопределенности не так существенны для $E_A - E_D$ в $P-Al_{0,3}Ga_{0,7}As$. Если выбрано значение $N_A^- - N_D^+$, то P и F_D могут быть получены из условия электронейтральности, выраженного формулой (3.7.9). Это может быть сделано на ЭВМ или с помощью графического представления зависимости N_D^+ , N_A^- и концентрации носителей от положения уровня Ферми. Графическое решение показывает также, что, вообще говоря, величины $N_A^- - N_D^+$ или $N_D^+ - N_A^-$ в области нейтрального полупроводника не те же самые, что и в области пространственного заряда, в которой доноры и акцепторы полностью ионизованы.

На рис. 4.3.10 показано, как графически определяется положение уровня Ферми в $P-Al_{0,3}Ga_{0,7}As$. По формулам из табл. 4.2.1

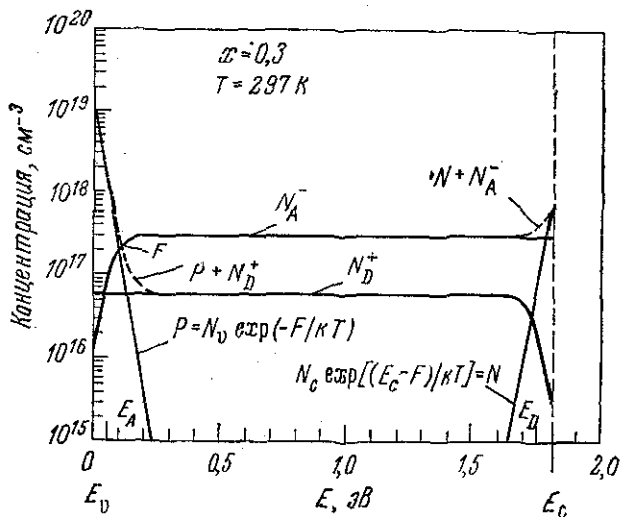


Рис. 4.3.10. Графический метод определения уровня Ферми в $P-Al_{0,3}Ga_{0,7}As$.

были получены следующие значения параметров: $E_g^\Gamma = 1,798$ эВ, $\Delta E^{L-\Gamma} = 0,103$ эВ, $\Delta E^{X-\Gamma} = 0,152$ эВ, $m_n^\Gamma = 0,092m_0$, $m_n^L = 0,58m_0$, $m_n^X = 0,83m_0$ и $m_p = 0,57m_0$. Выражения (4.3.24) и (4.3.39) при $T = 297$ К дают $N_c = 9,5 \cdot 10^{17}$ см $^{-3}$ и $N_v = 1,1 \cdot 10^{19}$ см $^{-3}$. Концентрации электронов и дырок, указанные на рис. 4.3.10, получены по формулам (4.3.25) и (4.3.30) соответственно. Считалось, что краю валентной зоны соответствует значение $E_v = 0$, так что $E_c = E_g$. При построении кривых на рис. 4.3.10 использовалось значение $E_A - E_v = E_c - E_D = 0,060$ эВ [43] при $N_A = 3 \cdot 10^{17}$ см $^{-3}$ и $N_D = 6 \cdot 10^{16}$ см $^{-3}$. Полный положительный заряд равен $P + N_D^+$, а полный отрицательный заряд равен $-N + N_A^-$. Соответствующие этим суммам кривые также построены на рис. 4.3.10. Условие электронейтральности выполняется в точке пересечения кривых $P + N_D^+$ и $N + N_A^-$, обозначенной на рисунке через F . Для $F = 0,107$ эВ в области нейтрального полупроводника $P = N_A^- - N_D^+ = 1,6 \cdot 10^{17}$ см $^{-3}$, и $N_A^- - N_D^+ = N_A - N_D = 2,4 \cdot 10^{17}$ см $^{-3}$ при полной ионизации примесей. В этом примере концентрация свободных носителей составляет примерно половину концентрации акцепторов. Описанный графический анализ показывает, как найти F , если N_A , N_D и $E_A - E_v$ (или N_D , N_A и $E_c - E_D$) известны. Он подчеркивает также тот факт, что обычно существует различие между степенями ионизации примесей в области пространственного заряда и в области нейтрального полупроводника.

На рис. 4.3.11, *а* изображена энергетическая зонная диаграмма GaAs—Al $_{0,3}$ Ga $_{0,7}$ As p — N -гетероперехода при нулевом смещении. Положение уровней Ферми определено по формуле (4.3.45). В области пространственного заряда с p -стороны, соответствующей GaAs, $p = 1 \cdot 10^{18}$ см $^{-3}$, $F_1 - E_v = 0,053$ эВ и $N_{A_1}^- - N_{D_1}^+ = 1,1p = 1,1 \cdot 10^{18}$ см $^{-3}$. С N -стороны для $N = 1,5 \times 10^{17}$ см $^{-3}$ $E_{c_2} - F_2 = 0,046$ эВ. Из сравнения результатов холловских измерений и измерений на барьере Шотки величина $N_{D_2}^+ - N_{A_2}^-$ в области пространственного заряда была взята равной $3 \cdot 10^{17}$ см $^{-3}$. На рис. 4.3.11, *б* показана энергетическая зонная диаграмма при прямом смещении, а на рис. 4.3.11, *в* — при обратном смещении. Примеры энергетических зонных диаграмм при нулевом смещении, отвечающих другим значениям концентраций, приведены на рис. 4.3.12. Концентрации и уровни Ферми указаны в подписи к рисунку. Рис. 4.3.12, *в* соответствует Al $_{0,08}$ Ga $_{0,92}$ As—Al $_{0,36}$ Ga $_{0,64}$ As p — N -гетеропереходу. Такие молярные доли AlAs характерны для ДГС-лазеров, используемых в системах волоконно-оптической связи. ДГС-лазер с $x = 0,08$ в активном слое излучает на длине волны $\sim 0,82$ мкм, на которой в оптическом волокне потери при передаче имеют минимум.

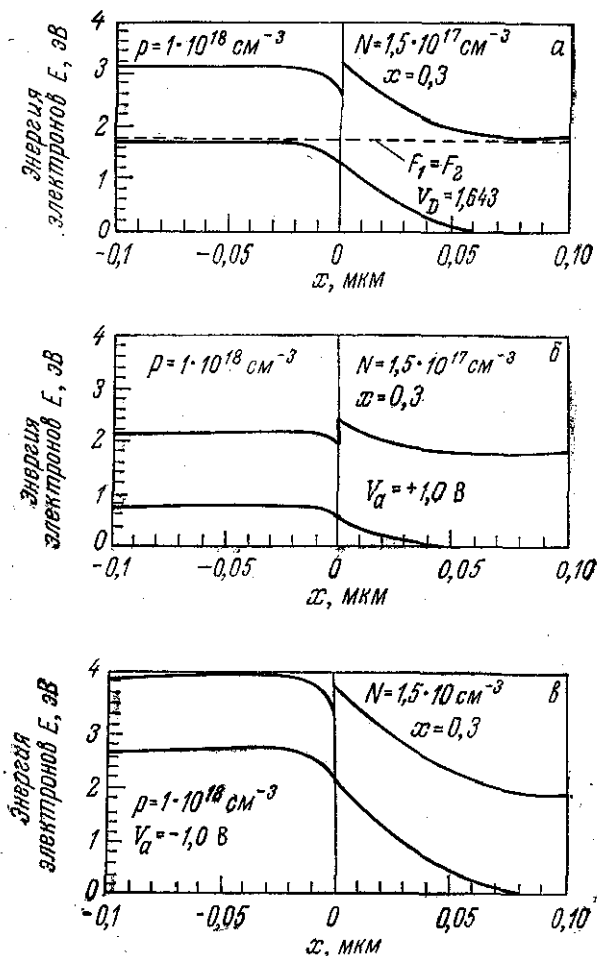


Рис. 4.3.11. Энергетическая зонная диаграмма GaAs — Al_{0,3}Ga_{0,7}As p — N-гетероперехода при T = 297 К. В области пространственного заряда $N_{A1}^- - N_{D1}^+ = 1,1 \cdot 10^{18} \text{ см}^{-3}$ и $N_{D2}^+ - N_{A2}^- = 3 \cdot 10^{17} \text{ см}^{-3}$. С p-стороны $F_1 - E_{v2} = 0,053 \text{ эВ}$, с N-стороны $E_{c2} - F_2 = 0,046 \text{ эВ}$. а — нулевое смещение; б — прямое смещение +1,0 В; в — обратное смещение -1,0 В.

GaAs — Al_xGa_{1-x}As n — P-гетеропереходы

Энергетические зонные диаграммы n — P-гетеропереходов строятся так же, как и диаграммы p — N-гетеропереходов. Строя для n — P-структуры диаграмму, подобную приведенной на рис. 4.3.1, но с F_1 вблизи E_{c1} и F_2 вблизи E_{v2} , получаем выражение для контактной разности потенциалов

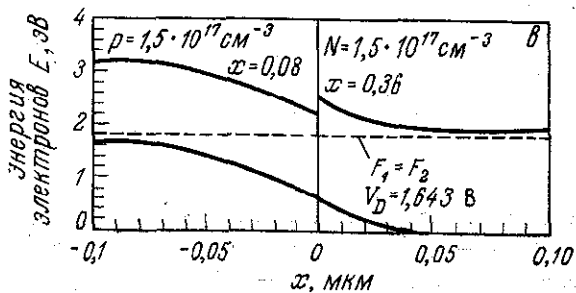
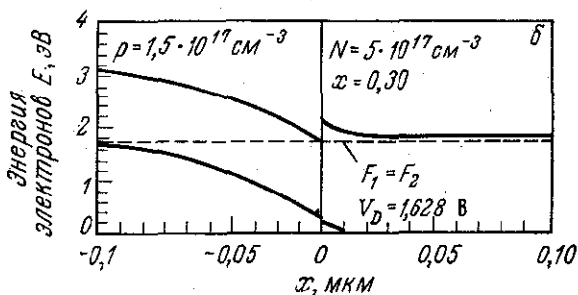
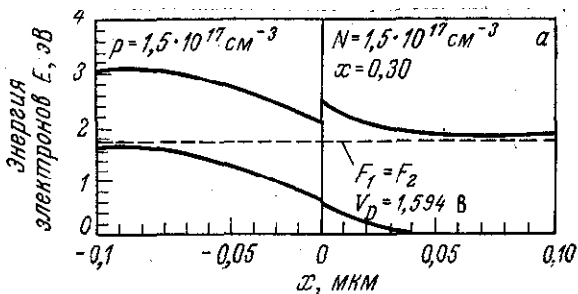


Рис. 4.3.12. Энергетические зонные диаграммы $p-N$ -гетеропереходов при нулевом смещении и $T = 297$ К. a — $\text{GaAs} - \text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ при $N_{A_1}^- - N_{D_1}^+ = 1,65 \cdot 10^{17} \text{ см}^{-3}$, $F_1 - E_{v_1} = 0,102$ эВ на p -стороне и $N_{D_2}^+ - N_{A_2}^- = 3 \cdot 10^{17} \text{ см}^{-3}$, $E_{c_2} - F_2 = 0,046$ эВ на N -стороне; b — $\text{GaAs} - \text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ при $N_{A_1}^- - N_{D_1}^+ = 1,65 \cdot 10^{17} \text{ см}^{-3}$, $F_1 - E_{v_1} = 0,102$ эВ на p -стороне и $N_{D_2}^+ - N_{A_2}^- = 1 \cdot 10^{18} \text{ см}^{-3}$, $E_{c_2} - F_2 = 0,012$ эВ на N -стороне; $в$ — $\text{Al}_{0,08}\text{Ga}_{0,92}\text{As} - \text{Al}_{0,36}\text{Ga}_{0,64}\text{As}$ при $N_{A_1}^- - N_{D_1}^+ = 1,65 \cdot 10^{17} \text{ см}^{-3}$, $F_1 - E_{v_1} = 0,104$ эВ на p -стороне и $N_{D_2}^+ - N_{A_2}^- = 3 \cdot 10^{17} \text{ см}^{-3}$, $E_{c_2} - F_2 = 0,059$ эВ на N -стороне.

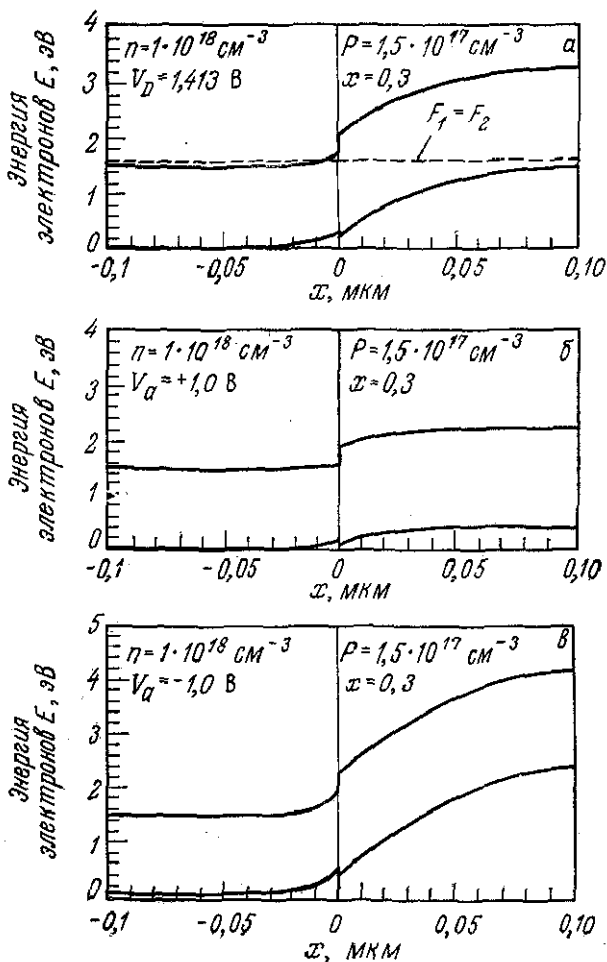


Рис. 4.3.13. Энергетическая зонная диаграмма GaAs — Al_{0,3}Ga_{0,7}As *n* — *P*-гетероперехода при $T = 297$ К с $F_1 - E_{c1} = 0,042$ эВ и $F - E_{v2} = 0,109$ эВ. В области пространственного заряда $N_{D1}^+ - N_{A1}^- = 1 \cdot 10^{18}$ см⁻³ и $N_{A2}^- - N_{D2}^+ = 2,3 \cdot 10^{17}$ см⁻³. а — нулевое смещение; б — прямое смещение +1,0 В; в — обратное смещение -1,0 В.

$$V_D(\text{В}) = (F_1 - F_2)/q = [E_{g2} - (F_{v2} - E_{v1}) - \Delta E_c - (E_{c1} - F_c)]/q. \quad (4.3.117)$$

Для E_{v1} снаружи обедненной области с *n*-стороны взято значение, равное ΔE_v , чтобы E_{v2} с *P*-стороны для $x > x_p$ обращалось в нуль при $V_a = V_D$. Такой выбор дает

$$E_{v1} = \Delta E_v \quad \text{при} \quad -\infty < x < -x_n. \quad (4.3.118)$$

$$E_{v_1} = \Delta E_v + V_{D_n} - V_1 - [q(N_{D_1}^+ - N_{A_1}^-)/2\epsilon_1][x_n^2 - (x_n + x)^2] \\ \text{при } -x_n \leq x < 0. \quad (4.3.119)$$

На P -стороне

$$E_{v_2} = V_{D_n} - V_1 + [q(N_{A_2}^- - N_{D_2}^+)/2\epsilon_2][x_p^2 - (x_p - x)^2] \\ \text{при } 0 < x \leq x_p \quad (4.3.120)$$

и

$$E_{v_2} = V_D - V_a \quad \text{при } x_p < x < \infty. \quad (4.3.121)$$

Край зоны проводимости вновь определяется по формулам $E_{c_1} = E_{g_1} + E_{v_1}$ и $E_{c_2} = E_{v_2} + E_{g_2}$. Контактные разности потенциалов находятся с помощью формулы (4.3.108), в которой V_{D_p} должно быть заменено на V_{D_n} , $N_{A_1}^- - N_{D_1}^+$ — на $N_{D_1}^+ - N_{A_1}^-$ и $N_{D_2}^+ - N_{A_2}^-$ — на $N_{A_2}^- - N_{D_2}^+$. Аналогичные замены должны быть произведены в формулах (4.3.110) и (4.3.111), чтобы определить толщины обедненных областей x_n и x_p . На рис. 4.3.13 показана при нулевом, прямом и обратном смещениях зонная диаграмма n — P -гетероперехода для $n = N_{D_1}^+ - N_{A_1}^- = 1 \cdot 10^{18} \text{ см}^{-3}$, $P = 1,5 \cdot 10^{17} \text{ см}^{-3}$, $N_{A_2}^- - N_{D_2}^+ = 2,3 \cdot 10^{17} \text{ см}^{-3}$ и $x = 0,3$.

Изотипные GaAs — Al_xGa_{1-x}As-гетеропереходы

При построении энергетических зонных диаграмм изотипных гетеропереходов используют те же выражения, что и для анизотипных гетеропереходов. Исходными пунктами при этом вновь являются известные значения ΔE_c и ΔE_v , а также требование, чтобы уровень Ферми при тепловом равновесии был постоянным во всем кристалле. Изотипные гетеропереходы рассматривались Андерсоном в его первых статьях [6, 7]. Подробное описание таких гетеропереходов можно найти в работах [10, 11].

Рассмотрим сначала p — P -гетеропереход. В этом случае на диаграмме, аналогичной изображенной на рис. 4.3.1, уровень F_2 располагается вблизи потолка валентной зоны E_{v_2} , что является единственным отличием между этими диаграммами. Как и в случае анизотипного гетероперехода, изгиб зон на гетерогранице определяется разностью уровней Ферми с p - и P -сторон. В изотипном p — P -гетеропереходе контактная разность потенциалов определяется выражением

$$V_D(V) = (F_1 - F_2)/q = [(F_1 - E_{v_1}) + \Delta E_v - (F_2 - E_{v_2})]/q. \quad (4.3.122)$$

Поскольку выбрана разность $F_1 - F_2$, прямое смещение соответствует положительному потенциалу на P -стороне. В этом случае V_D будет положительно, если $(F_1 - E_{v_1}) + \Delta E_v > (F_2 - E_{v_2})$, отри-

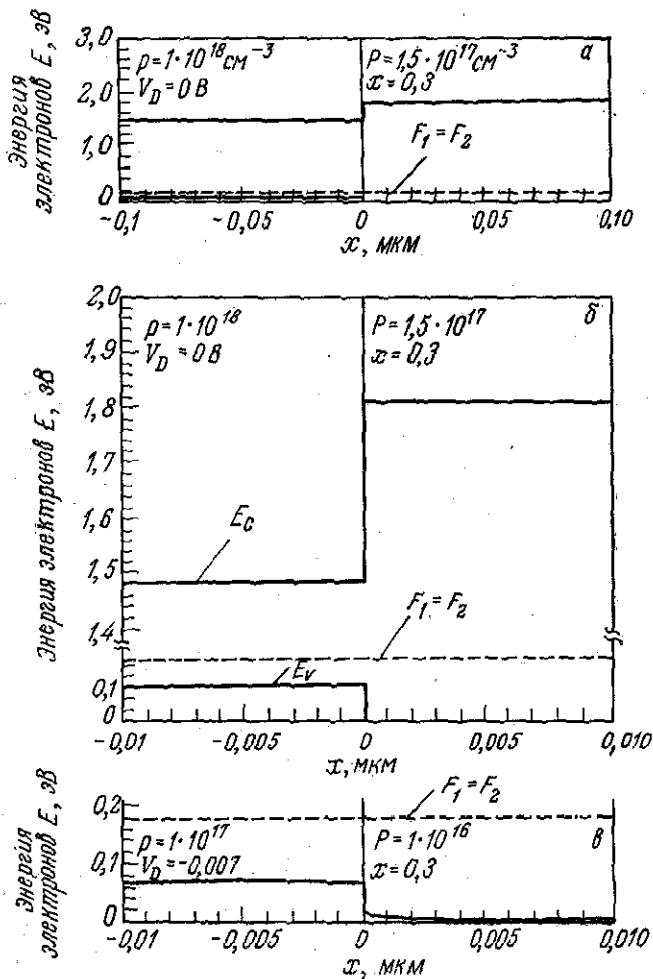


Рис. 4.3.14. Энергетическая зонная диаграмма GaAs — Al_{0,3}Ga_{0,7}As p — P-гетероперехода при T = 297 К. а — $N_{A_1}^- - N_{D_1}^+ = 1,1 \cdot 10^{18} \text{ см}^{-3}$ и $F_1 - E_{v_1} = 0,053 \text{ эВ}$ с p-стороны, $N_{A_2}^- - N_{D_2}^+ = 2,3 \cdot 10^{17} \text{ см}^{-3}$ и $F_2 - E_{v_2} = 0,109 \text{ эВ}$ с P-стороны; б — диаграмма рис. а в увеличенном масштабе; в — $N_{A_1}^- - N_{D_1}^+ = 1,1 \cdot 10^{17} \text{ см}^{-3}$ и $F_1 - E_{v_1} = 0,113 \text{ эВ}$ с p-стороны, $N_{A_2}^- - N_{D_2}^+ = 1,05 \cdot 10^{16} \text{ см}^{-3}$ и $F_2 - E_{v_2} = 0,178 \text{ эВ}$ с P-стороны.

цательно, если $(F_1 - E_{v_1}) + \Delta E_v < (F_2 - E_{v_2})$, и равно нулю, если $(F_1 - E_{v_1}) + \Delta E_v = (F_2 - E_{v_2})$. Пространственный заряд образован подвижными дырками с одной стороны гетерограницы и неподвижным зарядом с другой стороны.

Для обычно встречающихся концентраций примесей значение V_D находится вблизи нуля. Энергетическая зонная диаграмма определяется при этом выражениями, аналогичными выражениям для случая $n - P$ -гетероперехода. В формулах (4.3.119) и (4.3.120) $(N_{D_1}^+ - N_{A_1}^-)$ заменяется на $(N_{A_1}^- - N_{D_1}^+)$, а V_{D_n} — на V_{D_p} . Эти формулы для валентной зоны определены так, что $E_{V_2} = 0$ при $V_a = V_D$. Контактная разность потенциалов V_{D_p} , приходящаяся на p -сторону, определяется по формуле (4.3.89), разность $(V_{D_p} - V_1)$ — по формуле (4.3.108), а толщины обедненных слоев — по формулам (4.3.110) и (4.3.111) при условии, что во всех этих формулах сделана замена $(N_{D_1}^+ - N_{A_2}^-)$ на $(N_{A_2}^- - N_{D_1}^+)$. На рис. 4.3.14, а, б изображена энергетическая зонная диаграмма для $p = 1 \cdot 10^{18} \text{ см}^{-3}$, $P = 1,5 \cdot 10^{17} \text{ см}^{-3}$ и $x = 0,3$. Поскольку контактная разность потенциалов так мала, масштаб на рис. 4.3.14, б увеличен по обеим осям x и E .

Если $(F_2 - E_{v_2}) > (F_1 - E_{v_1}) + \Delta E_v$, то V_D отрицательна, и энергетическая зонная диаграмма определяется такими же формулами, как и для случая $p - N$ -гетероперехода, но в которых $(N_{D_2}^+ - N_{A_2}^-)$ заменено на $(N_{A_2}^- - N_{D_2}^+)$. Зависимость положения края валентной зоны от расстояния для $p = 1 \cdot 10^{17} \text{ см}^{-3}$ и $P = 1 \cdot 10^{16} \text{ см}^{-3}$ представлена на рис. 4.3.14, в. Масштабы по осям x и E увеличены, чтобы были видны небольшие изменения в E_v .

Рассмотрим теперь $n - N$ -гетеропереход. На диаграмме на рис. 4.3.1 уровень F_1 должен теперь находиться вблизи края зоны проводимости E_{c_1} . Контактная разность потенциалов определяется выражением

$$V_D = [\Delta E_c + (E_{c_1} - F_1) - (E_{c_2} - F_2)]/q. \quad (4.3.123)$$

Для обычно встречающихся примесных концентраций величина V_D будет положительна, поскольку ΔE_c , как правило, больше, чем разность $(E_{c_1} - F_1) - (E_{c_2} - F_2)$. Формулы, описывающие энергетические зонные диаграммы для рассматриваемого случая, аналогичны формулам для $p - N$ -гетероперехода, в которых $(N_{A_1}^- - N_{D_1}^+)$ заменено на $(N_{D_1}^+ - N_{A_1}^-)$. Прямому смещению соответствует отрицательный потенциал на N -стороне. На рис. 4.3.15, а изображена энергетическая зонная диаграмма $n - N$ -гетероструктуры с $n = 1,0 \cdot 10^{17} \text{ см}^{-3}$, $N = 1,5 \cdot 10^{17} \text{ см}^{-3}$ и $x = 0,3$ при нулевом смещении. Диаграмма гетероструктуры с $n = 1 \cdot 10^{18} \text{ см}^{-3}$ и $N = 1,5 \cdot 10^{17} \text{ см}^{-3}$ показана на рис. 4.3.15, б при нулевом смещении, а на рис. 4.3.15, в — при положительном смещении 0,15 В. Следует отметить, что энергетическая зонная диаграмма $n - N$ -гетероструктуры при $V_a \approx V_D$ (рис. 4.3.15, в) похожа на диаграмму $p - P$ -гетероструктуры (рис. 4.3.14, а).

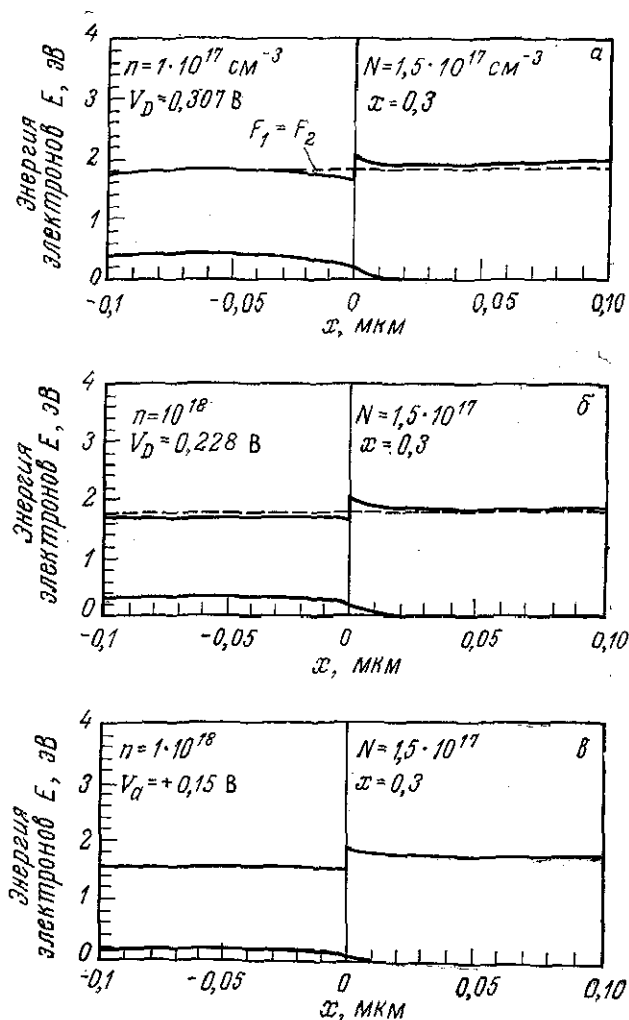


Рис. 4.3.15. Энергетическая зонная диаграмма GaAs — Al_{0,3}Ga_{0,7}As *p* — *N*-гетероперехода при $T = 297$ К. *а* — $N_{D_1}^+ - N_{A_1}^- = n$ и $E_{c_1} - F_1 = 0,035$ эВ с *p*-стороны, $N_{D_2}^+ - N_{A_2}^- = 3 \cdot 10^{17}$ см⁻³ и $E_{c_2} - F_2 = 0,046$ эВ с *N*-стороны; *б* — $N_{D_1}^+ - N_{A_1}^- = n$ и $F_1 - E_{c_1} = 0,042$ эВ с *p*-стороны, $N_{D_2}^+ - N_{A_2}^- = 3 \cdot 10^{17}$ см⁻³ и $E_{c_2} - F_2 = 0,046$ эВ с *N*-стороны; *в* — диаграмма для образца, соответствующего рис. б, при прямом смещении.

GaAs — Al_xGa_{1-x}As двусторонние гетероструктуры

Энергетические зонные диаграммы $N-p-P$ или $N-n-P$ ДГС получаются объединением обсужденных выше диаграмм анизотипных и изотипных гетеропереходов. Значения концентраций носителей и молярной доли AlAs, для которых построены диаграммы на рис. 4.3.11—4.3.15, характерны для реальных лазерных структур. Для удобства изображения диаграмм толщины слоев взяты равными 0,1 мкм. В ДГС-лазерах толщины N - или P -слоев должны быть равны 1,0—3,0 мкм, в то время как толщина n - или p -слоя для получения непрерывного режима работы при комнатной температуре должна лежать в пределах от 0,1 до 0,3 мкм.

$N-p-P$ ДГС-лазер при нулевом смещении показан на рис. 4.3.16, а, а при прямом смещении, равном 1,430 В, — на рис. 4.3.16, б. Разрыв зоны проводимости ΔE_c создает потенциальный барьер для электронов на $p-P$ -гетеропереходе и тем самым ограничивает инжектированные электроны активным слоем GaAs, в котором происходит излучательная рекомбина-

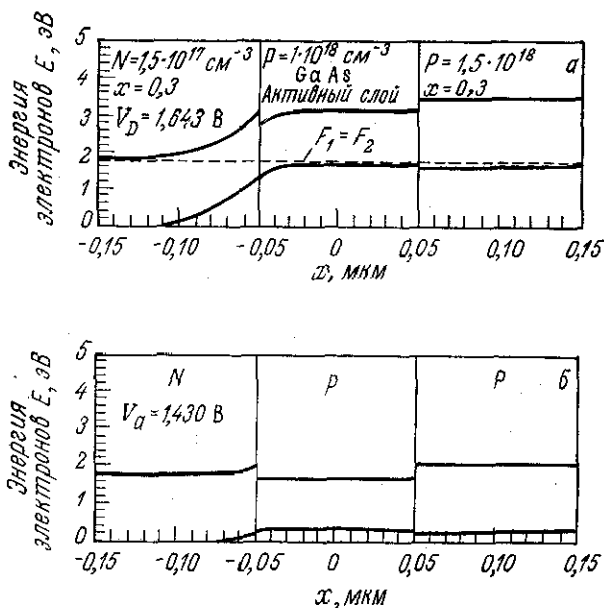


Рис. 4.3.16. Энергетическая зонная диаграмма GaAs — Al_{0,3}Ga_{0,7}As $N-p-P$ двусторонней гетероструктуры. При построении этой диаграммы взят $N-p$ -гетеропереход рис. 4.3.11, а и $p-P$ -гетеропереход рис. 4.3.14, а. а — нулевое смещение; б — прямое смещение 1,430 В.

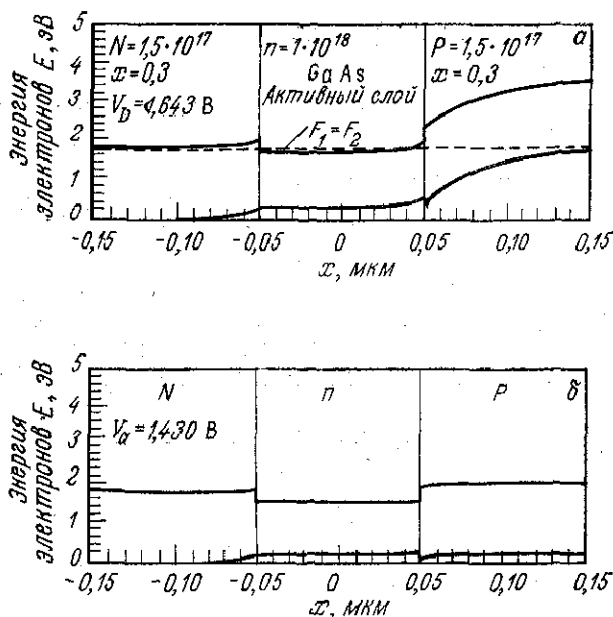


Рис. 4.3.17. Энергетическая зонная диаграмма GaAs — $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ N — n — P двусторонней гетероструктуры. При построении этой диаграммы взят N — n -гетеропереход рис. 4.3.15, б и n — P -гетеропереход рис. 4.3.13, а. а — нулевое смещение; б — прямое смещение 1,430 В.

ция. Разрыв валентной зоны ΔE_v вместе с оставшейся после приложения прямого смещения контактной разностью потенциалов создает потенциальный барьер для дырок на p — N -гетеропереходе и, таким образом, препятствуют инжекции дырок в N -слой. Эти барьеры приводят к ограничению как неосновных, так и основных носителей двусторонней гетероструктурой.

На рис. 4.3.17 показаны аналогичные диаграммы для N — n — P двусторонней гетероструктуры. Большая часть контактной разности потенциалов приходится на n — P -гетеропереход, однако некоторый изгиб зон присутствует и на N — n -гетеропереходе. Распределение V_a между N — n - и n — P -гетеропереходами определяется требованием непрерывности тока. В настоящее время распределение V_a между двумя гетеропереходами еще не установлено. На рисунке приложенное смещение произвольным образом распределено между n — P -гетеропереходом (1,280 В) и N — n -гетеропереходом (0,150 В). Следует также отметить, что n — N -гетеропереход между подложкой из GaAs и первым слоем $\text{Al}_x\text{Ga}_{1-x}\text{As}$, существующий в реальных приборах, не учитывался. Этот n — N -гетеропереход смещен в обратном направлении, но он, должно быть, находится в состоянии

пробоя, поскольку никаких данных о необходимости избыточного напряжения нет. Сравнение рис. 4.3.16, б и рис. 4.3.17, б показывает, что при больших прямых смещениях, характерных для ДГС-лазеров, энергетические зонные диаграммы $N-p-P$ и $N-n-P$ -структур имеют в основном одинаковый вид. Это частично объясняет схожесть поведения GaAs—Al_xGa_{1-x}As $N-p-P$ и $N-n-P$ ДГС-лазеров.

§ 4. ПЛАВНЫЕ ГЕТЕРОПЕРЕХОДЫ

Метод суперпозиции для построения энергетических зонных диаграмм

Удобным методом построения энергетических зонных диаграмм гетеропереходов является разделение выражений, описывающих зависимость положения краев зоны проводимости и валентной зоны от пространственной координаты x , на две части. Тогда энергетическая зонная диаграмма получается при сложении этих двух частей. Одна часть E_s связана с отличием контактной разности потенциалов от приложенного смещения, а другая обусловлена различием запрещенных зон и обозначается через ΔE_c и ΔE_v . Зависимость положения края валентной зоны от пространственной координаты x может быть представлена в виде

$$E_v(x) = E_s(x) + \Delta E_v(x), \quad (4.4.1)$$

в то время как для зоны проводимости

$$E_c(x) = E_s(x) + E_{g_1} + \Delta E_c(x), \quad (4.4.2)$$

где E_{g_1} — ширина запрещенной зоны более узкозонного материала, а $\Delta E_v(x)$ и $\Delta E_c(x)$ представляют пространственную зависимость ΔE_v и ΔE_c .

Сравнение выражения (4.4.1) с формулами (4.3.112) — (4.3.115) дает

$$E_s(x) = V_{D_N} + V_{D_p} - V_a \quad \text{при } -\infty < x < -x_p \quad (4.4.3)$$

и

$$E_s(x) = V_{D_N} + V_{D_p} - V_a - [q(N_{A_1}^- - N_{D_1}^+)/2\epsilon_1](x_p + x^2) \quad \text{при } -x_p \leq x < 0 \quad (4.4.4)$$

с p -стороны $p-N$ -гетероперехода. С N -стороны имеем

$$E_s(x) = (V_{D_N} - V_2) - [q(N_{D_2}^+ - N_{A_2}^-)/2\epsilon_2][x_N^2 - (x_N - x)^2] \quad \text{при } 0 < x \leq x_N \quad (4.4.5)$$

и

$$E_s(x) = 0 \quad \text{при } x_N < x < \infty. \quad (4.4.6)$$

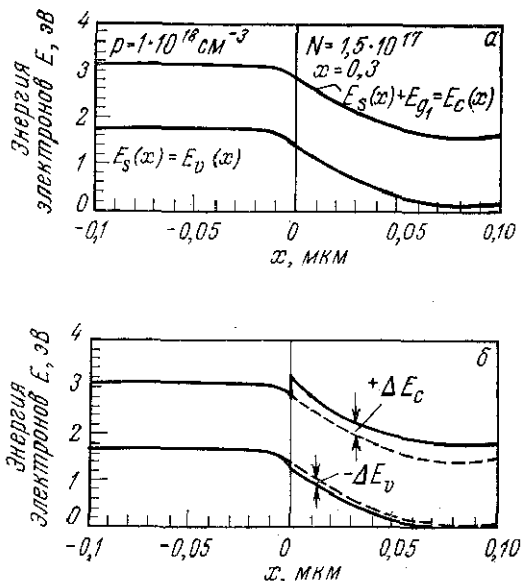


Рис. 4.4.1. Метод суперпозиции для построения изображенной на рис. 4.3.11 энергетической зонной диаграммы GaAs — Al_xGa_{1-x}As p — N -гетероперехода. a — «гомопереходная» часть энергетической зонной диаграммы; b — добавление $-\Delta E_v(x)$ к $E_v(x)$ и $\Delta E_c(x)$ к $E_c(x)$ для получения энергетической зонной диаграммы гетероперехода.

График зависимости E_s от x при этом будет таким же, как и для гомоперехода в полупроводнике с запрещенной зоной E_{g1} . Однако здесь есть одна тонкость. Для того чтобы получить правильное значение контактной разности потенциалов, нужно при определении уровня Ферми в полупроводнике с широкой запрещенной зоной использовать правильные значения параметров, например таких, как эффективная масса.

Применение этого метода на примере резкого p — N -гетероперехода, изображенного на рис. 4.3.11, a , продемонстрировано на рис. 4.4.1. На рис. 4.4.1, a $E_c(x)$ и $E_v(x)$ изображены с учетом только $E_s(x)$, что соответствует случаю гомоперехода. На рис. 4.4.1, b энергетическая зонная диаграмма, приведенная выше на рис. 4.3.11, a , получается добавлением к $E_c(x)$ при $x > 0$ постоянной величины ΔE_c и вычитанием из $E_v(x)$ постоянной величины ΔE_v . На рис. 4.4.1, a уровень отсчета выбран так, что $E_v(x)$ на рис. 4.4.1, b стремится к нулю на больших положительных расстояниях. Этот метод удобно применять при схематическом изображении энергетических зонных диаграмм для получения приближенных выражений. Его полезно использовать также в упрощенной модели плавных гетеропереходов.

Энергетические зонные диаграммы плавных гетеропереходов

В настоящей главе мы рассматривали до сих пор только резкие гетеропереходы; однако выражения (4.4.1) и (4.4.2) справедливы для любого вида пространственной зависимости $\Delta E_v(x)$ и $\Delta E_c(x)$. Измерения Вомэка и Редикера [58] фотоотклика GaAs — Al_xGa_{1-x}As-гетеропереходов, изготовленных жидкофазной эпитаксией (ЖФЭ), трактовались в рамках модели плавно меняющейся запрещенной зоны. В результате недавних измерений фотоотклика p — N -гетеропереходов [59], изготовленных эпитаксией из молекулярных пучков (ЭМП), было обнаружено такое же поведение этих гетеропереходов, как и приготовленных ЖФЭ. Измерения Дингла [39] на многослойных гетероструктурах, выращенных ЭМП, показали, что эти гетероструктуры резкие. Поэтому гетеропереходы, изготовленные методом ЖФЭ, также могут быть резкими. Несмотря на эту неопределенность, полезно рассмотреть эффекты, вызываемые плавностью гетеропереходов. Модель, которая излагается здесь, является упрощенной моделью для гетеропереходов с плавно меняющейся запрещенной зоной, предложенной Ченгом и др. [60], и основана на обобщенной модели Олдхэма и Милнса [61].

В модели Олдхэма и Милнса [61] необходимо решать на ЭВМ уравнение Пуассона. Применяя же метод суперпозиции, изложенный в первой части этого параграфа, Ченг и др. [60] показали, что использование простых выражений для величины $\Delta E_c(x)$, которая добавляется к $E_s(x)$, дает те же результаты, что и модель Олдхэма и Милнса. Они нашли, что плавное изменение удобно представлять гиперболическим тангенсом. Для $\Delta E_c(x)$ использовалось выражение [60]

$$\Delta E_c(x) = (\Delta E_c/2) \{1 + \text{th}[(x - x_0)/l]\}, \quad (4.4.7)$$

где x_0 соответствует центру области, в которой происходит изменение $\Delta E_c(x)$, а l — характерная длина изменения. Напомним, что гиперболический тангенс равен -1 при больших отрицательных значениях аргумента, $+1$ при больших положительных значениях аргумента и 0 , когда аргумент равен нулю. Если область переменной запрещенной зоны целиком лежит со стороны широкозонного слоя при $x \geq 0$, то удобно представление

$$\Delta E_c(x) = \Delta E_c \text{th}(x/l) \quad \text{при } x > 0. \quad (4.4.8)$$

Для $\Delta E_v(x)$ используются выражения, аналогичные выражениям (4.4.7) и (4.4.8).

В качестве первого примера рассмотрим случай, когда область переменной запрещенной зоны находится только с N -стороны, а $\Delta E_c(x)$ определяется формулой (4.4.8) с параметром

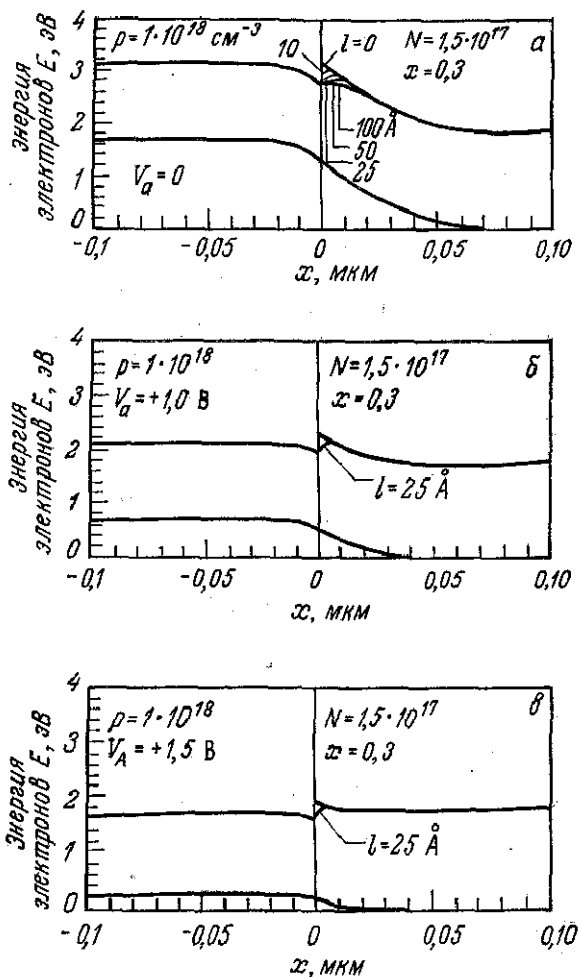


Рис. 4.4.2. Энергетическая зонная диаграмма плавного GaAs — Al_{0,3}Ga_{0,7}As p — N -гетероперехода, соответствующая резкому гетеропереходу на рис. 4.3.11, при $\Delta E_c(x) = \Delta E_c \text{th}(x/l)$ для $x > 0$. a — нулевое смещение и $l = 0, 10, 25, 50$ и 100 \AA ; $б$ — сравнение резкого и плавного гетеропереходов при $l = 25 \text{ \AA}$ и $V_a = +1,0 \text{ В}$; $в$ — сравнение резкого и плавного гетеропереходов при $l = 25 \text{ \AA}$ и $V_a = +1,5 \text{ В}$.

плавности изменения $l = 0, 10, 25, 50, 100 \text{ \AA}$. Получающаяся в этом случае энергетическая зонная диаграмма показана на рис. 4.4.2, a . На этом рисунке продемонстрированы изменения края зоны проводимости, обусловленные градиентом состава. При этом влияние, оказываемое на край валентной зоны, слиш-

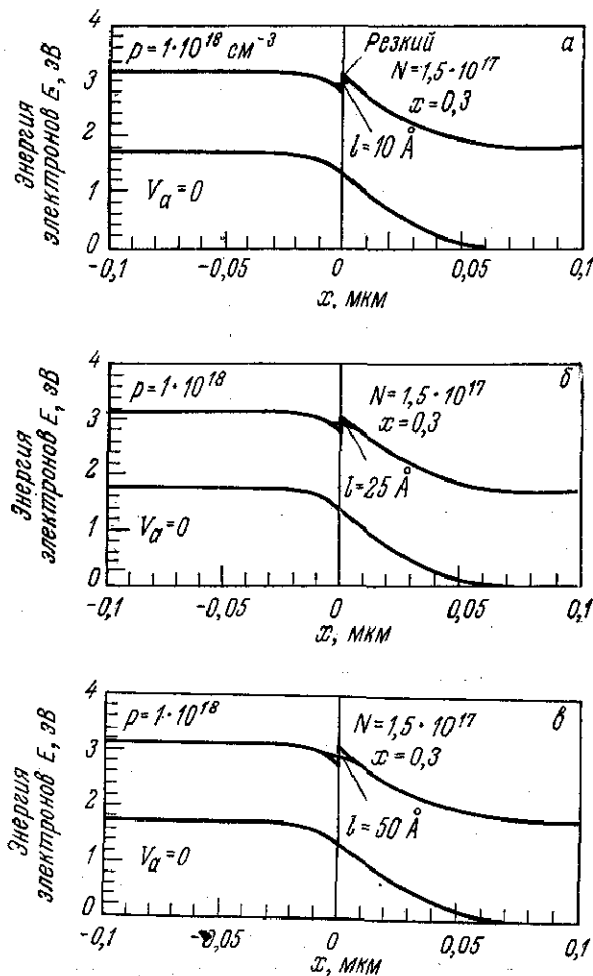


Рис. 4.4.3. Энергетическая зонная диаграмма плавного GaAs — $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ p — N -гетероперехода, соответствующего резкому гетеропереходу на рис. 4.3.11, при $\Delta E_c(x) = (\Delta E_c/2) [1 + \text{th}(x/l)]$. *a* — $l = 10 \text{ \AA}$; *б* — $l = 25 \text{ \AA}$; *в* — $l = 50 \text{ \AA}$.

ком мало, чтобы оно было заметно в том масштабе, который используется на рисунке. Зависимость вида энергетической зонной диаграммы от приложенного смещения показана на рис. 4.4.2, *б*, *в*, для которых $l = 25 \text{ \AA}$ и прямое смещение равно 1,0 и 1,5 В соответственно. Когда на плавный гетеропереход подается прямое смещение, барьеры, аналогичные тем, которые существуют в резком гетеропереходе, разрушаются. В этом случае барьер

распределяется по всей области переменной запрещенной зоны. Рассмотрение этой простой модели в широком диапазоне примесных концентраций может быть найдено в работе [60].

Энергетические зонные диаграммы для случая, когда область, в которой происходит плавное изменение ширины запрещенной зоны, лежит с обеих сторон $p-N$ -перехода, чему соответствует формула (4.4.7), показаны на рис. 4.4.3 для $l = 10, 25$ и 50 \AA . Как и в предыдущем примере, прямое смещение разрушает барьер, что хорошо видно на рис. 4.3.2, б, в.

§ 5. ЭЛЕКТРИЧЕСКИЕ СВОЙСТВА

Вольт-емкостная зависимость

Мы приведем здесь только выражения, полученные Андерсоном [6, 7] и описывающие вольт-емкостную зависимость. Влияние на емкость, оказываемое наличием состояний на границе раздела [62] и смещением гетероперехода относительно $p-n$ -перехода [63], не рассматривается, поскольку оно мало в гетеропереходах, подходящих для использования в инжекционных лазерах. В качестве примера рассмотрим $p-N$ -гетеропереход. Толщины обедненных слоев с p - и N -сторон определяются выражениями (4.3.110) и (4.3.111) соответственно. Подстановка выражения (4.3.108) для $V_{Dp} - V_1$ в выражение (4.3.110) дает x_p в виде

$$x_p = \left\{ \frac{2\epsilon_1}{q} \frac{\epsilon_2(V_D - V_a)(N_{D_2}^+ - N_{A_2}^-)}{(N_{A_1}^- - N_{D_1}^+)[\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)]} \right\}^{1/2} \quad (4.5.1)$$

Как уже говорилось выше, все доноры и акцепторы считаются полностью ионизованными в обедненной области. Для $V_{DN} - V_2$ из формул (4.3.106) и (4.3.107) можно получить

$$V_{DN} - V_2 = \frac{(V_D - V_a)(N_{A_1}^- - N_{D_1}^+)\epsilon_1}{\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)}. \quad (4.5.2)$$

Тогда из выражения (4.3.111) следует

$$x_N = \left\{ \frac{2\epsilon_1\epsilon_2(V_D - V_a)(N_{A_1}^- - N_{D_1}^+)}{q(N_{D_2}^+ - N_{A_2}^-)[\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)]} \right\}^{1/2} \quad (4.5.3)$$

Полная толщина обедненного слоя W равна

$$W = x_p + x_N = \left\{ \frac{2\epsilon_1\epsilon_2(V_D - V_a)[(N_{A_1}^- - N_{D_1}^+) + (N_{D_2}^+ - N_{A_2}^-)]^2}{q[\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)](N_{A_1}^- - N_{D_1}^+)(N_{D_2}^+ - N_{A_2}^-)} \right\}^{1/2} \quad (4.5.4)$$

Из этой формулы видно, что W меняется как квадратный корень из приложенного напряжения.

Как было показано на рис. 4.3.7, *a*, обедненная область состоит из двух заряженных слоев, содержащих равные и противоположные по знаку заряды. Это равенство зарядов выражалось формулой (4.3.86). Поэтому положительный или отрицательный пространственный заряд $|Q|$, приходящийся на единицу поверхности, равен

$$|Q| = q(N_{A_1}^- - N_{D_1}^+) x_p = q(N_{D_2}^+ - N_{A_2}^-) x_n. \quad (4.5.5)$$

Из этой формулы и выражений (4.5.1) для x_p и (4.5.3) для x_n следует

$$|Q| = \left[\frac{2q\epsilon_1\epsilon_2(V_D - V_a)(N_{D_2}^+ - N_{A_2}^-)(N_{A_1}^- - N_{D_1}^+)}{\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)} \right]^{1/2}. \quad (4.5.6)$$

Емкость C обедненной области на единицу площади a определяется как $C/a = |dQ/dV_a|$, где $d|Q|$ — приращение заряда на единице площади при бесконечно малом изменении приложенного напряжения dV_a . Следовательно,

$$\frac{C}{a} = \left| \frac{dQ}{dV_a} \right| = \left| \frac{d|Q|}{d(V_D - V_a)} \right|. \quad (4.5.7)$$

Беря производную от выражения (4.5.6) по $V_D - V_a$, получаем

$$\frac{C}{a} = \left\{ \frac{q\epsilon_1\epsilon_2(N_{A_1}^- - N_{D_1}^+)(N_{D_2}^+ - N_{A_2}^-)}{2(V_D - V_a)[\epsilon_1(N_{A_1}^- - N_{D_1}^+) + \epsilon_2(N_{D_2}^+ - N_{A_2}^-)]} \right\}^{1/2}. \quad (4.5.8)$$

Построив график зависимости $1/C^2$ от V_a , можно определить величину V_D , которая, как это видно из формулы (4.5.8), определяет точку пересечения графика с осью $1/C^2 = 0$. Измерив теперь C/a , величину ΔE_c можно найти по формуле (4.3.8), если известны уровни Ферми и ширины запрещенных зон с каждой стороны гетероперехода. Построить таким способом точное значение ΔE_c трудно из-за неопределенности численных значений величин, используемых для вычисления F_1 и F_2 . Ховарт и Фойхт [64] использовали измерения емкости Ge—Al_xGa_{1-x}As и GaAs—Al_xGa_{1-x}As-гетеропереходов для определения ΔE_c и ΔE_v . Их результаты показали, что действительно большая часть разности ширин запрещенных зон для GaAs—Al_xGa_{1-x}As-гетероперехода приходится на скачок в зоне проводимости. Полученное ими значение ΔE_v было меньше, чем значение, полученное Динглом [39].

Вольт-амперные характеристики

В настоящее время еще не ясно, правильно или нет модель Андерсона, дающая резкие скачки в зоне проводимости, описы-

вает реальные гетеропереходы. Как видно из рассмотрения изображенного на рис. 4.4.2, б, в плавного гетероперехода при большом прямом смещении, ограничение для носителей в гетеролазерах сохраняется независимо от того, является ли скачок в зонах резким или нет. Тот факт, что экспериментально не наблюдаются какие-либо эффекты, обусловленные скачкообразным изменением зонной структуры, например вызванные наличием выступающей над плоским краем зоны E_c , части «пики» на рис. 4.3.8, говорит о том, что современное представление энергетической зонной диаграммы не является полным. Независимо от того, как описывается разрыв запрещенной зоны, приведенные в § 3 настоящей главы выражения, определяющие ширину обедненной области и контактную разность потенциалов, полезны для описания важных свойств гетероперехода. Однако строгая теория вольт-амперных характеристик потребует знания правильной энергетической зонной диаграммы. Вольт-амперная зависимость, которая рассматривается здесь, описывает обычный диффузионный ток в гомопереходе [64]. В гетеропереходе диффузионный ток определяется главным образом инжекцией из широкозонного в узкозонный полупроводник. Этот диффузионный ток, инжектированный в активную область, поддерживает излучательную рекомбинацию, необходимую для вынужденного излучения. В приборах, использующих контактные явления, обычно встречаются и другие механизмы тока, например такие, как туннелирование и рекомбинация в области пространственного заряда. Однако они не являются доминирующими при высоких плотностях тока, необходимых для вынужденного излучения. Эти безызлучательные механизмы тока и эффекты, обусловленные наличием состояний на границе раздела, обсуждаются в работе [10] и не будут рассматриваться здесь.

В идеальной модели, описывающей вольт-амперную зависимость, делаются четыре предположения [65]:

1) обедненный слой пространственного заряда имеет резкие границы, как это показано на рис. 4.3.7, так что вне обедненной области полупроводник считается нейтральным;

2) концентрации носителей могут быть представлены с помощью экспоненциального приближения функции Ферми — Дирака, как это сделано в выражении (4.3.25);

3) концентрации инжектированных неосновных носителей малы по сравнению с концентрациями основных носителей;

4) в обедненной области нет токов генерации или рекомбинации, так что токи электронов и дырок постоянны в этой области.

Хотя приводимый ниже анализ относится к случаю низкого уровня накачки, он тем не менее позволяет определить все необходимые понятия. Зависимость тока от напряжения для предельного случая сильной накачки, когда $n \approx p$, может быть получена

распространением на этот случай результатов, найденных для низкого уровня накачки.

Рассмотрим сначала зависимость концентраций носителей от расстояния для изображенного на рис. 4.3.11, а p — N -гетероперехода, находящегося при тепловом равновесии. Концентрации электронов и дырок при экспоненциальном приближении функции распределения, которое называется распределением Больцмана, определены следующим образом [см. (4.3.25) и (4.3.40)]:

$$N \text{ или } n = N_c \exp[(F_c - E_c)/kT],$$

$$P \text{ или } p = N_v \exp[(E_v - F_v)/kT].$$

Поскольку молярные доли x арсенида алюминия различны с каждой стороны гетероперехода, для каждого полупроводника должны использоваться свои значения величин N_c и N_v . Для приведенного на рис. 4.3.11, а примера с $p = 1 \cdot 10^{18} \text{ см}^{-3}$, $N = 1,5 \cdot 10^{17} \text{ см}^{-3}$ и $x = 0,3$ зависимость концентрации основных носителей от расстояния имеет вид, изображенный на рис. 4.5.1. При увеличении в обедненной области разности $E_v - F_v$ или $E_c - F_c$ концентрации носителей быстро уменьшаются. Концентрация неосновных носителей — электронов с p -стороны n_p и концентрация неосновных носителей — дырок с N -стороны P_N , при тепловом равновесии равны

$$n_p = n_i^2/p_0, \quad (4.5.9)$$

$$P_N = N_i^2/N_0, \quad (4.5.10)$$

где N_i — собственная концентрация носителей с N -стороны, а индекс 0 говорит о том, что берутся концентрации при тепловом

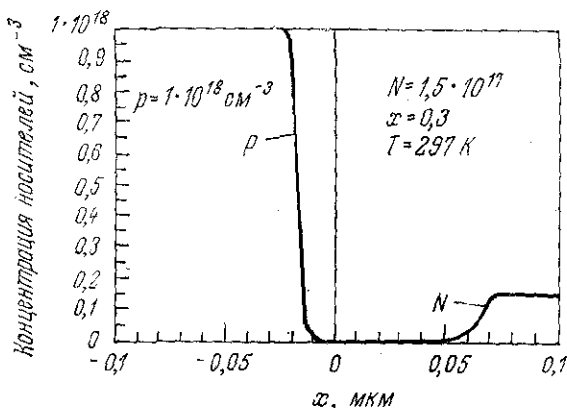


Рис. 4.5.1. Зависимость концентраций основных носителей от расстояния для изображенного на рис. 4.3.11 резкого $\text{GaAs}-\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ p — N -гетероперехода.

равновесии. Для квадрата собственной концентрации выше получена формула (4.3.43). При значениях эффективных масс в GaAs, приведенных в табл. 4.2.1, и $E_{g_i} = 1,424$ эВ при $T = 297$ К имеем $n_i^2 = 2,44 \cdot 10^{12}$ см⁻⁶ и $n_{p_0} = 2,44 \cdot 10^{-6}$ см⁻³, для $p_0 = 1 \cdot 10^{18}$ см⁻³, в то время как с N -стороны $N_i^2 = 3,3 \cdot 10^6$ см⁻⁶ и $P_{N_0} = 2,2 \cdot 10^{-11}$ см⁻³ для $N_0 = 1,5 \cdot 10^{17}$ см⁻³. Численные значения приведены с целью продемонстрировать порядки величин.

Для того чтобы рассмотреть отклонения от теплового равновесия, удобно записать формулу (4.3.25) в немного иной форме. Когда $F_c = F_v$ соответствует уровню Ферми в собственном кристалле, $n = p = n_i$ и $N = P = N_i$. Рассматривая только формулу (4.3.25), имеем

$$n_i = N_c \exp[(F_i - E_c)/kT] \quad (4.5.11)$$

и, следовательно,

$$n = n_i \exp[(F_c - F_i)/kT]. \quad (4.5.12)$$

Аналогично

$$p = n_i \exp[(F_i - F_v)/kT], \quad (4.5.13)$$

где n_i относится к p -стороне. Такие же выражения с N_i могут быть записаны для N -стороны. Далее, потенциал ψ для уровня Ферми в собственном кристалле и потенциал ϕ для уровня Ферми в реальном кристалле определяются как

$$\psi \equiv -F_i/q \quad (4.5.14)$$

и

$$\phi \equiv -F_v/q \equiv -F_c/q. \quad (4.5.15)$$

При этом формулы (4.5.12) и (4.5.13) принимают вид

$$n = n_i \exp[q(\psi - \phi)/kT], \quad (4.5.16)$$

$$p = n_i \exp[q(\phi - \psi)/kT]. \quad (4.5.17)$$

Произведение этих выражений дает $np = n_i^2$.

Приложение прямого смещения V_a уменьшает контактную разность потенциалов, как это выражено формулой (4.3.107). Для $V_a = +0,75$ В уровни Ферми, как показано на рис. 4.5.2, *a*, разделяются на эту величину. Концентрации неосновных носителей больше не являются равновесными величинами. Неравновесные уровни Ферми ϕ или F называются квазиуровнями Ферми и обозначаются ϕ_n и ϕ_p для электронов и дырок соответственно. Концентрации носителей при этом равны

$$n = n_i \exp[q(\psi - \phi_n)/kT] \quad (4.5.18)$$

и

$$p = n_i \exp[q(\phi_p - \psi)/kT]. \quad (4.5.19)$$

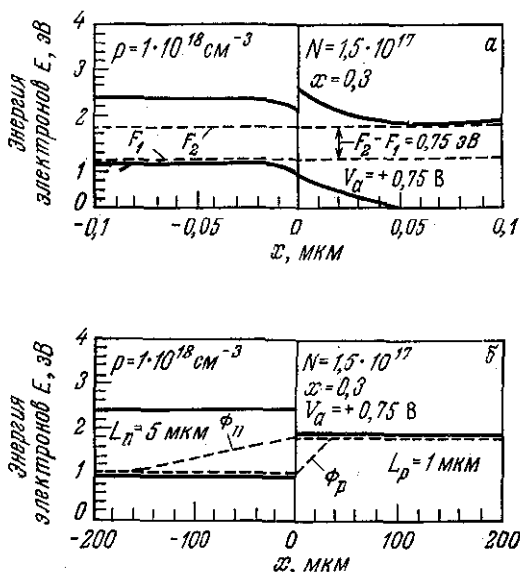


Рис. 4.5.2. Квазиуровни Ферми в изображенном на рис. 4.3.11 GaAs — $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ p — N -гетеропереходе. a — разделение квазиуровней Ферми вблизи области пространственного заряда; b — разделение квазиуровней Ферми в объеме нейтрального полупроводника. Для неосновных носителей длины диффузии электронов и дырок равны L_n и L_p соответственно.

Кроме того,

$$F_c - F_v = -q\phi_n + q\phi_p = -q(\phi_n - \phi_p) = qV_a, \quad (4.5.20)$$

или

$$V_a = \phi_p - \phi_n. \quad (4.5.21)$$

Произведение np теперь записывается в виде

$$np = n_i^2 \exp[q(\phi_p - \phi_n)/kT]. \quad (4.5.22)$$

Объединяя формулы (4.5.21) и (4.5.22), получаем

$$np = n_i^2 \exp(qV_a/kT). \quad (4.5.23)$$

Концентрация неосновных носителей n_p с p -стороны на краю обедненной области, которому соответствует $x = -x_p$, определяется из формулы (4.5.23):

$$n_p = (n_i^2/p) \exp(qV_a/kT) = n_{p0} \exp(qV_a/kT), \quad (4.5.24)$$

где n_{p_0} — равновесная концентрация неосновных носителей. С N -стороны в точке $x = x_N$ концентрация неосновных носителей — дырок равна

$$P_N = P_{N_0} \exp(qV_a/kT), \quad (4.5.25)$$

где P_{N_0} — значение равновесной концентрации. Для того чтобы в дальнейшем приступить к выводу выражений, определяющих вольт-амперные характеристики p — N -гетероперехода, необходимо сейчас рассмотреть, как перенос носителей заряда описывается уравнением непрерывности.

Уравнение непрерывности выражает тот факт, что производная по времени от концентрации электронов или дырок равна дивергенции потока носителей плюс разность между скоростями генерации и рекомбинации. В одномерном случае

$$\partial n/\partial t = (1/q) (\partial/\partial x) i_n + g(x) - [(n - n_0)/\tau_n] \quad (4.5.26)$$

и

$$\partial p/\partial t = -(1/q) (\partial/\partial x) i_p + g(x) - [(p - p_0)/\tau_p]. \quad (4.5.27)$$

В этих уравнениях i_n и i_p обозначают плотности тока электронов и дырок, $g(x)$ — скорость генерации носителей, n_0 и p_0 — равновесные значения концентраций n и p , а τ_n и τ_p — времена жизни электронов и дырок соответственно. Сохраняется электрическая нейтральность, выражаемая формулой $(n - n_0) = (p - p_0)$. Как показано ниже, уравнения (4.5.26) и (4.5.27) могут быть сведены к упрощенной форме. Более строгая трактовка этих уравнений дана в работе Мак-Келви [66].

В состоянии равновесия $\partial n/\partial t$ и $\partial p/\partial t$ равны нулю, и при отсутствии внешнего источника возбуждения, например света, $g(x)$ также равно нулю. Плотности диффузионных токов определены формулами (4.3.48) и (4.3.50), а токи дрейфа — формулами (4.3.51) и (4.3.52). В электронейтральном материале электрическое поле \mathcal{E} берется равным нулю; поэтому будут рассматриваться только диффузионные составляющие (4.3.50) и (4.3.48):

$$i_n = qD_n (dn/dx),$$

$$i_p = -qD_p (dp/dx).$$

Подстановка этих выражений в уравнения (4.5.26) и (4.5.27) для состояния равновесия и при $g(x) = 0$ дает

$$(d^2n/dx^2) - [(n - n_0)/D_n\tau_n] = 0, \quad (4.5.28)$$

$$(d^2p/dx^2) - [(p - p_0)/D_p\tau_p] = 0. \quad (4.5.29)$$

Для концентрации неосновных носителей — электронов в материале p -типа с $n_0 = n_{p_0}$ решение уравнения (4.5.28) имеет вид

$$n = C_1 \exp[-x/(D_n\tau_n)^{1/2}] + C_2 \exp[x/(D_n\tau_n)^{1/2}] + n_{p_0}. \quad (4.5.30)$$

При граничном условии $n(x = -\infty) = n_p$ имеем $C_1 = 0$. Для $n(x = -x_p)$ из формулы (4.5.24) следует $n = n_{p_0} \exp(qV_a/kT)$, так что $C_2 = n_{p_0} [\exp(qV_a/kT) - 1] \times \exp[x_p/(D_n\tau_n)^{1/2}]$. Выражение (4.5.30) при этом принимает вид

$$n - n_{p_0} = n_{p_0} [\exp(qV_a/kT) - 1] \exp[(x_p + x)/L_n], \quad (4.5.31)$$

где длина диффузии неосновных носителей L_n определяется как

$$L_n = (D_n\tau_n)^{1/2}. \quad (4.5.32)$$

Аналогичные выражения могут быть получены для дырок с N -стороны гетероперехода.

Объединяя выражения (4.3.25) и (4.5.31), можно получить зависимость положения квазиуровня Ферми для электронов от расстояния. При прямом смещении инжектированные неосновные носители диффундируют в направлении от перехода, и их концентрация убывает экспоненциально, как это выражено формулой (4.5.31). Обозначая $n - n_{p_0}$ через n , из выражений (4.3.25) и (4.5.31) для $x \leq -x_p$ получаем

$$F_c - E_c = \frac{(x_p + x)kT}{L_n} + kT \ln \left[\frac{n_{p_0} \exp(qV_a/kT)}{N_c} \right]. \quad (4.5.33)$$

Аналогично с N -стороны

$$E_v - F_v = -\frac{(x - x_N)kT}{L_p} + kT \ln \left[\frac{P_{N_0} \exp(qV_a/kT)}{N_v} \right]. \quad (4.5.34)$$

Для изображенного на рис. 4.5.2, а p - N -гетероперехода при $L_n = 5,0$ мкм и $L_p = 1,0$ мкм квазиуровни Ферми меняются с расстоянием, как показано на рис. 4.5.2, б.

Сила тока рассчитывается по формуле (4.3.50) в точке $x = -x_p$:

$$i_n = qD_n \frac{dn}{dx} \Big|_{x=-x_p} = -\frac{qD_n n_{p_0}}{L_n} [\exp(qV_a/kT) - 1]. \quad (4.5.35)$$

Это движение электронов в отрицательном направлении оси x представляет собой ток в положительном направлении оси x . Аналогично для дырок с N -стороны

$$i_p = (qD_p P_{N_0}/L_p) [\exp(qV_a/kT) - 1]. \quad (4.5.36)$$

Влияние гетероперехода можно проиллюстрировать, взяв отношение $|i_n|/|i_p|$:

$$\left| \frac{i_n}{i_p} \right| = \frac{D_n n_p L_p}{D_p P_{N_0} L_n} = \frac{D_n L_p n_i^2 N_0}{D_p L_n N_i^2 p_0}, \quad (4.5.37)$$

где n_{p_0} и P_{N_0} заменены по формулам (4.5.9) и (4.5.10). Заметим, что для примера на рис. 4.5.2 с $x = 0,3$ имеем $n_i^2/N_i^2 = 7,4 \cdot 10^5$.

Собственные концентрации определяются формулой (4.3.43), в которой N_c и N_v выражаются формулами (4.3.24) и (4.3.39), так что

$$\left| \frac{i_n}{i_p} \right| = \frac{D_n L_p N_0 (m_{p,p} m_{n,p})^{1/2}}{D_p L_n p_0 (m_{p,N} m_{n,N})^{1/2}} \exp[(E_{g_2} - E_{g_1})/kT]. \quad (4.5.38)$$

Здесь E_{g_2} и E_{g_1} — запрещенные зоны с N - и p -сторон соответственно, $m_{p,p}$ и $m_{n,p}$ — эффективные массы дырок и электронов с p -стороны, $m_{p,N}$ и $m_{n,N}$ — эффективные массы дырок и электронов с N -стороны. Поскольку при комнатной температуре $1/kT = 39 \text{ эВ}^{-1}$, экспоненциальный множитель в формуле (4.5.38) является доминирующим даже при малых различиях ширины запрещенной зоны, поэтому $|i_n| \gg |i_p|$. Таким образом, диффузионный ток обусловлен главным образом инжекцией основных носителей из полупроводника с широкой запрещенной зоной в более узкозонный полупроводник. Обычно ток записывается в виде

$$I = I_0 \exp(qV_a/kT). \quad (4.5.39)$$

Поскольку $|i_n| \gg |i_p|$, то $I_0 = (qD_n n_{p0}/L_n) a$ для площади поперечного сечения a . Для n — P -гетероперехода с $D_n/L_n > D_p/L_p$ неравенство $|i_p| \gg |i_n|$ справедливо тогда, когда E_{g_2} больше E_{g_1} по крайней мере на несколько kT . В этом случае для n — P -гетероперехода $I_0 = (qD_p p_{n0}/L_p) a$.

При сильном смещении, таком, что $n = p$ при $x = -x_p$, формула (4.5.24) дает

$$n_p = n_i \exp(qV_a/2kT) \quad (4.5.40)$$

и

$$I = I_0 \exp(qV_a/2kT), \quad (4.5.41)$$

где $I_0 = (qD_n n_i/L_n) a$. В этом случае ток меняется с напряжением как $\exp(qV_a/2kT)$ по сравнению с $\exp(qV_a/kT)$ для случая слабого возбуждения. Детальный анализ случая сильного возбуждения значительно более сложен, чем данный здесь, но тем не менее выражения (4.5.39) и (4.5.41) показывают, что при переходе к высокому уровню накачки ожидается изменение наклона вольт-амперной характеристики. В реальных приборах токи безызлучательной рекомбинации (обусловленные поверхностной рекомбинацией [67]), которые зависят от напряжения как $\exp(qV_a/2kT)$, доминируют при слабом возбуждении, а диффузионные токи с такой же зависимостью $\exp(qV_a/2kT)$ играют главную роль вблизи порога генерации. Зависимости тока от напряжения типа $\exp(qV_a/kT)$ и $\exp(qV_a/2kT)$ иллюстрируются на рис. 4.5.3. Как подчеркивалось выше, именно диффузионный ток приводит к излучательной рекомбинации.

В модели Андерсона выражение, определяющее вольт-амперную характеристику, записывается в несколько иной форме, чем

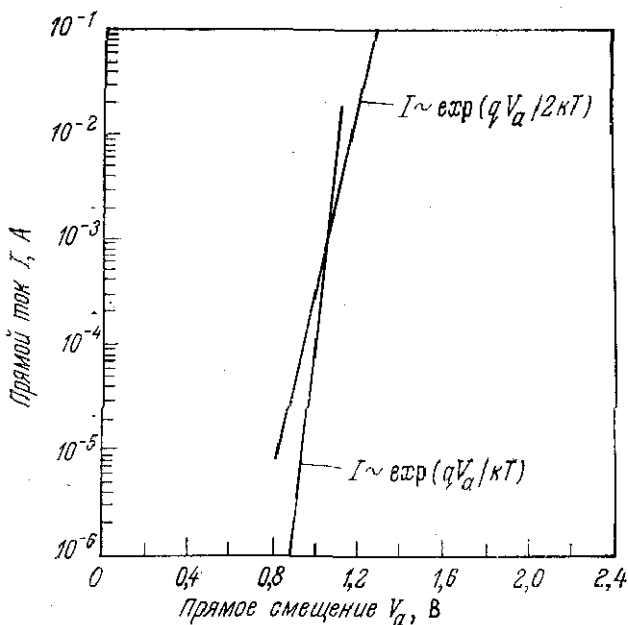


Рис. 4.5.3. Зависимости прямого тока от напряжения, описываемые выражениями $\exp(qV_a/kT)$ и $\exp(qV_a/2kT)$.

приведенное выше выражение (4.5.39), обычно используемое при описании гетеропереходов [6, 7, 10, 11]. На энергетической зонной диаграмме, изображенной на рис. 4.3.8, при тепловом равновесии барьер для электронов, движущихся из N - в p -сторону, равен V_{DN} в то время как для электронов, движущихся из p - в N -сторону, барьер равен $\Delta E_c - V_{Dp}$. Поскольку общий ток при тепловом равновесии равен нулю, потоки электронов в обоих направлениях должны быть равны друг другу:

$$B_1 \exp[-q(\Delta E_c - V_{Dp})/kT] = B_2 \exp[-qV_{DN}/kT], \quad (4.5.42)$$

где коэффициенты B_1 и B_2 зависят от концентраций примесей, эффективных масс носителей и механизма протекания тока.

Когда на $p-N$ -гетеропереход накладывается прямое смещение V_a , падение напряжения на p -стороне составляет V_1 , а на N -стороне — V_2 . При K , определенном выражением (4.3.116), V_{Dp} равно V_D/K , и, как следует из формулы (4.3.108), $V_1 = V_a/K$. Энергетические барьеры в случае прямого смещения становятся равными $V_{DN} - V_2$ и $\Delta E_c - (V_{Dp} - V_1)$. Суммарный

электронный поток из N - в p -сторону определяется выражением

$$\begin{aligned} \text{электронный поток} = & B_2 \exp[-q(V_{D_N} - V_2)/kT] - \\ & - B_1 \exp\{-q[(\Delta E_c - (V_{D_p} - V_1))/kT]\}. \end{aligned} \quad (4.5.43)$$

Коэффициент B_1 можно выразить из (4.5.42) через B_2 , и тогда выражение (4.5.43) может быть записано в виде

электронный поток =

$$= B_2 \exp[-qV_{D_N}/kT] [\exp(qV_2/kT) - \exp(-qV_1/kT)]. \quad (4.5.44)$$

Электрический ток записывается в форме, аналогичной форме записи для потока носителей. Если считать, что ток ограничивается скоростью диффузии электронов в p -слое [6, 7, 10, 11], то

$$I = q \left(\frac{XD_n N}{L_n} \right) a \exp(-qV_{D_N}/kT) [\exp(qV_2/kT) - \exp(-qV_1/kT)], \quad (4.5.45)$$

где X — коэффициент пропускания электронов через границу раздела. Остается продемонстрировать, насколько хорошее согласие с экспериментальными вольт-амперными характеристиками для $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$ $p - N$ -гетероперехода дает выражение (4.5.45). Аналогичное выражение получено и для $n - P$ -гетероперехода [6, 7, 10]. Для изотипных гетеропереходов механизм проводимости определяется термоэлектронной эмиссией; он обсуждается детально в гл. 4 книги [10].

§ 6. ОГРАНИЧЕНИЕ ДЛЯ НОСИТЕЛЕЙ ТОКА

Потенциальные барьеры, создающие ограничение для электронов и дырок

Одним из основных свойств гетероструктур является существующее в них ограничение для инжектированных неосновных носителей. Зависимость ограничения от молярной доли AlAs x в $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$ ДГС-лазерах была рассмотрена Роудом [32], в то время как Гудвин и др. [68] исследовали температурную зависимость утечки электронов из активной области в пассивный P -слой. Для того чтобы проиллюстрировать ограничение для носителей, необходимо установить характерные значения концентраций носителей в активном слое. Как показано на рис. 3.7.7, усиление в GaAs возникает при концентрациях инжектированных носителей, превышающих $\sim 1 \cdot 10^{18} \text{ см}^{-3}$. Хуань и Даймент [69] установили, что в полосковых лазерах для получения генерации необходима концентрация инжектированных электронов примерно $2 \cdot 10^{18} \text{ см}^{-3}$. Поэтому при рассмотрении $N - p - P$

двусторонних гетероструктур мы будем рассматривать в качестве характерной концентрации инжектированных неосновных носителей значение $2 \cdot 10^{18} \text{ см}^{-3}$.

Концентрация электронов в прямом минимуме зоны проводимости была определена выражением (4.3.16) как интеграл от плотности состояний

$$\rho_c(E - E_c) = (1/2\pi^2) (2m_n^\Gamma/\hbar^2)^{3/2} (E - E_c)^{1/2},$$

умноженной на функцию распределения Ферми $f_c = 1/\{\exp[(E - F_c)/kT] + 1\}$. Это произведение $\rho_c(E - E_c)f_c$ представляет собой распределение электронов $n(E)$ в прямом минимуме зоны проводимости. В GaAs при $T = 297 \text{ К}$ разность $F_c - E_c$ равна $0,079 \text{ эВ}$ для $n = 2 \cdot 10^{18} \text{ см}^{-3}$; получающееся при этом распределение $n(E)$ показано на рис. 4.6.1. При молярной доле AlAs $0,3$ разность между краями зон проводимости в активном p -слое и пассивном P -слое, определяемая формулой (4.3.5), составляет $\Delta E_c = 0,318 \text{ эВ}$. Это значение ΔE_c отмечено на рис. 4.6.1 штриховой линией, а область под кривой $n(E)$, соответствующая энергиям, превышающим ΔE_c , заштрихована. Концентрация носителей с энергиями, превышающими ΔE_c , обозначается через n^Γ , в принятых выше обозначениях она равна

$$n^\Gamma = \frac{1}{2\pi^2} \left(\frac{2m_n^\Gamma}{\hbar^2} \right)^{3/2} \int_{E_c + \Delta E_c}^{\infty} \frac{(E - E_c)^{1/2} dE}{1 + \exp[(E - F_c)/kT]}. \quad (4.6.1)$$

Для рассматриваемого примера $n^\Gamma = 1,6 \cdot 10^{14} \text{ см}^{-3}$. Второй интеграл в выражении (4.3.16) определяет концентрацию электро-

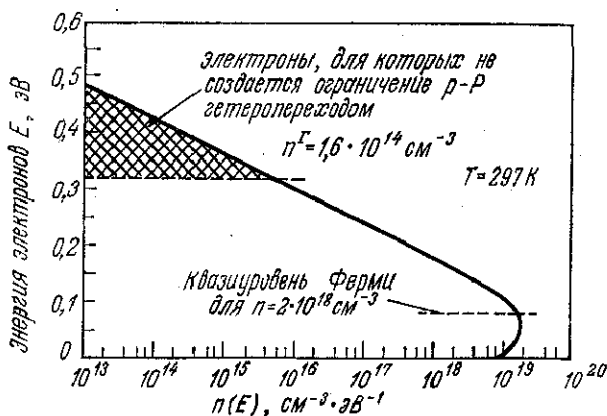


Рис. 4.6.1. Распределение электронов в прямом минимуме зоны проводимости GaAs, показывающее, какая часть электронов не ограничивается барьером ΔE_c , который при $x = 0,3$ равен $0,318 \text{ эВ}$.

нов в L -минимуме зоны проводимости, которая при $n = 2 \cdot 10^{18} \text{ см}^{-3}$ и $F_c - E_c = 0,079 \text{ эВ}$ равна $n_L = 3,4 \cdot 10^{15} \text{ см}^{-3}$. Аналогичные вычисления для X -минимума зоны проводимости дают $n_X = 1,5 \cdot 10^{12} \text{ см}^{-3}$. В L -минимуме концентрация электронов с энергиями, превышающими ΔE_c , составляет $\sim 1,5 \cdot 10^{15} \text{ см}^{-3}$, а в X -минимуме все электроны имеют энергию больше ΔE_c . Эти электроны могут «утекать» из активного слоя и приводить к появлению составляющей тока, которая не дает вклада в вынужденное излучение. Ток утечки мы рассмотрим ниже в этом параграфе.

Рассмотрим ограничение для дырок, создаваемое на $p-N$ -гетерогранице. На рис. 4.3.16, б показана $N-p-P$ двусторонняя гетероструктура при большом прямом смещении. Из этого рисунка видно, что потенциальный барьер для дырок в активном p -слое представляет собой сумму величин $V_D - V_a$ и ΔE_v . Поэтому в $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ -гетеропереходах, в которых ΔE_v мало, ограничение для дырок определяется главным образом разностью между контактной разностью потенциалов и приложенным смещением. Для увеличения концентрации электронов, инжектированных в активный слой, V_a необходимо увеличивать. Однако если V_a становится слишком большим, потенциальный барьер для дырок уменьшается. В $p-N$ -гетеропереходе с $N_{A1}^- - N_{D1}^+ = 1,1 \cdot 10^{18} \text{ см}^{-3}$ и $N_{D2}^+ - N_{A2}^- = 3 \cdot 10^{17} \text{ см}^{-3}$, изображенном на рис. 4.3.11, инжектирование в p -слой неосновных носителей до концентрации $n = 2 \cdot 10^{18} \text{ см}^{-3}$ требует инжектирования и основных носителей для поддержания электронейтральности. Условие электронейтральности будет выполнено при $p = n + (N_{A1}^- - N_{D1}^+) = 3,1 \cdot 10^{18} \text{ см}^{-3}$; при этом $F_v - E_{v1} = 0,022 \text{ эВ}$. Необходимое для получения $n = 2 \cdot 10^{18} \text{ см}^{-3}$ и $F_c - E_c = 0,079 \text{ эВ}$ прямое смещение определяется из положения квазиуровня Ферми для электронов, как это иллюстрируется на рис. 4.6.2. Это прямое смещение равно

$$V_a = [(F_{c1} - E_{c1}) + (E_{c2} - F_{c2}) + qV_D - \Delta E_c] / q. \quad (4.6.2)$$

Концентрацию дырок p , определяемую выражением (4.3.29), в котором m_p берется из формулы (4.3.38), можно представить как интеграл от плотности состояний

$$\rho_v(E_v - E) = (1/2\pi^2)(2m_p/\hbar^2)^{3/2}(E_v - E)^{1/2},$$

умноженной на функцию распределения Ферми для дырок. Это произведение $\rho_v(E_v - E)f_v$ описывает распределение дырок $p(E)$ в валентной зоне. На рис. 4.6.3 показано это распределение $p(E)$ в GaAs при $T = 297 \text{ К}$ и $p = 3,1 \cdot 10^{18} \text{ см}^{-3}$, чему соответствует $F_v - E_v = 0,022 \text{ эВ}$. В этом случае квазиуровень Ферми находится внутри запрещенной зоны. Как и для зоны проводимости,

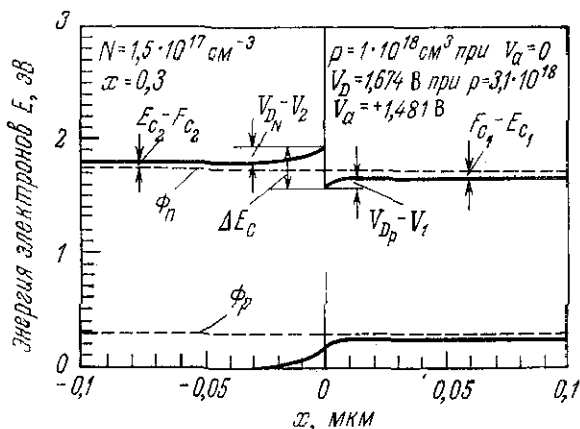


Рис. 4.6.2. Энергетическая зонная диаграмма GaAs — Al_xGa_{1-x}As *p* — *N*-гетероперехода при прямом смещении, дающем концентрацию инжектированных электронов $n = 2 \cdot 10^{18} \text{ см}^{-3}$ при $T = 297 \text{ К}$. Обозначения пояснены в тексте.

энергия потенциального барьера для дырок $(V_D - V_a) + \Delta E_v$ отмечена на рисунке штриховой линией, а область под кривой $\rho(E)$, отвечающая энергиям, меньшим энергии барьера, заштрихована. Концентрация дырок с энергиями, меньшими энергии барьера, обозначена через ρ_{Γ} и определяется выражением

$$\rho_{\Gamma} = \frac{1}{2\pi^2} \left(\frac{2m_p}{\hbar^2} \right)^{3/2} \int_{-\infty}^{(v_D - v_a + \Delta E_v)} \frac{(E_v - E)^{1/2} dE}{1 + \exp[(F_v - E)/kT]} \quad (4.6.3)$$

Для рассматриваемого примера $\rho_{\Gamma} = 7,1 \cdot 10^{14} \text{ см}^{-3}$. Как и электроны, эти дырки могут утекать из активного слоя в пассивный *N*-слой и приводить к возникновению составляющей тока, не дающей вклада в вынужденное излучение. Этот ток утечки также будет рассмотрен ниже в этом параграфе.

Гетеропереход, изображенный на рис. 4.6.2 при большом прямом смещении, демонстрирует еще одно уникальное свойство гетеропереходов. Электроны с *N*-стороны, где $N = 1,5 \cdot 10^{17} \text{ см}^{-3}$, инжектируются на *p*-сторону, где тем самым поддерживается концентрация $n = 2 \cdot 10^{18} \text{ см}^{-3}$. Это явление, при котором носители инжектируются из эмиттера до концентраций, больших, чем концентрация в самом эмиттере, было названо Алферовым «суперинжекцией» [70]. Ченг подробно рассмотрел этот эффект [71]. Суперинжекция возможна независимо от того, является ли пик в зоне проводимости резким или плавным. Высокие уровни инжекции должны влиять на энергетическую зонную диаграмму посредством таких эффектов, как, например, сужение запрещен-

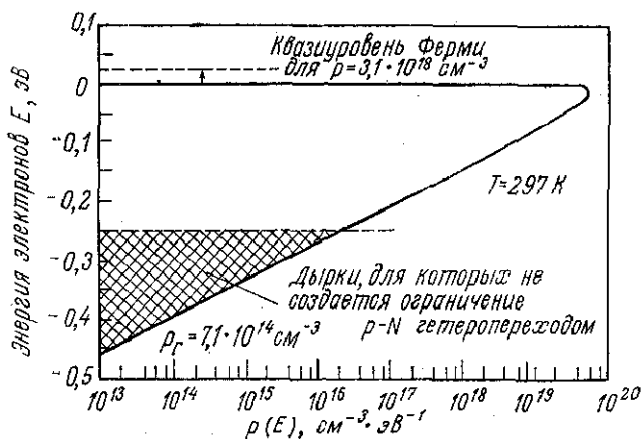


Рис. 4.6.3. Распределение дырок в валентной зоне GaAs при $x=0,3$, показывающее, какая часть дырок не ограничивается барьером $(V_D - V_a) + \Delta E_v$.

ной зоны, определяемое выражением (3.7.2). Однако ни о каких исследованиях влияния высоких уровней инжекции на энергетическую зонную диаграмму не сообщалось. Судя по характеристикам ДГС-лазеров при высоких уровнях инжекции, основные свойства гетеропереходов, по-видимому, сохраняются.

Распределение электронов между прямыми и непрямыми минимумами зоны проводимости

С возрастанием молярной доли AlAs x в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ от нуля энергетические зазоры между минимумами зоны проводимости уменьшаются. При x вблизи 0,45 происходит переход прямая — непрямая зона проводимости, и концентрации электронов в не-прямых минимумах зоны проводимости n_L и n_X становятся сравнимыми с концентрацией электронов в прямом минимуме зоны проводимости n_T . Поскольку электроны в непрямах минимумах не участвуют в процессе вынужденного излучения, генерация далее становится невозможной. В этой ситуации уровень Ферми определяется из выражения (4.3.16) с учетом членов для каждого минимума. Затем можно найти электронную заселенность каждого минимума. Часть γ общего числа электронов, которую составляют электроны, находящиеся в прямом минимуме зоны проводимости, будет равна

$$\gamma = n_T / (n_T + n_L + n_X). \quad (4.6.4)$$

Зависимость γ от x для $n = 2 \cdot 10^{18}$ см⁻³ показана на рис. 4.6.4 при $T = 24$ и 50°C . При значениях x , превышающих 0,30, более половины электронов находится в непрямах минимумах зоны проводимости.

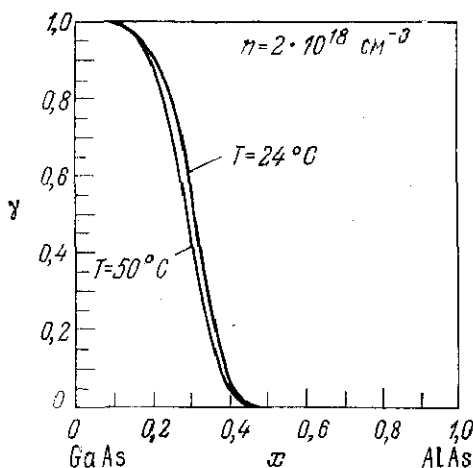


Рис. 4.6.4. Доля электронов, заполняющих, прямой минимум зоны проводимости в $\text{Al}_x\text{Ga}_{1-x}\text{As}$, как функция x при $T = 297$ К.

Ток утечки, обусловленный носителями, для которых не создается ограничение

Барьеры, создающие ограничение для электронов и дырок, рассмотрены в первой части этого параграфа. Максимальный ток утечки на границе $x = 0$ представляет собой тепловой поток электронов с энергиями, превышающими энергию потенциального барьера на гетеропереходе. Этот тепловой поток электронов на $p-P$ -границе определяется произведением тепловой скорости электронов на концентрацию создающих ток утечки электронов $n^T + n^L + n^X$, которая рассмотрена выше. Эта концентрация определяет ту часть электронов, инжектированных в активный слой $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$ $N-p-P$ двусторонней гетероструктуры, которая проходит через $p-P$ -гетеропереход и становится неосновными носителями в P -слое. В дальнейшем они рекомбинируют с дырками по мере диффузии по направлению к положительному контакту. Роуд [32] показал, что максимальным током утечки является тепловой поток носителей. Однако ток утечки ограничивается скоростью, с которой носители могут дрейфовать и диффундировать в P -слое. При полном анализе этого вопроса нужно было бы учесть как диффузионную, так и дрейфовую составляющие тока утечки при решении уравнений непрерывности (4.5.26) и (4.5.27). Мы будем здесь рассматривать только диффузионный ток с целью продемонстрировать влияние молярной доли AlAs и температуры на ток утечки. Дальнейшее обсуждение вопроса о токах утечки см. в работе [72].

Если наличие поля игнорируется или оно берется равным нулю, то уравнение непрерывности (4.5.28) имеет общее решение вида (4.5.30). Переписав выражение (4.5.30) для P -стороны, получим

$$N_P(x) = C_1 \exp(-x/L_N) + C_2 \exp(x/L_N) + N_P. \quad (4.6.5)$$

Постоянные интегрирования вновь определяются из граничных условий. На границе $x = x_P$ электронная концентрация должна быть равна N_P . Поскольку для GaAs—Al_xGa_{1-x}As p — P -гетероперехода $x_P \ll L_N$, величины $\exp(-x_P/L_N)$ и $\exp(x_P/L_N)$ могут быть положены равными единице. Другое условие состоит в том, что $N_P(x)$ обращается в нуль на контакте, расположенном на расстоянии w от плоскости $x = 0$. При этих граничных условиях выражение (4.6.5) принимает вид

$$N_P(x) = \frac{N_P \{ \exp[(w-x)/L_N] - \exp[-(w-x)/L_N] \}}{\exp(w/L_N) - \exp(-w/L_N)}. \quad (4.6.6)$$

Здесь член N_P опущен. Диффузионный ток электронов, как и в выражении (4.3.50), определяется величиной $qD_N dN/dx$, вычисленной при $x = x_P \approx 0$, и равен

$$i_N = - \frac{qD_N N_P}{L_N \operatorname{th}(w/L_N)}. \quad (4.6.7)$$

Если w/L_N лежит в пределах от $\sim 0,1$ до $\sim 0,5$, то $\operatorname{th}(w/L_N)$ приближенно равен w/L_N , и выражение (4.6.7) может быть записано в виде

$$i_N = - qD_N N_P / w. \quad (4.6.8)$$

Оценка этого выражения была проведена Гудвином и др. [68]. При значениях w вблизи 2 мкм и L_N вблизи 4 мкм выражение (4.6.8) дает разумное приближение.

Как и в случае гомоперехода, диффузионный ток вычисляется при $x = x_P$, и для используемой в этом выражении концентрации неосновных носителей N_P берется значение при $x = x_P$. Для определения N_P необходимо рассмотреть, как энергетическая зонная диаграмма p — P -гетероперехода меняется при сильной накачке двусторонней гетероструктуры N — p — P . Для рассмотренного на рис. 4.6.2 примера с $p = 3,1 \cdot 10^{18}$ см⁻³ и $P = 1,5 \cdot 10^{17}$ см⁻³ энергетическая зонная диаграмма в этом случае показана на рис. 4.6.5; она отличается от изображенного на рис. 4.3.14, б случая теплового равновесия. Контактная разность потенциалов, определенная формулой (4.3.122), изменяется от 0 при тепловом равновесии до $-0,031$ В в случае сильной накачки. При $(F_P - E_{v_1}) > \Delta E_v + (F_P - E_{v_1})$ удобнее записать V_D в виде (рис. 4.6.5)

$$V_D = [(F_P - E_{v_2}) - \Delta E_v - (F_P - E_{v_1})] / q. \quad (4.6.9)$$

В верхней части рис. 4.6.5 показано, что положение квазиуровня Ферми для электронов в P -слое можно определить из уравнения

$$(E_{c_2} - F_N) = \Delta E_c - qV_D - (F_n - E_{c_1}). \quad (4.6.10)$$

Отсюда с учетом выражения (4.6.9) получаем

$$(E_{c_2} - F_N) = \Delta E_g + (F_p - E_{v_1}) - (F_p - E_{v_2}) - (F_n - E_{c_1}). \quad (4.6.11)$$

Затем N_p находится по формуле (4.3.25), в которой N_c определяется выражением (4.3.24) для значений x , соответствующих области прямозонных составов, и выражением (4.3.36), когда x лежит в области непрямозонных составов. При этом для данного значения x используются соответствующие значения m_n^r , m_n^l и m_n^x .

Без знания величин ω или L_N трудно рассчитать характерные значения i_N . Было сообщено [73], что в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ L_N лежит в пределах от 3 до 6 мкм. Величина ω обычно составляет 1—2 мкм; таким образом, можно использовать формулу (4.6.8). В рассматриваемом здесь типичном примере величина ω взята равной 2 мкм. Коэффициент диффузии электронов вычисляется по формуле $D_N = \mu_N kT/q$ с $\mu_N = 10^3 \text{ см}^2/(\text{В}\cdot\text{с})$, что дает $D_N = 25,6 \text{ см}^2/\text{с}$. При этих значениях параметров $qD_N/\omega = 2 \cdot 10^{-14} \text{ см}/\text{с}$. На рис. 4.6.6 показан рассчитанный по формуле

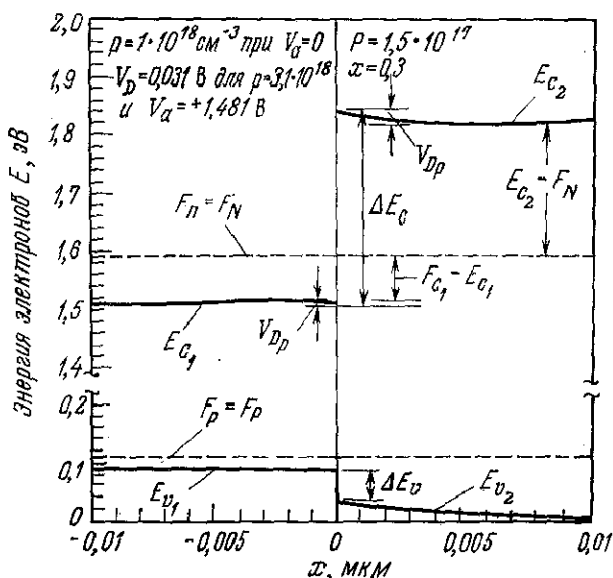


Рис. 4.6.5. Энергетическая зонная диаграмма $p-p$ -гетероперехода в $n-p-p$ двусторонней гетероструктуре при высоком уровне накачки. Эта диаграмма иллюстрирует задание положения квазиуровня Ферми для электронов в P -слое.

(4.6.8) электронный ток утечки для $n = 2 \cdot 10^{18} \text{ см}^{-3}$ при $T = 24$ и 50°C и для $n = 3 \cdot 10^{18} \text{ см}^{-3}$ при $T = 50^\circ\text{C}$. Величина N_P вычислялась описанным выше способом. При повышении температуры усиление уменьшается и для увеличения n необходимо усиливать инжекцию. На рис. 4.6.7 измеренная Роудом [32] экспериментальная зависимость пороговой плотности тока $J_{\text{пор}}$ от x сравнивается с расчетной $J_{\text{пор}}$, полученной прибавлением тока утечки к значению $J_{\text{пор}}$ для $x > 0,3$. Ток утечки определялся по кривой на рис. 4.6.6, соответствующей $n = 2 \cdot 10^{18} \text{ см}^{-3}$ и $T = 24^\circ\text{C}$. Расхождение представленных результатов лежит в пределах, объяснимых неточностью в оценке тока утечки. Такое согласие результатов говорит о том, что приближение, используемое при оценке N_P , и представление i_N выражениями (4.6.7) и (4.6.8) разумны. Следует отметить, что в случае, когда $L_N \ll \omega$, $\text{th}(\omega/L_N)$ стремится к единице и

$$i_N = qD_N N_P / L_N \quad \text{при } L_N \ll \omega. \quad (4.6.12)$$

При малых L_N значение тока утечки может быть много большим, чем те значения, которые приведены на рис. 4.6.6.

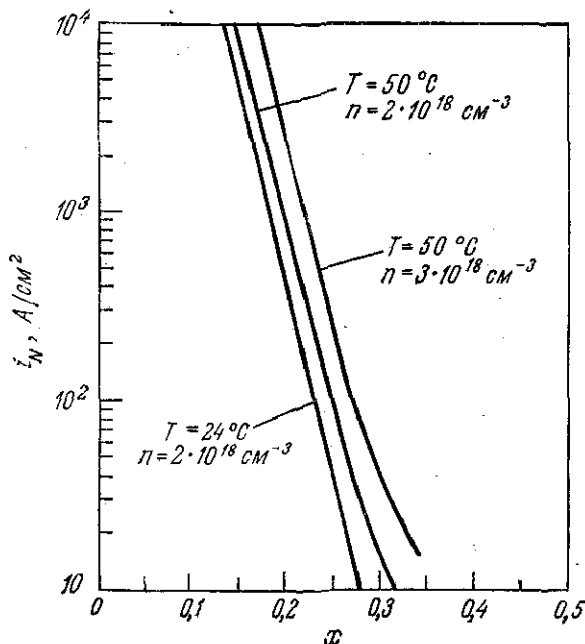


Рис. 4.6.6. Зависимость тока утечки электронов от молярной доли AlAs x в $P - \text{Al}_x\text{Ga}_{1-x}\text{As}$ -слое $N - p - P$ двусторонней гетероструктуры при указанных значениях n в активном p -слое при $T = 24$ и 50°C .

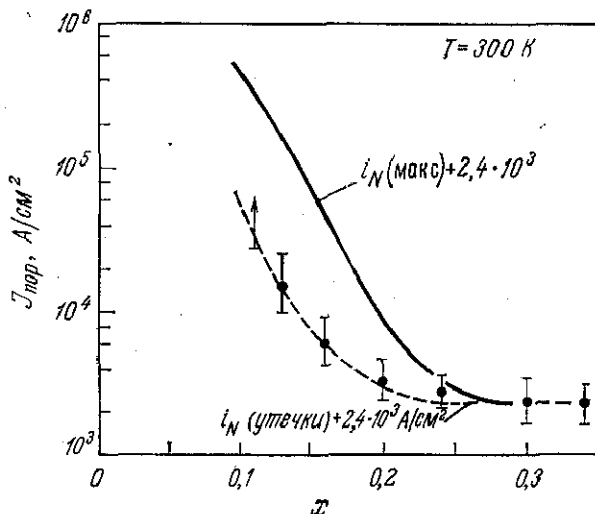


Рис. 4.6.7. Сравнение экспериментальной зависимости $J_{\text{пор}}$ от x [32] с расчетной зависимостью от x тока утечки. Для $x \geq 0,3$ величина $J_{\text{пор}} = 2,4 \cdot 10^3 \text{ А/см}^2$ и представляет собой $J_{\text{пор}}$ при пренебрежимых токах утечки. Сплошная линия в области токов, больших $2,4 \cdot 10^3 \text{ А/см}^2$, соответствует максимальному току утечки $i_N(\text{макс})$, являющемуся тепловым потоком электронов с концентрацией $n^{\Gamma} + n^X$, для которых не создается ограничение [32]. Штриховая кривая, лежащая выше значения тока $2,4 \cdot 10^3 \text{ А/см}^2$, представляет собой диффузионный ток утечки i_N (утечки) и взята с рис. 4.6.6 для $n = 2 \cdot 10^{18} \text{ см}^{-3}$ при $T = 24^\circ \text{C}$.

Выражение, аналогичное выражению (4.6.7), может быть получено для дырок, инжектированных из активного p -слоя в N -слой:

$$i_p = \frac{qD_p P_N}{L_p \text{th}(\omega/L_p)}. \quad (4.6.13)$$

Поскольку в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ $\mu_p \approx \mu_n/10$, то $L_p < L_n$, и тогда $\omega \geq L_p$. В этих условиях $\text{th}(\omega/L_p) \approx 1$, и дырочный ток утечки определяется выражением

$$i_p = qD_p P_N / L_p, \quad (4.6.14)$$

где P_N определяется по формуле (4.3.40) с соответствующим данному x значению массы m_p . Как показано на рис. 4.6.2, величина $F_p - E_{v_2}$ зависит от приложенного смещения и равна

$$F_p - E_{v_2} = q(V_D - V_a) + \Delta E_v + (F_p - E_{v_1}). \quad (4.6.15)$$

Это выражение с учетом формулы (4.3.8), определяющей V_D , принимает вид

$$F_p - E_{v_2} = E_{g_2} - qV_a - (E_{c_2} - F_{c_2}), \quad (4.6.16)$$

где V_a задается выражением (4.6.2). Для примера, приведенного на рис. 4.6.2, $F_p - E_{v_2} = 0,272$ эВ и $P_N = 2,6 \cdot 10^{14}$ см⁻³. Если взять $D_p = 3$ см²/с и $L_p = 1$ мкм, то по формуле (4.6.14) получим $i_p = 1,3$ А/см². При уменьшении молярной доли AlAs x в двусторонней гетероструктуре, симметричной по x , ток i_p остается меньшим, чем ток i_N .

§ 7. ЗАКЛЮЧЕНИЕ

В этой главе приведены выражения, описывающие энергетическую зонную диаграмму гетероперехода, и рассмотрено поведение гетероперехода при приложении внешнего напряжения смещения. Хотя еще неясно, насколько хорошо андерсоновская модель резкого гетероперехода описывает реальный гетеропереход она предсказывает его ограничивающие и инжекционные свойства, существенные для работы гетеролазеров. Уникальным свойством гетеропереходов является возможность работы в режиме так называемой суперинжекции, при котором концентрация основных носителей в широкозонном эмиттере меньше, чем концентрация инжектированных носителей в активном слое.

В этой главе рассматривались только GaAs—Al_xGa_{1-x}As-гетеропереходы, так как они наиболее широко используются в полупроводниковых лазерах и для них имеются обширные экспериментальные данные относительно свойств, представляющих интерес с точки зрения применения в гетеролазерах. Приведены многочисленные примеры, основанные на имеющихся данных для Al_xGa_{1-x}As и иллюстрирующие свойства гетеропереходов в соединениях типа A^{III}B^V без учета состояний на границе раздела. Можно ожидать, что численные значения величин будут уточняться по мере поступления новых данных для Al_xGa_{1-x}As. Кроме того, другие системы на соединениях типа A^{III}B^V начинают иметь важное значение для применения в лазерах, и приведенный здесь анализ можно использовать для этих систем при соответствующих значениях параметров. В следующей главе рассматриваются другие полупроводниковые материалы, используемые в гетеролазерах, и приводятся энергетические зонные диаграммы гетеропереходов на некоторых других системах.

ЛИТЕРАТУРА

1. Губанов А. И., ЖТФ, 21, 304 (1951).
2. Губанов А. И., ЖЭТФ, 21, 721 (1951).
3. Губанов А. И., ЖТФ, 22, 729 (1952).
4. Shockley W., US Patent 2 569 347, 1951.
5. Kroemer H., Proc. IRE, 45, 1535 (1957).

6. *Anderson R. L.*, IBM Journ. Res. Dev., 4, 283 (1960).
7. *Anderson R. L.*, Solid-State Electron., 5, 341 (1962).
8. *Kroemer H.*, Crit. Rev. Solid State Sci., 5, 555 (1975).
9. *Frenslley W. R., Kroemer H.*, Journ. Vac. Sci. Tech., 13, 810 (1976).
10. *Milnes A. G., Feucht D. L.*, Heterojunctions and Metal-Semiconductor Junctions, Academic Press, New York, 1972. (Имеется перевод: *Милнс А., Фойхт Д.* Гетеропереходы и переходы металл — полупроводник. — М.: Мир, 1975.)
11. *Sharma B. L., Purohit R. K.*, Semiconductor Heterojunctions, Pergamon Press, Oxford, 1974. (Имеется перевод: *Шарма Б. Л., Пурохит Р. К.* Полупроводниковые гетеропереходы. — М.: Сов. радио, 1979.)
12. Proc. Conf. Phys. Chem. Semiconductor Heterojunctions and Layered Structures, Vol. 1—5, Akadémiai Kiadó, Budapest, 1971.
13. *Long D.*, Energy Bands in Semiconductors, Wiley (Interscience), New York, 1968.
14. *Harrison W. A.*, Solid State Theory, McGraw-Hill, New York, 1970. (Имеется перевод: *Харрисон У.* Теория твердого тела. — М.: Мир, 1972.)
15. *Pollak F. H., Higginbtham C. W., Cardona M.*, Proc. Intl. Conf. Phys. Semiconductors, Kyoto, 1966, p. 20 (Journ. Phys. Soc., Jpn., 21, Supplement, 1966).
16. *Sell D. D., Casey H. C., Jr., Wecht K. W.*, Journ. Appl. Phys., 45, 2650 (1974).
17. *Thurmond C. D.*, Journ. Electrochem. Soc., 122, 1133 (1975).
18. *Aspnes D. E.*, Phys. Rev., B14, 5331 (1976).
19. *Aukerman L. W., Willardson R. K.*, Journ. Appl. Phys., 31, 939 (1960).
20. *Smith R. A.*, Semiconductors, Cambridge Univ. Press, London, 1961, p. 40. (Имеется перевод 1-го издания: *Смит Р.* Полупроводники. — М.: ИЛ, 1962.)
21. *Blakemore J. S.*, Semiconductor Statistics, Pergamon, New York, 1962, p. 40. (Имеется перевод: *Блэкмор Дж.* Статистика электронов в полупроводниках. — М.: Мир, 1964.)
22. *Berolo O., Wooley J. C.*, Proc. 11th, Intern. Conf. Phys. Semiconductors, Vol. 2, Polish Scientific, Warsaw, 1972, p. 1420.
23. *Lawaetz P.*, Phys. Rev., 4, 3460 (1971).
24. *Vrehen Q. H. F.*, Journ. Phys. Chem Solids, 29, 129 (1968).
25. *Mears A. L., Stradling R. A.*, Journ. Phys., C4, L22 (1971).
26. *Pitt G. D., Lees J.*, Phys. Rev., B2, 4144 (1970).
27. *Pitt G. D., Lees J.*, Solid State Comm., 8, 491 (1970).
28. *Onton A.*, Proc. 10th Intern. Conf. Phys. Semiconductors, Cambridge, Mass., 1970, p. 107 (US Atomic Energy Commission, Springfield, Virginia, 1970).
29. *Monemar B.*, Phys. Rev., B8, 5711 (1973).
30. *Dingle R., Logan R. A., Arthur J. R.*, частное сообщение.
31. *Stukel D. J., Ewens R. N.*, Phys. Rev., 188, 1193 (1969).
32. *Rode D. L.*, Journ Appl. Phys., 45, 3887 (1974).
33. *Aspnes D. E.*, частное сообщение.
34. *Blakemore J. S.*, Semiconductor Statistics, Pergamon, New York, 1962, p. 76. (Имеется перевод: *Блэкмор Дж.* Статистика электронов в полупроводниках. — М.: Мир, 1964.)
35. *Casey H. C., Jr., Panish M. B.*, Journ. Appl. Phys., 40, 4910 (1969).
36. *Berolo O., Wooley J. C.*, Can. Journ. Phys., 49, 1335 (1971).
37. *Onton A., Lorenz M. R., Woodall J. M.*, Bull. Amer. Phys. Soc., 16, 371 (1971).
38. *Monemar B., Shihk K., Pettit G. D.*, Journ. Appl. Phys., 47, 2604 (1976).
39. *Dingle R.*, Festkörper — Probleme XV — Advances in Solid State Physics, Pergamon-Vieweg, 1975, 21.
40. *Hwang C. J., Brews J. R.*, Journ. Phys. Chem. Solids, 32, 837 (1971).
41. *Casey H. C., Jr., Stern F.*, Journ. Appl. Phys., 47, 631 (1976).

42. *Panish M. B.*, Journ. Appl. Phys., **44**, 2667 (1973).
43. *Spring Thorpe A. J., King F. D., Becke A.*, Journ. Electron. Mat., **4**, 101 (1975).
44. *Casey H. C., Jr. et al.*, Bull. Am. Phys. Soc., **22**, 317 (1977).
45. *Blakemore J. S.*, Semiconductor Statistics, Pergamon, New York, 1962, p. 79. (Имеется перевод: Блэкмор Дж. Статистика электронов в полупроводниках. — М.: Мир, 1964.)
46. *Blakemore J. S.*, Semiconductor Statistics, Pergamon, New York, 1962, p. 346. (Имеется перевод: Блэкмор Дж. Статистика электронов в полупроводниках. — М.: Мир, 1964.)
47. *Joyce W. B., Dixon R. W.*, Appl. Phys. Lett., **31**, 354 (1977).
48. *Darke L. S., Gurry R. W.*, Physical Chemistry of Metals, McGraw-Hill, New York, 1953, p. 189.
49. *Swalin R. A.*, Thermodynamics of Solids (2nd ed.), Wiley, New York, 1974, p. 296.
50. *Strzalkowski I., Joshi S., Crowell C. R.*, Appl. Phys. Lett., **28**, 350 (1976).
51. *Fern R. E., Onton A.*, Journ. Appl. Phys., **42**, 3499 (1971).
52. *McKelvey I. P.*, Solid State and Semiconductor Physics, Harper, New York, 1966, p. 390.
53. *Sze S. M.*, Physics of Semiconductor Devices, Wiley (Interscience), New York, 1969. (Имеется перевод: Су С. М. Физика полупроводниковых приборов. — М.: Энергия, 1973.)
54. *McKelvey I. P.*, Solid State and Semiconductor Physics, Harper, New York, 1966, p. 199.
55. *Van der Pauw L. J.*, Philips Res. Rep., **13**, 1 (1958).
56. *Sze S. M.*, Physics of Semiconductor Devices, Wiley (Interscience), New York, 1969, p. 363. (Имеется перевод: Су С. М. Физика полупроводниковых приборов. — М.: Энергия, 1973.)
57. *Ermanis F., Wolfstirn K.*, Journ. Appl. Phys., **37**, 1963 (1966).
58. *Womac J. F., Rediker R. H.*, Journ. Appl. Phys., **43**, 4129 (1972).
59. *Casey H. C., Jr., Cho A. Y.*, не опубликовано.
60. *Cheung D. T., Chiang S. Y., Pearson G. L.*, Solid-State Electron., **18**, 263 (1975).
61. *Oldham W. G., Milnes A. G.*, Solid-State Electron., **6**, 121 (1963).
62. *Donnelly J. P., Milnes A. G.*, IEEE Trans. Electron. Devices, ED-14, 63 (1967).
63. *Добрынин С. Н.*, ФТП, **6**, 1001 (1972).
64. *Howarth D. S., Feucht D. L.*, Appl. Phys. Lett., **23**, 365 (1973).
65. *Sze S. M.*, Physics of Semiconductor Devices, Wiley (Interscience), New York, 1963, p. 96. (Имеется перевод: Су С. М. Физика полупроводниковых приборов. — М.: Энергия, 1973.)
66. *McKelvey I. P.*, Solid State and Semiconductor Physics, Harper, New York, 1966, Chapter 10.
67. *Henry C. H., Logan R. A., Merritt F. R.*, Appl. Phys. Lett., **31**, 454 (1977).
68. *Goodwin A. R. et al.*, Journ. Appl. Phys., **46**, 3126 (1975).
69. *Hwang C. J., Dymert J. C.*, Journ. Appl. Phys., **44**, 3240 (1973).
70. *Алферов Ж. И., Халфин В. Б., Казаринов Р. Ф.*, ФТП, **8**, 3102 (1966).
71. *Cheung D. T., Pearson G. L.*, Journ. Appl. Phys., **46**, 2313 (1975).
72. *Casey H. C., Jr.*, Journ. Appl. Phys. (May 1978).
73. *Kawakami T., Sugiyama K.*, Jpn. Journ. Appl. Phys., **12**, 151 (1973).
74. *Summers C. J., Dingle R., Hill D. E.*, Phys. Rev., **B1**, 1603 (1970).
75. *Fetterman H. R. et al.*, Appl. Phys. Lett., **21**, 434 (1972).
76. *Williams E. W.*, Phys. Rev., **168**, 922 (1968).
77. *Haisty R. W., Mehal E. W., Stratton R.*, Journ. Phys. Chem. Solids, **23**, 829 (1962).
78. *Stadek R. J.*, Phys. Rev., **140**, A1345 (1965).
79. *Ashen D. J. et al.*, Journ. Phys. Chem. Solids, **36**, 1041 (1975).

80. Schairer W., Grobe E., Solid State Commun., 8, 2017 (1970).
81. Hegems M., Dingle R., GaAs and Related Compounds, 1974 Symp. Proc., Inst. Phys., London, 1975, p. 1.
82. Rosi F. D., Meyerhofer D., Jensen R. V., Journ. Appl. Phys., 31, 1105 (1960).
83. Baxter R. D., Reid F. J., Journ. Electrochem. Soc., 110, 187C (1963).
84. Haisty R. W., Cronin G. R., Proc. 7th Intern. Conf. Phys. Semiconductors, Dunod, Paris, 1964, p. 1161.
85. Шишняну Ф. С., Болтакс Б. И., ФТТ, 8, 1312 (1966).
86. Blatte M., Schairer W., Willmann F., Solid State Commun., 8, 1265 (1970).

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- $AlAs$, энергетические зазоры 222, 223
 — эффективные массы носителей 222, 223
 $Al_xGa_{1-x}As$, диэлектрическая проницаемость 243
 — плотность состояний 233
 — показатель преломления 57—60
 — энергетические зазоры 222, 224, 225
 — энергия ионизации примесей 231
 — эффективная плотность состояний 235, 237
 — эффективные массы носителей 225
 Андерсона модель 218, 246
- Блоха волны 120
 — функция 170
 Бозе — Эйнштейна закон распределения 138
 Брэгга условие 113
 Бурштейна сдвиг 185
- Вероятность перехода 139, 140, 149, 150
 Волновая функция 145
 — — в параболической зоне 172
 — — локализованного состояния 170—172
 Волновод асимметричный трехслойный 84—87
 — — — выражения для электрического поля 85
 — — — дисперсионное уравнение 80
 — — — коэффициент оптического ограничения 87
 — — — распределение интенсивности 86
 — — — условие отсечки 80, 81
 — симметричный трехслойный, графическое решение дисперсионного уравнения 62—64
- Волновод симметричный трехслойный, дисперсионное уравнение для ТЕ-мод 54, 55
 — — — — — ТМ-мод 55, 56
 — — — коэффициент оптического ограничения 67—74
 — — — распределение интенсивности света 70
 — — — решения для ТЕ-мод 53—55
 — — — — — ТМ-мод 55, 56
 — — — условие отсечки 64
 — — — эффективный показатель преломления 65
 Волновое уравнение для плоских волн в среде без потерь 37, 38
 — — — — — с потерями 41
 — — — трехмерное 48, 49
 Время жизни безызлучательное 188
 — — излучательное 185—189
 Вынужденное излучение 132
 — — необходимое условие 12, 132, 141
 — — скорость 141, 142, 154
 — — — связь с коэффициентом поглощения 143
 — — — со скоростью спонтанного излучения 144
- $GaAs$, зонная структура 220—222
 — излучательное время жизни 186, 187
 — коэффициент поглощения 61, 62
 — показатель преломления 57—59
 — потери на свободных носителях 204
 — энергии ионизации примесей 230
 — эффективные массы носителей 222, 223
 Гетеропереходы 17, 23, 24, 218
 — анизотипные 218
 — вольт-амперные характеристики 271—280

- Гетеропереходы, вольт-емкостная зависимость 270, 271
 — изгиб зон 246
 — изотипные 218
 — контактная разность потенциалов 240—246
 — модель Андерсона 218, 246
 — разрывы зон 226, 227
 — толщина области пространственного заряда 270
 Граничные условия для электрического и магнитного полей 51, 52
 — — периодические 135, 136
 Гуса — Хэнхена сдвиг 81—83
 Гюйгенса угловой фактор 94
- ДГС-лазеры (лазеры на двусторонней гетероструктуре) 20, 26
 — картина излучения в дальней зоне 89—95
 — коэффициент оптического ограничения 67—74, 88
 — — отражения на торцевых гранях 98—102
 — расходимость луча 94—98
 — сдвиг Гуса — Хэнхена 81—83
 — энергетические зонные диаграммы 263—265
 ДГС — РО-лазеры (ДГС-лазеры с раздельным ограничением) 102—107
 — — дисперсионное уравнение 105
 — — коэффициент оптического ограничения 107
 — — распределение интенсивности 106
 Дебая длина экранирования 159
 Дипольный матричный элемент 150
 Дисперсионное уравнение 54—56
 — — в модели зигзагообразных волн 78—80
 — — графическое решение 62—65
 Дифференциальная квантовая эффективность 208, 210
 Диффузионный ток дырок 240
 — — электронов 240
 Диэлектрическая проницаемость 36
 — — комплексная 42
 — — связь с оптическими константами 42
 Длина диффузии 188
- Запрещенная зона, зависимость от концентрации 184, 191
 — — — состава в четверных твердых растворах соединений $A^{III}B^V$ 29, 30
- Излучение черного тела 135
 — — — спектральная плотность 138
- Картина излучения в дальней зоне 89—97
 Контактная разность потенциалов 240—246
 — — — n — N -гетеропереход 261
 — — — n — P -гетеропереход 256
 — — — p — N -гетеропереход 247
 — — — p — P -гетеропереход 259
 Концентрация дырок в валентной зоне 235, 237, 238
 — носителей собственная 187, 238
 — электронов в зоне проводимости 233—235, 284
 к-отбора правило 133, 170
 Коэффициент активности дырок 186
 — — электронов 186
 — диффузии 188, 240
 — оптического ограничения 23, 67
 — — — двусторонней гетероструктуры асимметричной 88
 — — — — с раздельным ограничением 107
 — — — — симметричной 68, 70
 — — — — односторонней гетероструктуры 88
 — — — приближенное выражение 73
 — отражения комплексный 78
 — поглощения 43, 153
 — — связь с суммарной скоростью поглощения 143
 — — — со скоростью вынужденного излучения 143
 — — — — спонтанного излучения 143
 — — сравнение расчетных и экспериментальных значений 177—185
 — усиления, зависимость от номинальной плотности тока 195—202
 — — расчет 192, 195—202
 — — температурная зависимость 202, 203
- Кристаллические твердые растворы 28
 Критический угол 21, 77
- Лазеры с РВО (распределенным брэгговским отражателем) 112
 — — РВ (расширенным волноводом) 107—111
 — — РОС (распределенной обратной связью) 111
 — — — метод блоховских волн 120—121

- Лазеры с РОС, периодическое изменение показателя преломления в гетероструктурах 122—128
 — — — с раздельным ограничением 128
 — — — теория связанных волн 112—120
 — — — — — мод 121—122
- Метод стационарной фазы 91
- n — N -гетеропереход 261—262
 n — P -гетеропереход 256, 258, 259
- Область пространственного заряда 241, 242, 270
 Ограничение для носителей тока 280
 ОГС-лазеры (лазеры на односторонней гетероструктуре) 19, 20, 87, 88
 Оператор Гамильтона 146, 148
 — импульса 146
 Отражение на границе раздела диэлектриков 75—78
- Параметры излучательной рекомбинации 185—189
 Плоская волна 39, 40
 Плотность порогового тока 16—20, 208
 — — — расчет 211—214
 — состояний 136, 137
 — — в валентной зоне 151
 — — — зоне проводимости 151
 — — зависимость от концентрации примесей 155, 160
 — тока номинальная 196
 p — N -гетеропереход 246—257
 Пойнтинга вектор 43, 149
 Показатель преломления 40
 — — дисперсия 138
 — — зависимость от концентрации носителей 45
 — — — температуры 45
 — — — комплексный 42
 Постоянная распространения 39
 — — комплексная 41
 — — связи 115, 122, 128
 Потери в гетеролазерах за счет связи 206
 — — — на рассеяние 206
 — — — — свободных носителей 204—206
 p — P -гетеропереход 259—261
 $PpnN$ -структура 108—111
- Примесная проводимость металлического типа 157
 Пуассона уравнение 242, 243
- Работа выхода 226
 Расходимость луча 95, 96
 — — приближенное выражение 97, 98
- Сила осциллятора 150
 Скорость излучательной рекомбинации 186, 196
 Снеллиуса закон 76
 Спонтанное излучение 140
 — — скорость 140, 153
 — — — связь с коэффициентом поглощения 143, 144
 — — — со скоростью вынужденного излучения 144
 — — спектры 181, 182
 Стерна матричный элемент 175—177
 Суперинжекция 283
 Суперлюминесценция 208—210
- Твердые растворы 28—30
 Теория связанных волн 112—120
 — — — дисперсионное уравнение 117
 — — — постоянная связи 115
 — — — спектр излучения 119
 — — — спектральная селективность 119
 — — — условие Брэгга 113
 — — — — генерации 118
 Ток утечки 285—290
- Уравнение для связанных волн 112—115
 — непрерывности 276
 Условие генерации 192—195
 — — амплитудное 194, 195
 — — лазера с РОС 118
 — — фазовое 195
 — — электронейтральности 190, 228
- Фактор усиления 201
 Ферми — Дирака интеграл 235, 236, 238
 — — функция распределения 139

- Ферми золотое правило 144, 145
— квазиуровни 139, 189, 238, 274, 275
— уровень 226, 240
- Хвосты зон 133
— — модель Гальперина — Лэкса 164
— — — — в гауссовой форме 167, 168
— — — Кейна 161
— моды 54
- Эйнштейна соотношения 139—141
- Электронное сродство 226, 227
Энергия ионизации примесей 230—232
— — — зависимость от концентрации 155, 157
Эффективная плотность состояний валентной зоны 237
— — — зоны проводимости 234, 235
Эффективность квантовая 196
Эффективные массы 221, 224
— — в AlAs 223
— — в $Al_xGa_{1-x}As$ 225
— — в GaAs 221—223
Эффективный показатель преломления 65

ОГЛАВЛЕНИЕ

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Предисловие редактора перевода | 5 |
| Литература | 7 |
| Предисловие | 8 |
| От авторов | 10 |
| Глава 1. Введение | 11 |
| § 1. Вводные замечания | 11 |
| § 2. Получение вынужденного излучения в инжекционных лазерах и возможность их работы в непрерывном режиме при комнатной температуре | 11 |
| § 3. Принцип действия инжекционного лазера | 21 |
| § 4. Инжекционные лазеры | 25 |
| § 5. Материалы | 27 |
| Литература | 30 |
| Глава 2. Оптические поля и распространение волн | 32 |
| § 1. Введение | 32 |
| § 2. Основы теории | 35 |
| § 3. Плоские диэлектрические волноводы | 44 |
| § 4. Распространение волн в симметричном трехслойном плоском волноводе | 48 |
| § 5. Решение дисперсионного уравнения в случае симметричного трехслойного плоского волновода | 56 |
| § 6. Модель зигзагообразных волн и асимметричные трехслойные плоские волноводы | 74 |
| § 7. Расходимость луча | 89 |
| § 8. Коэффициент отражения на торцевых гранях лазера | 98 |
| § 9. Волноводы в четырех- и пятислойных гетероструктурах | 102 |
| § 10. Лазеры с распределенной обратной связью | 111 |
| § 11. Заключение | 128 |
| Литература | 129 |
| Глава 3. Вынужденное излучение в полупроводниках | 132 |
| § 1. Введение | 132 |
| § 2. Соотношения между поглощением, вынужденным и спонтанным излучением | 134 |
| § 3. Вероятность перехода | 144 |
| § 4. Скорости оптического поглощения и излучения в полупроводни- ках | 150 |

| | |
|--------------------------------------------------------------------------------------------------|------------|
| § 5. Зависимость плотности состояний от концентрации | 155 |
| § 6. Оптический матричный элемент | 169 |
| § 7. Рассчитанные оптические спектры | 177 |
| § 8. Пороговая плотность тока | 192 |
| § 9. Заключение | 214 |
| Литература | 214 |
| Глава 4. Гетеропереходы | 218 |
| § 1. Введение | 218 |
| § 2. Зонная структура гетероперехода $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ | 220 |
| § 3. Энергетические зонные диаграммы гетеропереходов | 226 |
| § 4. Плавные гетеропереходы | 265 |
| § 5. Электрические свойства | 270 |
| § 6. Ограничение для носителей тока | 280 |
| § 7. Заключение | 290 |
| Литература | 290 |
| Предметный указатель | 294 |