

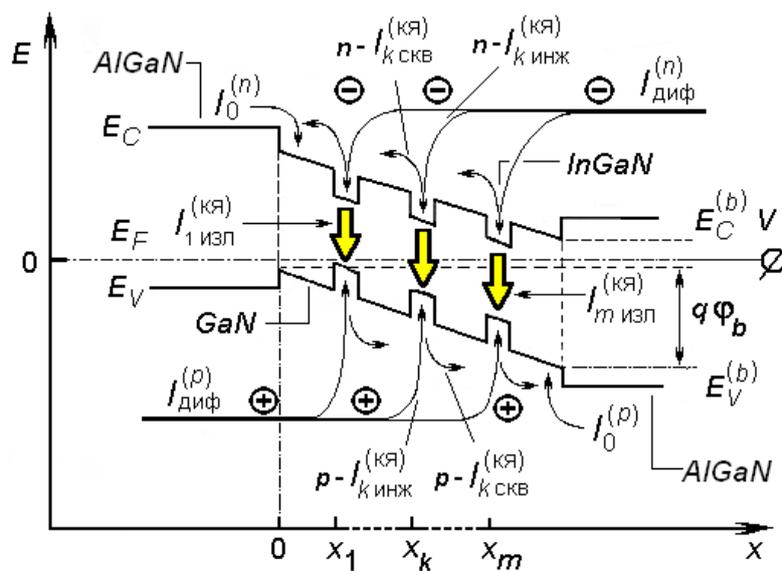
Министерство науки и высшего образования
Российской Федерации

Томский государственный университет
систем управления и радиоэлектроники

В.Н. Давыдов

ОСНОВЫ ОПТОЭЛЕКТРОНИКИ

Учебное пособие для студентов,
обучающихся по физико-техническим
направлениям бакалавриата



ТОМСК
2021

УДК 621.383(075.8)
ББК 32.86я73
Д 138

Рецензенты:

Брудный В.Н., директор НОЦ «Наноэлектроника» НИ ТГУ,
профессор, доктор физико-математических наук
Гермогенов В.П., профессор, заведующей кафедрой
полупроводниковой электроники НИ ТГУ,
доктор физико-математических наук

Давыдов, Валерий Николаевич

Основы оптоэлектроники: учебное пособие / В.Н. Давыдов. - Томск: Томск. гос. ун-т систем упр. и радиоэлектроники, 2021. – 95 с.
ISBN 978-5-86889-938-6

Излагаются физические основы, принципы работы элементов и приборов современной полупроводниковой оптоэлектроники, использующих свойства объёмных материалов, гетероструктур, *p-n* переходов. В изложении материала сделан акцент на описание физических процессов формирования электрических, оптических, флуктуационных и фотоэлектрических свойств объектов оптоэлектроники. Для проверки понимания полученных знаний каждая глава пособия заканчивается десятью вопросами.

В первых главах пособия в сжатой форме излагаются основы зонной теории полупроводников, необходимых для понимания принципа работы элементов оптоэлектроники. При этом особое внимание уделено описанию и объяснению важных понятий, используемых для понимания зонной модели твердого тела и объяснению которых в учебной литературе уделяется недостаточно внимания. Это понятие длины свободного пробега и время жизни фотона, времени жизни неравновесных носителей заряда, понятие квазиуровней Ферми, дисперсии случайной величины и другие. Рассмотрение эффекта двулучепреломления в анизотропных кристаллах в наглядной форме объясняется возникновением набега фазы, а также предсказывается эффект самоограничения двулучепреломления. Для полноты рассмотрения эффектов поворота плоскости поляризации световой волны рассматриваются ориентационные и оптические свойства жидких кристаллов. В дополнение к эффекту двулучепреломления в качестве возможного способа управления световым излучением рассматривается эффект Франца-Келдыша.

В последующих главах описаны эмиссионные и фотоэлектрические свойства полупроводниковых элементов на основе фотопроводимости и фотоэдс в *p-n* переходах, позволяющих аргументировано выбрать оптимальный способ эмиссии излучения и его регистрации. Приведённая далее информация о флуктуационных свойствах полупроводниковых элементов поясняет физическое содержание дисперсии шума и позволяет минимизировать шумы оптоэлектронного прибора для достижения максимальной обнаружительной способности.

Одобрено на заседании каф. электронных приборов, протокол № 15 от 01.09.2020.

УДК 621.383(075.8)
ББК 32.86я73

© Давыдов В.Н., 2021
© Томск. гос. ун-т систем упр.
и радиоэлектроники, 2021

ОГЛАВЛЕНИЕ

| | |
|--------------------------------------------------------------------------|-------|
| 1 ВВЕДЕНИЕ..... | 5-6 |
| 2 ЭЛЕМЕНТЫ ЗОННОЙ ТЕОРИИ ТВЕРДЫХ ТЕЛ..... | 7-21 |
| 2.1 Общие положения..... | 7 |
| 2.2 Модельные представления..... | 7-9 |
| 2.3 Зонная диаграмма полупроводника..... | 9-12 |
| 2.4 Квазиимпульс электрона. Долины энергии и зона Бриллюэна..... | 12-14 |
| 2.5 Распределение частиц по энергии. Функция Ферми – Дирака | 14-16 |
| 2.6 Положительно заряженные квазичастицы – дырки..... | 16-17 |
| 2.7 Движение частиц под действием электрического поля..... | 17-19 |
| 2.8 Механизмы формирования тока в полупроводниках..... | 19-21 |
| Вопросы по главе 2..... | 21 |
| 3 ВЗАИМОДЕЙСТВИЕ ОПТИЧЕСКОГО ИЗЛУЧЕНИЯ С ТВЕРДЫМИ ТЕЛАМИ..... | 22-30 |
| 3.1 Основные параметры поглощения излучения..... | 22-24 |
| 3.2 Основные механизмы поглощения излучения..... | 24-25 |
| 3.3 Собственное поглощение. Прямые и непрямые переходы..... | 25-27 |
| 3.4 Примесное поглощение излучения..... | 27-28 |
| 3.5 Поглощение свободными носителями | 28-29 |
| 3.6 Решеточное поглощение..... | 29 |
| Вопросы по главе 3..... | 29 |
| 4 ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ..... | 30-50 |
| 4.1 Основные понятия и параметры..... | 30-32 |
| 4.2 Фотопроводимость полупроводников: собственная и примесная | 32-37 |
| 4.3 Полевые свойства фотопроводимости..... | 37-38 |
| 4.4 Частотные свойства фотопроводимости..... | 38-40 |
| 4.5 Фотовольтаические эффекты в полупроводниках..... | 40-43 |
| 4.6 Квазиуровни Ферми в полупроводниках..... | 43-45 |
| 4.7 Барьерная фотоэдс в полупроводниковых $p-n$ переходах..... | 45-49 |
| 4.7.1 Электронно-дырочные переходы | 45-47 |
| 4.7.2 Фотоэлектрические свойства $p-n$ переходов. Барьерная фотоэдс..... | 47-49 |
| 4.8 Лавинные фотодиоды..... | 49-50 |
| Вопросы по главе 4 | 50 |

| | |
|----------------------------------------------------------------|-------|
| 5 ЭМИССИЯ ИЗЛУЧЕНИЯ ИЗ ТВЕРДЫХ ТЕЛ..... | 51-67 |
| 5.1 Описание излучательных процессов в полупроводниках..... | 51-53 |
| 5.2 Зависимость эмиссии от уровня легирования..... | 53-55 |
| 5.3 Спонтанное и вынужденное излучение атома..... | 55-58 |
| 5.4 Стимулированное излучение твердых тел..... | 58-59 |
| 5.5 Принцип создания генератора света..... | 59-61 |
| 5.6 Линия излучения твердых тел..... | 61-63 |
| 5.7 Светодиоды на основе гомогенного <i>p-n</i> перехода..... | 64-65 |
| 5.8 Светоизлучающие структуры с высоким квантовым выходом..... | 65 |
| 5.9 Лазеры на полупроводниковых гетероструктурах..... | 65-67 |
| Вопросы по главе 5..... | 67-68 |
| 6 ФЛУКТУАЦИОННЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ..... | 69-76 |
| 6.1 Типы и природа шумов..... | 69-71 |
| 6.2 Основные параметры и характеристики шумов..... | 71-72 |
| 6.3 Полевые и фоновые свойства шумов в полупроводниках..... | 72-73 |
| 6.4 Частотные свойства шумов различной природы..... | 73-75 |
| Вопросы по главе 6..... | 75-76 |
| 7 ЭЛЕКТРООПТИЧЕСКИЕ ЭФФЕКТЫ ДЛЯ УПРАВЛЕНИЯ ИЗЛУЧЕНИЕМ..... | 77-81 |
| 7.1 Общие сведения об электрооптических эффектах..... | 77 |
| 7.2 Эффект Франца – Келдыша в оптоэлектронике | 77-80 |
| 7.3 Двухлучепреломление в анизотропных 3D кристаллах..... | 80-81 |
| Вопросы по главе 7..... | 81 |
| 8 ЖИДКИЕ КРИСТАЛЛЫ В ОПТОЭЛЕКТРОНИКЕ..... | 82-94 |
| 8.1 Общие сведения о жидких кристаллах..... | 82 |
| 8.2 Основные свойства жидких кристаллов..... | 82-83 |
| 8.3 Типы и структура жидких кристаллов..... | 83-86 |
| 8.4 Ориентационные эффекты в жидких кристаллах..... | 86-88 |
| 8.5 Оптические свойства жидких кристаллов..... | 88-91 |
| 8.6 Применение жидких кристаллов в оптоэлектронике..... | 91-94 |
| Вопросы по главе 8..... | 94 |
| 9 ПЕРСПЕКТИВЫ РАЗВИТИЯ ОПТОЭЛЕКТРОНИКИ..... | 95 |
| 10 ЛИТЕРАТУРА..... | 95 |

1 ВВЕДЕНИЕ

При решении технических задач, в своей повседневной деятельности мы широко используем различные приборы и устройства твердотельной электроники. Применение электронных приборов в различных сферах деятельности человека не случайно: они отличаются малыми габаритами и малым энергопотреблением, высоким быстродействием и высокой чувствительностью к внешним воздействиям (оптическому и корпускулярному излучению, электрическому и магнитному полям, упругим напряжениям, адсорбции-десорбции примесных молекул и т.д.), простотой и гибкостью управления, универсальностью выполняемых функций. Фактически, все задачи по функциональному преобразованию электрических, магнитных и других сигналов принципиально могут быть решены с помощью средств твердотельной электроники.

Как и все объекты природы, твердотельная электроника имеет свою диалектику развития, сопровождающегося накоплением количественных изменений и последующим качественным переходом с сохранением в новом состоянии перспективных элементов предыдущего своего состояния.

Чтобы представить путь развития твердотельной электроники и лучше понять изменение требований к ней в будущем, обратимся к истории. Первые элементы электронной техники на основе твердого тела в виде детекторов на основе контактов металлической иглы к полупроводниковому кристаллу появились в двадцатые годы прошлого столетия. Исследование этих контактов привело к созданию первых приборов твердотельной электроники: диодов на основе контактов металл-полупроводник. Первый качественный скачок в развитии твердотельной электроники произошел в пятидесятые годы XX века в связи с изобретением американскими учеными Шокли и Брэдштейном нового элемента – биполярного транзистора на основе структур $p-n-p$ и $n-p-n$. С помощью транзисторов многие операции по преобразованию сигналов удалось значительно упростить, переведя их с вакуумной техники на твердотельную, а также создать принципиально новые приборы, не имеющие аналогов среди ламповой техники. Это различные датчики внешних сигналов, тиристоры, симисторы и т.д.

Однако в конце шестидесятых годов XX века в твердотельной электронике возникла кризисная ситуация, связанная с низкой надёжностью сверхсложных электронных схем (прежде всего электронно-вычислительных машин и периферии к ним), их большими габаритами и высоким энергопотреблением. Эта кризисная ситуация была преодолена с помощью идей и технологий интегральной микроэлектроники - в результате их развития появились микросхемы различного назначения, содержащие в себе сотни тысяч транзисторов, диодов, планарных резисторов и емкостей.

Однако последующее развитие и совершенствование техники, ориентированное преимущественно на решение информационных задач, создание современной вычислительной техники, показало, что многие проблемы решаются средствами микроэлектроники неоптимальным способом. В первую очередь это касалось принципов функционирования устройств ввода и вывода информации, увеличения памяти компьютера, отображения информации, увеличения ёмкости соединительных проводов, повышения быстродействия и снижения энергопотребления. Более того, с развитием фундаментальной науки появились материалы, обладающие электронными свойствами, которые могут быть использованы в решении задач твердотельной электроники. К ним относятся жидкие кристаллы, сверхпроводники, синтезированные с помощью новейшей технологии различные квантовые структуры: сверхрешетки, квантовые ямы и т.д. В результате включения этих материалов и новых физических явлений и свойств в арсенал материалов и методов твердотельной электроники стали развиваться ее новые направления: оптоэлектроника, акустоэлектроника, пьезоэлектроника, магнитоэлектроника, криоэлектроника, хемотроника, молекулярная электроника, наноэлектроника.

Развитие этих направлений в ближайшие годы позволит решить проблемы современной микроэлектроники: дальнейшей интеграции электронных устройств, наращивания их функциональных возможностей, надежности, снижения габаритов и энергопотребления.

В данном учебном пособии приведено описание основных электрических и оптических явлений, происходящих в объеме твердого тела или в области его контакта с веществами другого физико-химического состава, а также в жидких кристаллах, которые используются или могут быть использованы для решения фундаментальных задач оптоэлектроники по регистрации, преобразованию и передаче информации на основе оптических и электрических явлений.

2 ЭЛЕМЕНТЫ ЗОННОЙ ТЕОРИИ ТВЕРДЫХ ТЕЛ

2.1 Общие положения

Ввиду того что подавляющее число приборов и устройств оптоэлектроники используют свойства полупроводников, начнём изучение физического базиса данного курса с рассмотрения фундаментальных свойств твердых тел. Нас будут интересовать преимущественно их электрические и оптические свойства.

Если рассматривать какое-либо явление в твердом теле, то для понимания физических процессов, которые формируют данное явление, необходимо знать поведение элементарной частицы: электрона в заданных условиях - именно он определяет свойства твердых тел при действии различных физических полей. Исторически первой средой, в которой были изучены электронные явления, был газ. Последующее развитие науки и техники привело к необходимости изучения электронных процессов в твердых телах. Однако уже первые эксперименты с электропроводностью твердых тел показали, что она не может быть объяснена развитой к тому времени классической электронной теорией, хорошо объясняющей поведение электронов в вакууме и газах. Было непонятно: как, имея огромное количество электронов (находящихся у атомов), твердые тела различного физико-химического состава могут иметь различающуюся на многие порядки электропроводность, а для отдельных веществ вообще крайне низкую; почему их электропроводности по-разному зависят от температуры, освещения и т.д.? На эти и другие вопросы классическая электронная теория ответить не могла, и потому были предприняты попытки построения новой теории - теории твердого тела.

2.2 Модельные представления

Поскольку атом является квантовой системой, то строгий подход к объяснению электропроводности твердых тел должен строиться на основе квантовых представлений об их строении. Исторически первая попытка в этом направлении была предпринята немецким физиком Зоммерфельдом. Он считал, что твердое тело состоит из невзаимодействующих между собой атомов – как в газах. Принципиальное же отличие предложенной им модели электропроводности твердых тел от модели газов и вакуума состояло в том, что в этой модели атом считался квантовой частицей: электроны в твердом теле могли иметь энергии, определяемые квантовой природой изолированного атома. Предполагалось также, что свободные электроны не принадлежат отдельным атомам, а могут свободно перемещаться по кристаллу, не испытывая влияния кулоновского поля от заряда его ионов. Это допущение было основано на том умозаключении Зоммерфельда, что суммарное поле ионов внутри кристалла точно компенсируется суммарным полем свободных электронов (их количества

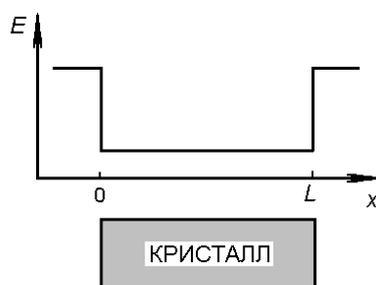


Рисунок 1 - Энергетическая диаграмма кристалла в вакууме по Зоммерфельду

должны быть одинаковы), за исключением поверхности. Таким образом, заряд каждого иона должен быть компенсирован находящимся рядом зарядом свободного электрона. Поэтому движущийся через кристалл свободный электрон не испытывал электрического взаимодействия ни с ионами, ни с другими электронами. Поле же у поверхности кристалла

удерживает электроны в кристалле, затрудняя их выход в вакуум. Энергия, необходимая для создания свободного электрона в твердом теле, равна энергии ионизации изолированного атома – E_{ion} . Зависимость энергии электрона в кристалле от координаты x по модели Зоммерфельда показана на рис. 1.

Модель электропроводности Зоммерфельда оказалась малополезной. Важным вопросом, на который она не давала ответ, являлся вопрос о причинах различия электропроводности металлов, полупроводников и диэлектриков. В этой связи Зоммерфельд считал, что в металлах валентные электроны легко отщепляются от атома, становясь свободными, и участвуют в электропроводности (E_{ion} мала), тогда как у диэлектриков это сделать крайне трудно (E_{ion} велика). Однако это объяснение различия электропроводностей различных кристаллов не согласуется с тем фактом, что численные значения энергий ионизации изолированных атомов E_{ion} металлов и диэлектриков слабо различаются между собой, а иногда у металлов могут быть даже больше, чем у диэлектриков.

Следующий шаг в уточнении модельных представлений о поведении электрона в кристалле сделан Блохом, который при решении квантово-механической задачи о поведении свободного электрона в объёме твердого тела учёл взаимодействие атомов между собой, а также взаимодействие свободного электрона с ионами в узлах кристаллической решетки. Характер изменения энергии свободного электрона, движущегося вдоль ионных узлов решётки, имеет вид, показанный на рис. 2. Как можно видеть из рисунка, в объёме кристалла изменение энергии электрона носит периодический характер с периодом, равным постоянной решётки a . Исследуя поведение свободного электрона в таком поле с помощью аппарата квантовой механики, Блох показал, что в идеальной периодической решётке электрон движется свободно, как в вакууме, не испытывая столкновений с ионами кристалла - он их не “видит”! Но самое главное, оказалось, что свободный электрон в кристалле не может иметь произвольную энергию: весь диапазон ее значений оказался разделен на чередующиеся полосы разрешенных и запрещенных энергий. Их назвали зонами энергий. Более того, ширина зон запрещенных и разрешенных энергий определяется не только природой атомов решетки, но и степенью перекрытия электронных оболочек соседних

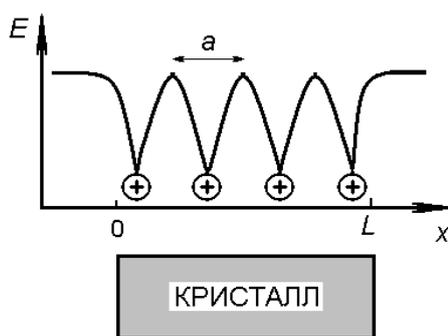


Рисунок 2 - Энергетическая диаграмма кристалла в вакууме по Блоху

атомов. В итоге ширина запрещенной зоны оказалась отличной от энергии ионизации изолированного атома E_{ion} , и это отличает модель Блоха от модели Зоммерфельда, где ширина запрещенной зоны отождествлена с E_{ion} .

Другой важный вывод, который наряду с теорией Блоха использован для объяснения электропроводности твердых тел, состоит в том, что, не говоря о внешних электронах, даже электроны с внутренних оболочек атома нельзя считать принадлежащими конкретному атому кристалла. Эти электроны одних атомов могут меняться местами с электронами других атомов, идентичными данным электронам, например, с внутренней d - орбиты одного атома на такую же d - орбиту другого, если у первого атома на этой орбите имеется вакансия

(так называемый «механизм Варли»).

Именно эти два физических механизма переноса заряженных частиц в твердом теле легли в основу теории электропроводности полупроводников, разработанной Вильсоном, и которая до сих пор используется для объяснения электрических свойств твердых тел.

2.3 Зонная диаграмма полупроводника

Для объяснения электронных процессов в твердых телах удобно пользоваться так называемой зонной диаграммой твердого тела (полупроводника, диэлектрика). Она представляет собой зависимость энергии электрона от координаты в твердом теле. Чтобы нарисовать его зонную диаграмму, начнем с простейшей модели, описывающей поведение электрона в твердом теле, модели Зоммерфельда. Согласно ей, разрешенные уровни валентных электронов (т.е. электронов, находящихся на внешних орбитах атомов, именно они в первую очередь взаимодействуют с внешними полями, так как являются внешними электронами, как и человек, общается с окружающей средой через свою внешнюю оболочку - кожу) в кристалле расположены близко друг от друга и образуют систему уровней, простирающихся от дна потенциальной ямы до больших значений энергии. Эту систему уровней можно назвать «валентной зоной», поскольку она описывает энергии электронов на валентных орбитах атома. Уровни энергии электронов с внутренних оболочек остаются невозмущенными (напомню, что эти электроны не взаимодействуют ни друг с другом, ни с ионами решетки) и совпадают с уровнями энергии изолированного атома. Эти уровни одни и те же для каждого атома кристалла, и потому они оказываются N -кратно вырождены, где N - число атомов в кристалле, т.е. одному значению энергии соответствует N физически различных состояний. Разрешенные уровни энергии электронов в кристалле по модели Зоммерфельда показаны на рис.3 в левой его части. Здесь за нулевую энергию выбрана энергия свободного электрона с нулевой кинетической энергией.

В теории Блоха вследствие введения периодического потенциала решетки энергетические уровни группируются в определенные полосы, называемые «зонами разрешенных энергий», разделенные областями, в которых нет разрешенных значений энергий, «запрещенные зоны». Для внутренних электронов разрешенные зоны чрезвычайно

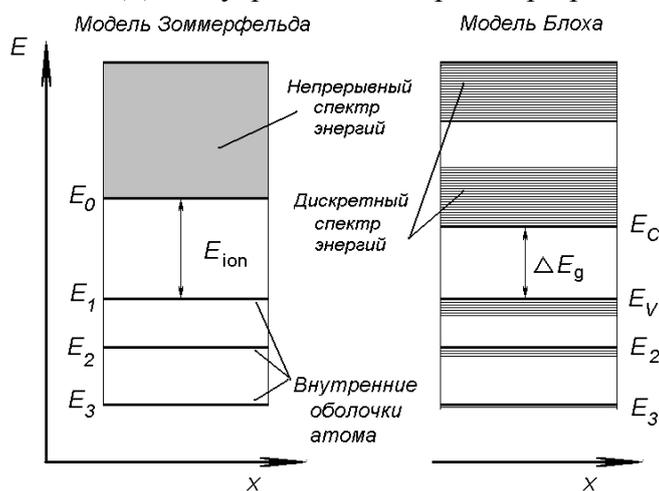


Рисунок 3 – Сравнение энергетических диаграмм электрона в твердом теле по Зоммерфельду (левый рисунок) по Блоху (правый рисунок)

узкие и соответствуют атомным уровням энергии. Зоны энергий для внешних (валентных) электронов оказываются широкими. Расположение этих уровней энергии показано на правой части рис. 3. В блоховской модели каждая зона состоит из множества близко расположенных уровней, так что во многих практических приложениях их можно рассматривать как непрерывный спектр.

К такой же картине энергетических зон в твердом теле можно прийти и другим путем. Для этого рассмотрим два одинаковых атома, расположенных на большом удалении друг от

друга, и предположим, что каждый из них обладает одним невырожденным уровнем энергии (это равносильно тому, что из спектра энергий изолированного атома выбирается только один уровень). В этом случае спектр энергии системы из двух атомов содержит один уровень энергии, вырожденный дважды. Теперь будем сближать эти атомы. В результате усиливающегося взаимодействия между ними вырожденный уровень энергии расщепится на два уровня, энергетический зазор между которыми будет увеличиваться с уменьшением расстояния (рис.4). Этот результат известен в классической теории из задачи по взаимодействию двух осцилляторов: при сближении двух одинаковых колебательных контуров образующаяся колебательная система будет иметь две резонансные частоты, отличающиеся от собственной частоты изолированных контуров на некоторую величину в сторону больших и меньших значений частот. Значит, если расстояние между атомами a_0 , то верхний уровень энергии валентной зоны будет иметь энергию, равную E_1 . Нижнее значение энергии зоны проводимости будет E_2 .

Данный подход применим также к системе из N атомов. Предположим, что рассматриваемый кристалл растянут и все межатомные расстояния в нём велики. Тогда разрешённые уровни энергии для электрона в таком кристалле совпадают со спектром отдельного атома, но каждый уровень при этом N -кратно вырожден. Если уменьшать расстояние между атомами, то из-за межатомного взаимодействия (из-за перекрытия электронных оболочек сближаемых атомов) каждый уровень расщепится на серию из N одиночных, невырожденных уровней. В результате вместо вырожденного уровня получим зону разрешённых энергий для электрона из N плотно расположенных уровней (рис.5). Для

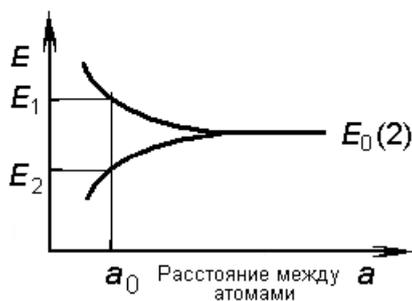


Рисунок 4 – Расщепление уровня энергии при сближении атомов в кристалле

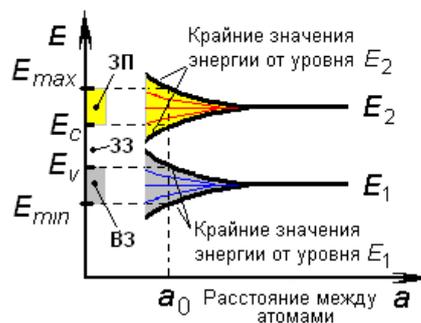


Рисунок 5 – Формирование энергетической диаграммы для электронов в кристалле

внутренних электронов из-за их экранировки внешними электронами возмущение оказывается слабым по сравнению с взаимодействием их с ядром и потому расщепление этих уровней будет малым. Для валентных электронов оно может быть большим: вплоть до перекрытия зон. Рисунок 5 наглядно демонстрирует образование зоны проводимости (ЗП) полупроводника, его запрещенной зоны (ЗЗ) и валентной зоны (ВЗ). Можно видеть, что увеличение расстояния между атомами приведет к сужению валентной зоны и зоны проводимости, но увеличению запрещенной зоны.

Однако энергетическая диаграмма состояний электрона в твердом теле показывает возможности размещения электронов (подобно креслам в зрительном зале). Но будут ли они заняты и как – этот вопрос требует отдельного рассмотрения.

Проанализируем случай, когда валентная зона и зона проводимости не перекрываются. Чтобы ответить на вопрос о распределении электронов по состояниям в зонах, необходимо воспользоваться принципом Паули, согласно которому на каждом уровне в зоне может находиться не более двух электронов с противоположно ориентированными спинами. Начнём с глубоко лежащих атомных уровней. В любом атоме каждый такой уровень должен быть занят двумя электронами, так что в кристалле полностью заполненными оболочкам

атомов можно поставить в соответствие полностью заполненные зоны. Частично заполнены будут зоны, соответствующие внешним (валентным) электронам. Различие между полностью и частично занятыми зонами играет важную роль в функционировании полупроводникового прибора. Оно иллюстрируется рис.6. Здесь пустые уровни соответствуют возбужденным состояниям валентных электронов; спин электрона показан стрелкой.

С помощью этих модельных представлений можно понять, почему электроны в низших зонах, соответствующих атомным остовам, не могут принять участия в электропроводности. Для участия в электропроводности электрону необходимо приобрести дополнительную энергию за счёт действия приложенного электрического поля. На языке квантовой механики

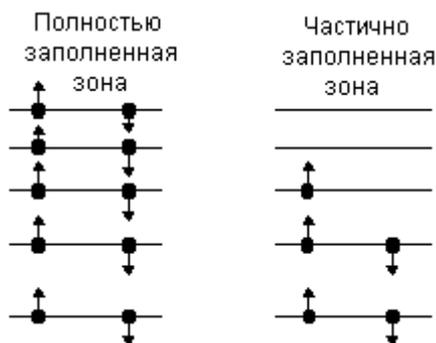


Рисунок 6 – Варианты заполнения уровней энергии электронов в зонах разрешённых значений энергии

это означает, что электрон должен перейти на более высокий энергетический уровень. Если же все уровни в данной зоне заняты, то электрон не может ускориться в электрическом поле до тех пор, пока в результате возбуждения не перейдёт в лежащую выше зону, в которой имеются свободные уровни. Процесс возбуждения, однако, маловероятен, так как электрону необходимо сообщить достаточно большую энергию (обычно единицы эВ). Поэтому, чтобы электроны приняли участие в электропроводности (а значит, обнаружили себя в электрических измерениях), необходимо наличие близлежащих пустых энергетических уровней, на которые рассматриваемые электроны могли бы перейти при действии приложенного электрического поля. Именно это обстоятельство составляет принципиальное различие в электропроводности между металлами, полупроводниками и диэлектриками.

Рассмотрим теперь наивысшую валентную зону, в которой имеются как занятые электронами, так и свободные уровни. В щелочном металле, например, каждый атом содержит по одному валентному электрону. В невозбуждённом состоянии N электронов займут $N/2$ наиболее глубоких уровней в этой валентной зоне. В результате зона окажется заполненной лишь наполовину и в ней останется ещё много свободных уровней энергии, необходимых для возникновения электропроводности. Значит, при наличии лишь одного валентного электрона у атома ситуация подобна модели металла по теории Зоммерфельда.

Обратимся к кристаллу, построенному из атомов, содержащих на внешней электронной оболочке два электрона с противоположно ориентированными спинами. В этом случае валентная зона кристалла, т.е. наивысшая из всех содержащих электроны зон, оказывается полностью занятой. Если к тому же между валентной зоной и следующей за ней верхней пустой зоной имеется энергетический зазор - запрещённая зона, то такой кристалл не будет обладать электропроводностью и соответствующее вещество будет диэлектриком, если запрещённая зона широкая. Однако если она невелика, то существует возможность теплового возбуждения электронов, приводящего к их забросу из валентной зоны в пустую зону проводимости, после чего эти электроны могут принимать участие в электропроводности. Число возбуждённых электронов будет увеличиваться с ростом температуры, что приведёт к росту и электропроводности. Так ведут себя кристаллические полупроводники и диэлектрики, различие между которыми заключается только в величине запрещённой зоны. Обычно полупроводник - это кристалл, у которого в химически чистом

состоянии величина ΔE_g является положительной и не превышает 2-3 эВ.

Сказанное иллюстрируется рис. 7, а и 7, б на которых показано энергетическое состояние элементарных полупроводников типа кремний и германий при абсолютном нуле температуры и при комнатной. Сообщив электрону энергию, равную ширине запрещенной зоны, можно перевести его в зону проводимости и тем самым заставить его принять участие в электропроводности. Оказалось, что можно добиться высокой электропроводности полупроводника и без столь значительных затрат энергии: введение малого количества специальных примесей может заметно повлиять на число свободных электронов в зоне проводимости и дырок в валентной зоне, так что материалы даже с $\Delta E_g < 3$ эВ ведут себя как низкоомные полупроводники. Эти примеси называются донорными (элементы V группы таблицы Менделеева) или акцепторными (элементы III группы) в зависимости от того, отдаёт или забирает атом примеси электрон у атома полупроводника.

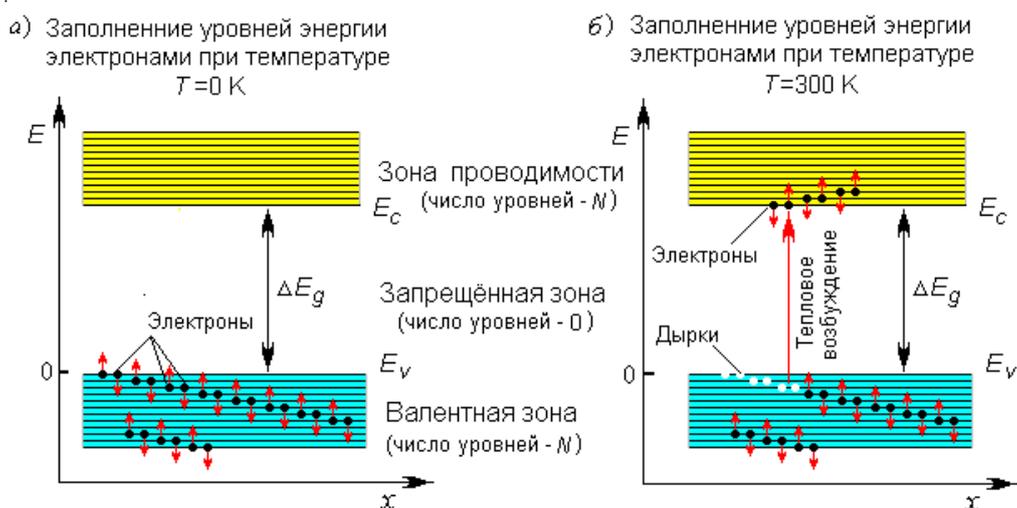


Рисунок 7 – Заполнение валентной зоны и зоны проводимости электронами при разных температурах кристалла

Пользуясь энергетической диаграммой, можно объяснить электропроводность диэлектриков, полупроводников и металлов. У диэлектриков энергетическая диаграмма аналогична приведенной на рис. 8, но у них ширина запрещенной зоны велика: $\Delta E_g > 3$ эВ. У полупроводников ширина запрещенной зоны лежит в диапазоне $0.1 < \Delta E_g < 2$ эВ. У металлов $\Delta E_g = 0$ или даже отрицательная величина.

2.4 Квазиимпульс электрона. Долины энергии и зона Бриллюэна

Из классической механики известно, что для характеристики поведения движущейся с соударениями частицы наряду с энергией частицы необходимо введение нового физического параметра – импульса частицы, который определяют, как произведение массы частицы на ее скорость: $p = mv$. Для характеристики движения квантовой частицы (например, электрона в зоне проводимости), которая является одновременно частицей и волной и потому относится к квазичастицам, также вводят понятие импульса. Однако, учитывая двойственную природу квантовой частицы, его называют квазиимпульсом, т.е. «почти импульсом». Двойственная природа квантовой частицы приводит к двум определениям ее квазиимпульса p . Квазиимпульс электрона-частицы вычисляется через ее кинетическую энергию:

$$E = \frac{m_n^* v^2}{2} = \frac{p^2}{2m_n^*}$$

где m_n^* – эффективная масса электрона, которая может отличаться от массы электрона в

состоянии покоя. В квантовой механике квазиимпульс электрона-волны таков:

$$p = k\hbar, \quad (2.1)$$

где \hbar – постоянная Дирака, равная $h/2\pi$; h – постоянная Планка; k – волновой вектор, модуль которого определяется длиной волны: $k = 2\pi/\lambda$, λ – длина волны. Если рассматривать кристалл в виде прямоугольной потенциальной ямы длиной L и с барьером на ее концах бесконечной высоты, то частица-волна в кристалле будет описываться в виде пакета стоячих волн. Это возможно, если по длине кристалла укладывается целое число полувольт, входящих в волновой пакет частицы-волны: $L = n\lambda/2$ или $\lambda = 2L/n$, где число $n = \pm 1, \pm 2, \dots, \pm N$ указывает, сколько раз на длине кристалла $L = N \cdot a$ укладывается полувольт. Подставив это условие в выражение для волнового вектора, получим, что его модуль может принимать только дискретные значения:

$$k = n(\pi/L), \quad n = \sqrt{n_x^2 + n_y^2 + n_z^2}. \quad (2.2)$$

Здесь n_x, n_y, n_z – целые числа, показывающие количество полувольт, укладываемых по размерам кристалла L_x, L_y, L_z вдоль координатных осей x, y, z (положительные значения n_x, n_y, n_z) и осей $-x, -y, -z$ (отрицательные значения n_x, n_y, n_z) соответственно. Все возможные значения и направления вектора \vec{k} в трехмерном пространстве (k_x, k_y, k_z) заполняют некоторый объем, образуя симметричную трехмерную фигуру, которую называют зоной Бриллюэна.

Существование максимального значения $n_0 = N$ связано с тем обстоятельством, что на самой короткой длине волны λ_0 , описывающей квантовую частицу, должно укладываться не менее двух атомов – один в максимуме, а другой в минимуме волны:

$$L = n_0\lambda_0/2 = Na, \quad \lambda_0 = 2a.$$

При меньшем числе атомов гармоническое колебание не будет распознано как волна. В этом случае $k_{\max} = \lambda/a$ есть максимальное значение волнового вектора. Этому максимальному значению соответствует край зоны Бриллюэна. В случае изотропной среды, параметры которой не зависят от направления их измерения, эта зона представляет собой шар радиусом $k_{\max} = \lambda/a$. В многокомпонентных кристаллах расстояние между соседними атомами a может зависеть от выбора направления, и потому в них зона Бриллюэна имеет сложный вид.

Используя волновое определение квазиимпульса, кинетическую энергию электрона можно представить в виде

$$E = \frac{\hbar^2 k^2}{2m_n^*},$$

где k принимает дискретные значения, определяемые по выражению (2.2). Тогда в кристалле, имеющем форму куба с ребром длиной L и рассматриваемом как потенциальная яма, разрешенные значения энергии квантовой частицы-электрона будут равны:

$$E = \left(\frac{\hbar^2}{8m_n^*L^2} \right) \cdot (n_x^2 + n_y^2 + n_z^2). \quad (2.3)$$

Поскольку квантовые числа n_x, n_y, n_z могут принимать только целые значения, то и величина E будет принимать только дискретные значения. Это означает, что спектр энергий свободного электрона как квантовой частицы в кристалле, строго говоря, дискретен. Однако в реальных кристаллах получающиеся энергетические расстояния между соседними уровнями энергии столь малы ($\sim 10^{-18}$ эВ для кристалла с $L = 1$ см), что спектр разрешенных значений энергий свободного квантового электрона в кристалле с достаточной точностью можно

считать непрерывным, что позволяет рассматривать его как классическую частицу.

Рассмотрим зависимость энергии от волнового вектора для одного направления в пространстве волновых векторов. Обычно в качестве начала отсчета энергии принимают энергию потолка валентной зоны. Тогда дну зоны проводимости отвечает более высокая потенциальная энергия, соответствующая ширине запрещенной зоны. Как это следует из выражения (2.2), зависимость E от k в пределах зоны разрешенных энергий является параболической. На рис. 8 показана зависимость E от k реальных полупроводников, у которых эффективная масса дырки больше эффективной массы электрона в зоне проводимости. Следовательно, крутизна зависимости E от k валентной зоны оказывается

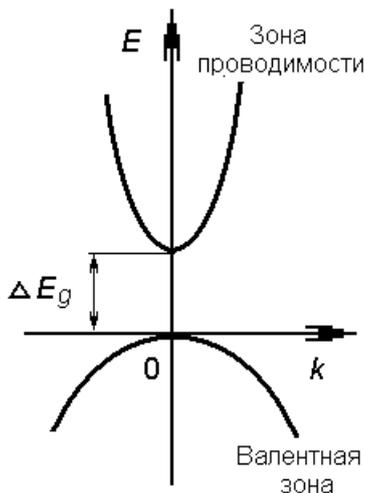


Рисунок 8 – Долинный спектр электронов в прямозонном кристалле



Рисунок 9 – Долинный спектр электронов в непрямоzonном кристалле

меньше, чем зоны проводимости. Такое распределение состояний называется параболической долиной. Смысл этого названия несколько проясняется при трехмерном представлении зависимости E от k_x и k_y – получающаяся при этом пространственную фигуру $E(k)$ можно рассматривать как долину среди гор.

Показанное на рис. 8 взаимное положение экстремумов зон разрешенных энергий электрона в кристалле не является единственно возможным. Из-за сильного взаимодействия с ближайшими соседями, со следующими за ними атомами и с атомами, более удаленными по отношению к рассматриваемому атому решетки полупроводника, минимум долины может быть смещен относительно точки $k = 0$ в некотором кристаллографическом направлении, например, k_x (рис.9). Полупроводники, у которых минимум зоны проводимости имеет место при том же значении квазиимпульса, что и максимум валентной зоны, называют прямоzonными. При их несовпадении полупроводники называют непрямоzonными. Встречаются оба типа полупроводников.

2.5 Распределение частиц по энергии. Функция Ферми-Дирака

Наличие разрешенных энергетических уровней не является достаточным условием возникновения электропроводности кристалла. Для того чтобы кристалл проводил электрический ток, наряду с зонами разрешенных энергий необходимо наличие в них свободных носителей заряда. Рассмотрим заполнение электронных состояний в зонах разрешенных энергий носителями заряда. Концентрация электронов на уровне с данной энергией E определяется плотностью состояний (их количеством в единичном энергетическом интервале энергии) в зоне. Выражение для плотности электронных состояний приведено в литературе по физике полупроводников. Однако важным является не только наличие разрешенных для электронов состояний (рядов кресел в зрительном зале), но

и вероятность их заполнения электронами (в нашей аналогии электронами). Она описывается распределением Ферми-Дирака и имеет вид

$$f_n(E) = \frac{1}{1 + \exp\left(\frac{E - F}{kT}\right)}$$

Ход этой функции в зависимости от энергии электронного состояния показан на рис. 11, а. Она показывает вероятность заполнения электронами состояния с энергией E . Здесь k – постоянная Больцмана; T – абсолютная температура; F – уровень энергии Ферми, т.е. уровень энергии, вероятность заполнения которого электронами равна 1/2. Если рассматриваемый уровень энергии на несколько единиц kT лежит выше уровня Ферми, то

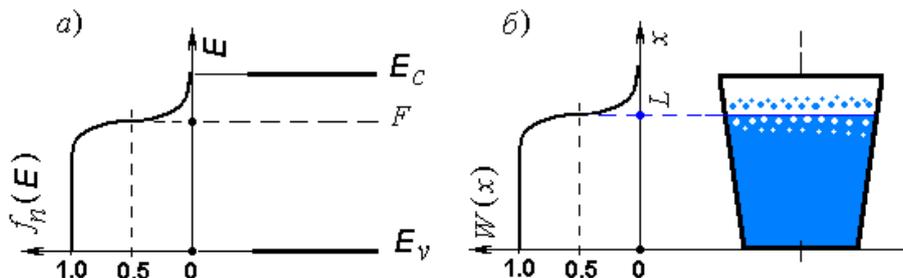


Рисунок 10 – Пояснение к пониманию уровня Ферми в кристалле

вероятность заполнения этого уровня электронами равна нулю. Если же уровень энергии находится ниже уровня Ферми на несколько единиц kT , то вероятность его заполнения электронами равна единице.

Для того чтобы лучше понять смысл распределения Ферми-Дирака, обратимся к аналогии. Если имеется стакан с газированной водой, заполненный ею частично, то функция заполнения стакана водой $W(x)$, где x – координата вдоль стакана, отсчитываемая от его дна, имеет значение $W(x) = 1$ при значениях координаты x , соответствующих заполненной части стакана. Если же значения x находятся выше уровня воды, то $W(x) = 0$ (рис.10, б). Такое распределение воды по высоте стакана аналогично заполнению уровней энергии в полупроводнике (рис.10, а). Если уровень энергии лежит ниже уровня Ферми, то он заполнен электронами (в рассмотренном примере - водой). Если же уровень энергии лежит выше, то он свободен от электронов. Уровень Ферми разделяет эти области заполнения (но между ними существует тонкая переходная область), в приведенном примере в качестве такой разделительной линии – уровня Ферми выступает уровень воды в стакане.

Согласно квантово-механическому принципу Паули, на каждом энергетическом уровне может находиться не более двух электронов с противоположно направленными спинами. Нередко бывает, что рассматриваемому энергетическому уровню соответствует более чем одно состояние (говорят, что уровень вырожден). Тогда при определении количества электронов на данном уровне функцию Ферми-Дирака умножают на фактор вырождения – число, показывающее, сколько состояний имеют данное значение энергии.

2.5 Положительно заряженные квазичастицы – дырки

В идеальном полупроводнике, когда все электроны находятся в наинизших энергетических состояниях, в зоне проводимости нет электронов. Однако такое положение теоретически возможно лишь при абсолютном нуле температуры. При реальных температурах, соответствующих применению полупроводников, в зоне проводимости всегда находится некоторое количество электронов, заброшенных туда из заполненной валентной зоны в результате передачи им тепловой энергии решетки. Поэтому в таком полупроводнике может течь электрический ток, обусловленный наличием в зоне проводимости свободных

электронов. Мгновенная плотность тока \bar{j} , обусловленная движением какого-либо электрона, пропорциональна его скорости \bar{v} и совпадает с ним по направлению. Величину \bar{j} можно вычислить следующим образом.

Пусть имеется N_V электронов (с вектором скорости, равным \bar{v}) в объеме V . Число электронов, пересекающих в единицу времени единичную площадку, перпендикулярную направлению скорости, определяется как $N_V \cdot |\bar{v}|/V$. Поэтому плотность тока равна $-qN_V \cdot |\bar{v}|/V$, а плотность тока, обусловленная движением одного электрона, будет

$$\bar{j} = -\left(\frac{q}{V}\right) \cdot \bar{v} .$$

Предположим теперь, что речь идет об электроне, состояние которого характеризуется волновым вектором \bar{k} и соответственно скоростью \bar{v} . Плотность электрического тока, создаваемого многими движущимися электронами, равна сумме плотностей токов, создаваемых отдельными движущимися электронами:

$$\bar{j} = -\left(\frac{q}{V}\right) \cdot \sum_s \bar{v}_s ,$$

где вектор скорости \bar{v}_s соответствует состояниям, занятым электронами. В частности, если электроны полностью заполняют какую-либо зону, то полный ток электронов в этой зоне $\bar{j} = 0$, так как каждому значению \bar{k}_s в зоне обязательно найдется равный по величине и обратный по направлению волновой вектор « $-\bar{k}_s$ », порождающий равный по величине и обратный по направлению ток. Это подтверждает симметричность зоны Бриллюэна относительно положительных и отрицательных значений квазиимпульсов.

Предположим теперь, что вся зона полностью заполнена электронами, за исключением единственного состояния, характеризуемого волновым вектором \bar{k}_j , которому соответствует скорость \bar{v}_j . Суммарный ток всех электронов в зоне записывается в виде

$$\bar{j} = -\left(\frac{q}{V}\right) \cdot \sum_{j \neq s} \bar{v}_j = -\left(\frac{q}{V}\right) \cdot \sum_s \bar{v}_s + \left(\frac{q}{V}\right) \cdot \bar{v}_j .$$

В данном выражении первое слагаемое равно нулю, так как в нем суммирование ведется по всем состояниям зоны. Поэтому имеем

$$\bar{j} = \left(\frac{q}{V}\right) \cdot \bar{v}_j . \quad (2.4)$$

Из данного выражения следуют три вывода: во-первых, если один электрон убрать из валентной зоны, переместив в зону проводимости, то суммарный ток всех электронов в валентной зоне будет эквивалентен току одной свободной частицы в валентной зоне; во-вторых, эта свободная частица имеет положительный заряд $+e$; в-третьих, эта частица на физическом уровне представляет собой вакантное место, откуда ушел электрон, которое способно перемещаться в валентной зоне (в выражении (2.4) фигурирует ее скорость). Поскольку эта частица представляет собой вакансию электрона на валентной оболочке атома вещества полупроводника, то ее называют дыркой. Ей можно приписать эффективную массу, равную по абсолютной величине эффективной массе того электрона, который занял бы это вакантное состояние.

Ясно, что при наличии некоторого количества положительных дырок плотность соответствующего им электрического тока равна

$$\bar{j} = \left(\frac{q}{V}\right) \cdot \sum_s \bar{v}_s .$$

При наличии и электронов, и дырок в полупроводнике мгновенное значение плотности тока определяется суммой токов, создаваемых движением электронов и дырок:

$$\bar{j} = -\left(\frac{q}{V}\right) \cdot \left(\sum_n \bar{v}_n - \sum_s \bar{v}_s\right),$$

где сумма по n относится к электронам, а сумма по s – к дыркам.

Для описания процессов рождения и уничтожения электронов и дырок удобно пользоваться той же энергетической диаграммой, которую рисовали только для электронов, с той лишь разницей, что в случае дырок энергию для них следует считать возрастающей по вертикали вниз (рис. 11). Движение вниз в валентной зоне по оболочкам атомов означает движение к внутренним оболочкам, близким к ядру атома. Из физических соображений ясно, что для тонизации внутренних оболочек требуется большая энергия, чем ионизация внешней.

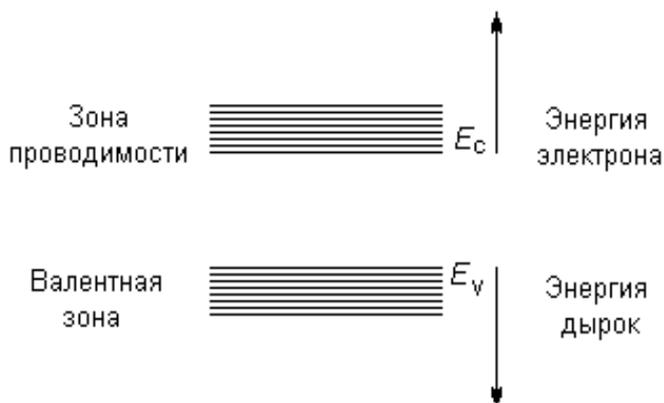


Рисунок 11 - Объединение энергетических диаграмм для электронов и дырок в полупроводнике

Таким образом, в одном представлении можно объединить энергии электронов и дырок, что при переходе к графическому представлению энергетических преобразований в полупроводнике означает: одну и ту же схему энергетических уровней можно использовать как для электронов, так и для дырок, если энергию дырки отсчитывать в направлении, обратном направлению отсчета энергии электрона.

2.6 Движение электронов и дырок в электрическом поле

Напомним, что основной задачей, решаемой в данной главе, является задача объяснения механизма электропроводности кристаллов. В предыдущих параграфах выяснено, что частицами, способными создать электрический ток в кристаллах, являются электроны и дырки, находящиеся в зонах разрешенных энергий. Но как заставить их двигаться по своим зонам и тем самым создать электрический ток, из энергетической диаграммы полупроводника неясно. Для этого рассмотрим, как влияет наличие внешнего электрического поля на величину средней скорости электрона, находящегося в зоне разрешенных значений энергии.

Работа, совершаемая внешним электрическим полем \bar{F} над свободным электроном или дыркой в единицу времени, равна изменению их энергии, определяемому кулоновским взаимодействием поля и движущегося заряда:

$$\frac{\partial E}{\partial t} = \pm q(\bar{F} \cdot \bar{v}) = \pm q \left(\bar{F} \cdot \frac{\partial \bar{x}}{\partial t} \right).$$

Здесь знак «+» соответствует энергии свободного электрона, а знак «-» - свободной дырки. Для простоты будем рассматривать одномерный случай, когда электрическое поле

направлено по оси x и движение частиц также происходит вдоль этой оси. Тогда данное выражение можно переписать следующим образом:

$$\frac{\partial E}{\partial x} \frac{\partial x}{\partial t} = \pm qF \frac{\partial x}{\partial t} \quad (2.5)$$

откуда следует, что

$$\frac{\partial E}{\partial x} = \pm qF \quad (2.6)$$

Отсюда следует, что энергия электронов и дырок во внешнем электрическом поле изменяется с расстоянием в полупроводнике по линейному закону. Скорость изменения энергии определяется только величиной приложенного электрического поля.

Теперь выведем закон изменения квазиимпульса (дрейфовой скорости) свободной частицы в электрическом поле – за счет чего энергия частиц изменяется? Вектор скорости можно найти, продифференцировав выражение для кинетической энергии частицы $E(k)$:

$$v = \frac{1}{\hbar} \frac{\partial E}{\partial k}$$

Поэтому выражение (2.6) для изменения энергии частицы в электрическом поле будет иметь следующий вид:

$$\hbar \frac{\partial k}{\partial t} = \frac{\partial p}{\partial t} = \pm qF,$$

где p – квазиимпульс свободной частицы. Данное соотношение показывает, что под действием электрического поля с течением времени квазиимпульс частицы изменяется пропорционально напряженности поля:

$$p = p_0 \pm qF \cdot t \quad (2.7)$$

Рассмотрим энергетический спектр свободных носителей заряда в постоянном

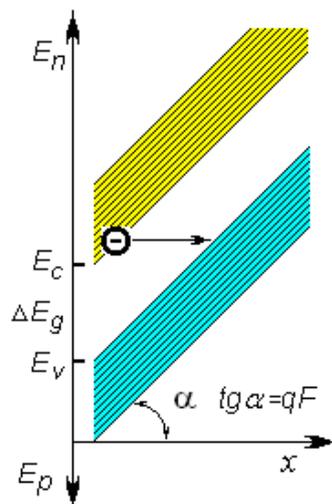


Рисунок.12 – Энергетическая диаграмма для электронов и дырок в полупроводнике при наложении электрического поля.

однородном электрическом поле. В случае, когда электрическое поле медленно изменяется вдоль кристалла, потенциальную энергию электрона можно записать в виде:

$$E_c(x) = E_c(0) + q(\vec{F} \cdot \vec{x}). \quad (2.8)$$

Данное выражение показывает, что при наложении на полупроводник внешнего электрического поля сохраняется представление о зонном энергетическом спектре электрона. Однако зонная диаграмма полупроводника в поле отличается от спектра в отсутствии внешних воздействий на кристалл. Как следует из выражения (2.8), при больших значениях x изменение второго слагаемого может оказаться сравнимым с энергией электрона

в невозмущенном полупроводнике. Поскольку энергия электрона зависит от x по линейному закону, то это означает, что зоны, оставаясь практически неизменными по форме, должны наклониться и принять вид, показанный на рис. 12. Согласно выражению (2.8), тангенс угла наклона энергетических зон определяется величиной приложенного электрического поля F . Из рисунка следует вывод: при действии электрического поля в полупроводнике как таковой запрещенной зоны нет. Действительно, для любого значения энергии, показанного на рис. 12 пунктирной линией, можно найти область пространства, в которой это значение попадает в зону проводимости или в валентную зону. Представление же о границах зон, тем не менее, сохраняет свой смысл.

Из сказанного следует, что во внешнем поле электрон может, не совершая работы, перейти из одной разрешенной зоны в другую, преодолев барьер в виде запрещенной зоны. Вероятность такого перехода зависит от расстояния, разделяющего разрешенные зоны энергий, и величины запрещенной зоны: чем они меньше, тем больше вероятность перехода между зонами без энергетических затрат. Это явление аналогично известному в квантовой механике туннельному переходу электрона сквозь потенциальный барьер. Соответственно и в теории полупроводников данный эффект, вызванный действием электрического поля, называют туннельным эффектом. На этом эффекте основано действие туннельных диодов.

2.7 Механизмы формирования тока в полупроводниках

Дрейфовый ток. При помещении полупроводника p - или n - типа в электрическое поле величиною ε_0 в зонах разрешенных значений энергии возникает направленное движение электронов и дырок: электроны, находясь в зоне проводимости, перемещаются навстречу полю, а дырки в валентной зоне движутся по полю. Поэтому полный ток в полупроводнике равен сумме этих токов:

$$j = j_n + j_p.$$

Каждый из этих токов можно рассчитать исходя из следующих соображений. Если концентрация электронов равна n_0 и они перемещаются навстречу полю со средней скоростью v_n , то переносимый ими заряд за время t через единичную площадку вдоль оси x , перпендикулярную направлению их скорости, будет

$$Q_n = qn_0v_n \cdot t,$$

где q – заряд электрона. Тогда протекающий электронный ток будет иметь плотность

$$j_n = \frac{dQ_n}{dt} = qn_0v_n.$$

Этот ток должен подчиняться закону Ома

$$j_n = \sigma \cdot F_0 = qn_0 \frac{v_n}{\varepsilon_0} F_0 = q\mu_n n_0 F_0, \quad (2.9)$$

где обозначено: $\mu_n = \frac{v_n}{F_0}$ - фундаментальный параметр, называемый подвижностью электронов, который имеет постоянную величину в данном в полупроводнике при заданной его температуре. Поясним обоснованность его введения.

В соответствии с законами классической механики скорость электрона, помещенного в постоянное электрическое поле, должна линейно возрастать как с ростом величины электрического поля, так и временем его воздействия t :

$$v_n = at = \frac{qF_0}{m_n} t.$$

В действительности же в твердых телах имеется большое количество разнообразных структурных дефектов, с которыми при движении электрон сталкивается. Из-за

множественных столкновений скорость движения электрона определяется средним временем между соседними соударениями $\langle \tau_n \rangle$. Поэтому

$$v_n = \frac{q \langle \tau_n \rangle}{m_n} F_0,$$

В результате она остается только линейной функцией величины электрического поля. Поэтому подвижность электронов, как и дырок, оказывается практически постоянной величиной для данного полупроводника:

$$\mu_n = \frac{q \langle \tau_n \rangle}{m_n}.$$

Воспользовавшись законом Ома ($j_n = \sigma_n F_0$), из выражения (2.9) можно найти электропроводность полупроводника, созданную движением свободных электронов:

$$\sigma_n = q \mu_n n_0.$$

Аналогичным образом находится дырочная составляющая электропроводности кристалла, в которой фигурируют подвижность дырок μ_p и их концентрация в валентной зоне p_0 .

Поэтому полная электропроводность полупроводника равна сумме электронной и дырочной электропроводностей:

$$\sigma = q \mu_n n_0 + q \mu_p p_0. \quad (2.10)$$

Теперь обратимся к объяснению возникновения тока проводимости в полупроводниках при приложении к нему электрического поля на основе энергетической диаграммы (рис.13).

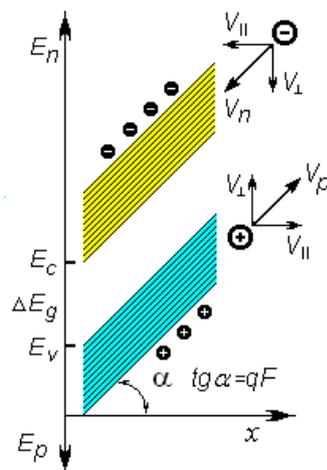


Рисунок 13 – Объяснение возникновения дрейфового тока в полупроводнике при наложении электрического поля разложением векторов скорости

Пусть имеется полупроводник n -типа, донорная примесь которого полностью ионизована. В результате в зоне проводимости находятся свободные электроны с концентрацией n_0 , равной концентрации введенной донорной примеси N_d . В валентной зоне также имеются свободные носители заряда – дырки, но их концентрация меньше концентрации свободных электронов:

$$p_0 = \frac{n_i^2}{N_d}.$$

Пусть теперь к этому полупроводнику приложили электрическое поле, причем его положительный полюс находится на левом торце полупроводникового образца, а отрицательный – на правом. Как установлено в предыдущем параграфе данной главы, наложение электрического поля приводит к наклону энергетической диаграммы полупроводника на угол, определяемый величиной поля. Энергетическая диаграмма

полупроводника в поле показана на рис.13. Здесь E_n , E_p - направления отсчитывания энергий электронов и дырок, x - координата в объеме полупроводника в направлении действия электрического поля, v_n , v_p - скорости дрейфа электронов и дырок, j_n , j_p - дрейфовые токи, создаваемые движением электронов и дырок соответственно. Стремясь понизить свою энергию, электроны будут двигаться налево, а дырки – направо. Имея противоположно направленные скорости и противоположные заряды, создаваемые ими токи, будут направлены в одном направлении, одинаково. Поэтому полный ток дрейфового движения носителей заряда будет равен сумме дрейфовых токов электронов и дырок.

Диффузионный ток. Электрический ток в твердых телах может возникнуть не только под действием электрического поля. Направленное движение электронов и (или) дырок может возникнуть при неоднородном их распределении в пространстве. В результате из мест с высокой концентрацией свободные носители устремятся в места, где их концентрация меньше. Возникающий при этом ток называется диффузионным, т.к. причина возникновения движения - диффузия носителей заряда. Величина диффузионного тока пропорциональна градиенту концентрации свободных носителей заряда. В одномерном случае диффузионные токи электронов и дырок равны:

$$j_n^{(диф)} = qD_n \frac{\partial n}{\partial x}, \quad j_p^{(диф)} = qD_p \frac{\partial p}{\partial x}, \quad (2.11)$$

где D_n , D_p – коэффициенты диффузии электронов и дырок соответственно. Они связаны с подвижностью свободных носителей заряда соотношениями:

$$D_n = \frac{kT}{q} \mu_n, \quad D_p = \frac{kT}{q} \mu_p.$$

Поскольку коэффициенты диффузии являются константами вещества, то, зная градиенты концентраций, по выражениям (2.11) можно найти величины диффузионных токов.

ВОПРОСЫ ПО ГЛАВЕ 2

1. В чем различия физических моделей твердого тела Зоммерфельда и Блоха?
2. В чем различие энергетической диаграммы для электронов, полученной Зоммерфельдом и Блохом?
3. Что такое квазиимпульс и какой физический процесс в твердом теле он описывает?
4. Почему значения квазиимпульса электрона оказываются дискретными?
5. Как называется область возможных значений квазиимпульса и чем задаются её границы?
6. Какой физический смысл заложен в вводимые в полупроводниках дополнительные к электронам заряженные частицы - дырки?
7. Что происходит с энергетической диаграммой для электронов и дырок при наложении электрического поля на кристалл?
8. Что такое подвижность электронов и какой физический процесс они отражают?
9. Что такое дрейфовый ток электронов и чем он отличается от дрейфового тока дырок?
10. Почему дрейфовые токи электронов и дырок текут в одинаковом направлении, а диффузионные токи в противоположные?

3 ВЗАИМОДЕЙСТВИЕ ОПТИЧЕСКОГО ИЗЛУЧЕНИЯ С ТВЕРДЫМИ ТЕЛАМИ

3.1 Основные параметры поглощения

При рассмотрении поглощения электромагнитного излучения в полупроводниках будем вести речь только о слабых потоках энергии. Слабая электромагнитная волна не изменяет энергетический спектр носителей заряда, а лишь создает новые пары «электрон-дырка» или вызывает перераспределение носителей по состояниям зоны проводимости или валентной зоны. Поэтому параметры, характеризующие оптические свойства среды, не зависят от интенсивности падающего излучения. Будем также считать, что длина падающей электромагнитной волны значительно превышает период трансляции решетки полупроводника a – расстояние между соседними атомами.

1. При падении излучения интенсивностью I_0 на полупроводник (рис. 14) в его объем проникает не весь поток, т.к. часть излучения отражается от поверхности. Для характеристики отражения света вводится **коэффициент отражения**. Обозначим интенсивность излучения через I – количество энергии, переносимое лучом через единичную площадку в единицу времени. Если падающая интенсивность I_0 , а отраженная – I_R , то коэффициент отражения определяется как

$$R = \frac{I_R}{I_0}.$$

Зависимость коэффициента отражения R от длины волны падающего излучения λ – $R(\lambda)$ называется спектром отражения твердого тела.

Экспериментальные исследования показывают, что часть излучения, проникающая в объем полупроводника, по мере распространения в его объеме уменьшается. Причиной уменьшения интенсивности излучения с координатой в объеме является взаимодействие излучения с атомами и решеткой полупроводника, со свободными носителями и т.д. Это уменьшение описывается законом Бугера-Ламберта:

$$I(x) = I_0(1 - R) \cdot \exp(-\alpha \cdot x),$$

где α – некоторый параметр среды, который имеет размерность обратной длины и называется коэффициентом поглощения. Для того чтобы выяснить физический смысл коэффициента поглощения, будем считать показатель экспоненты в законе равным единице: $\alpha \cdot x_0 = 1$. Отсюда $x_0 = 1/\alpha$ – это расстояние от освещаемой поверхности полупроводника, при прохождении которого интенсивность излучения уменьшается в e - раз (рис. 14). Обычно это расстояние обозначают $l_\phi = x_0$ и называют длиной свободного пробега фотона, т.е. это расстояние, которое проходит фотон в твердом теле до момента поглощения: $l_\phi = \alpha^{-1}$.

Почему же интенсивность уменьшается по экспоненциальному закону, а не по линейному? Чтобы ответить на этот вопрос, начнем с построения физической картины поглощения. Для этого мысленно разделим весь полупроводник на тонкие слои и будем

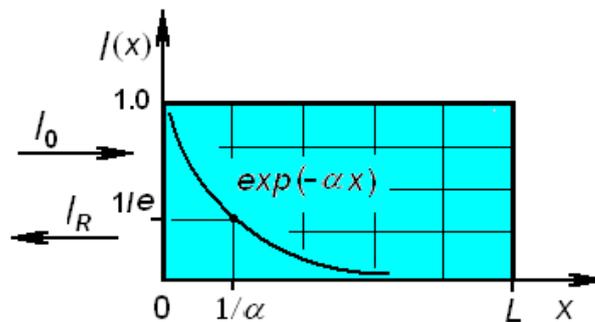


Рисунок 14 – Разделение падающего потока фотонов на потоки отражённых и прошедших в полупроводник

определять долю поглощенной интенсивности света в выбранном слое по отношению к падающей на этот слой интенсивности. Коэффициент поглощения в каждом слое предполагаем одинаковым. Тогда можно считать, что в каждом слое поглощается одинаковая доля мощности, например, 20% от падающей интенсивности. Поскольку к каждому очередному слою полупроводника подходит интенсивность по величине меньшая, чем к предыдущему, то и абсолютное значение поглощенной в рассматриваемом слое интенсивности будет меньше, чем в предыдущем. Значит, с увеличением расстояния от облучаемой поверхности величина поглощенной интенсивности будет уменьшаться. Если бы в каждом слое поглощалась одна и та же интенсивность, то закон уменьшения интенсивности излучения с координатой был бы линейным. В данном случае уменьшение интенсивности от слоя к слою непостоянно. Следовательно, закон уменьшения интенсивности с расстоянием должен быть нелинейным.

Для количественного ответа на поставленный вопрос выведем математический закон уменьшения интенсивности излучения по мере распространения в полупроводнике, исходя из квантовой теории излучения. Для этого выразим интенсивность I через число фотонов в падающем излучении. Пусть n – число фотонов в единичном объеме светового потока, тогда $n \cdot c$ есть поток фотонов (количество фотонов, переносимых через единичную площадку за единицу времени). Здесь c – скорость света. Каждый фотон несет энергию $\hbar\omega$, поэтому интенсивность света равна:

$$I = n \cdot c \cdot \hbar\omega.$$

Обозначив $n \cdot c = q$, получим:

$$I = q \cdot \hbar\omega. \quad (3.1)$$

Из наблюдений известно, что интенсивность излучения с координатой x уменьшается. Зададимся вопросом: за счет изменения какого параметра это может происходить? Из выражения (3.1) нетрудно видеть, что таковым может быть только параметр $q = n \cdot c$, т.к. ни энергия кванта $\hbar\omega$, ни скорость света c в принципе не могут зависеть от координаты. Следовательно, выражение (3.1) можно переписать в виде:

$$I(x) = q(x) \cdot \hbar\omega. \quad (3.2)$$

Из выражения (3.2) следует достаточно важный вывод: с точки зрения фотонной теории света, уменьшение его интенсивности при распространении в полупроводнике возможно только из-за уменьшения числа фотонов. Уменьшение числа фотонов может происходить только за счет их поглощения, т.к. фотоны не могут покинуть полупроводник (эта возможность в данной главе не рассматривается). Поглощение фотонов, вероятно, осуществляется некими центрами поглощения. Как показывает эксперимент, существование их в чистых и структурно совершенных кристаллах наталкивает на мысль, что этими центрами с большой долей вероятности могут выступать атомы основного вещества твердого тела. Примем эту гипотезу о возможной природе центров поглощения.

Чтобы сделать какие-либо выводы относительно причин и механизма поглощения фотонов твердым телом, рассмотрим случай, когда фотон попадает в площадь поперечного сечения атома полупроводника. Пусть σ – площадь поперечного сечения поглощающего центра, т.е. это вероятность исчезновения фотона при его попадании в одиночную мишень сечением σ . Если N – это число центров, поглощающих фотоны в единичном объеме, то уменьшение потока фотонов на величину δq при прохождении им слоя полупроводника толщиной δx как это следует из физических соображений, должно быть пропорционально площади поперечного сечения поглощающего центра σ , концентрации N , величине потока и толщине слоя:

$$-\delta q = \sigma N \cdot q(x) \delta x.$$

Здесь $\sigma \cdot N$ – вероятность поглощения фотонов всеми центрами в единичном объеме. Из этого уравнения найдем

$$\frac{\partial q}{q(x)} = -\sigma N \cdot dx.$$

Интегрирование этого уравнения дает:

$$q(x) = q(0) \cdot \exp(-\sigma N \cdot x).$$

Если обе части этого решения умножить на $\hbar\omega$, то получим:

$$I(x) = I(0) \cdot \exp(-\sigma N \cdot x).$$

Сравнивая это выражение с выражением (3.1) находим, что коэффициент поглощения равен: $\alpha = \sigma N$. Нетрудно убедиться, что он имеет размерность обратной длины.

Если считать, что квант света имеет энергию $\hbar\omega \geq \Delta E_g$, то поперечное сечение поглощения атомами основного вещества можно считать равным поперечному сечению самих атомов: $\sigma \approx S_{am}$. Это означает, что достаточно фотону попасть в атом, чтобы быть поглощенным. Поэтому можно считать, что $\sigma \approx (10^{-16} \div 10^{-17}) \text{ см}^2$, $N \approx 10^{22} \text{ см}^{-3}$. Тогда

$$\alpha \approx (10^5 \div 10^6) \text{ см}^{-1}, l_{\phi} \approx (10^{-5} \div 10^{-6}) \text{ см} = (0.1 \div 0.01) \text{ мкм}.$$

Отсюда следует важный физический вывод: излучение в полупроводнике поглощается в тонком приповерхностном слое, а объем полупроводника в поглощении излучения никак не участвует. Следовательно, поглощение в полупроводниках должно быть чувствительным к состоянию приповерхностного слоя и самой поверхности полупроводника.

Если в твердом теле обнаруживается несколько механизмов поглощения со своими значениями σ_i и N_i , действующими независимо друг от друга, то результирующий коэффициент поглощения излучения будет равен алгебраической сумме коэффициентов поглощения различных механизмов

$$\alpha = \sum_i \alpha_i = \alpha(\lambda),$$

т.к. вероятности независимых событий складываются. Здесь сумма берется по различным сортам поглощающих центров.

3.2. Основные механизмы поглощения излучения

В приведенном выше объяснении физической сущности закона Бугера-Ламберта дана феноменологическая картина поглощения фотона, показавшая его зависимость от концентрации поглощающих атомов N и их площади σ . Ясно, что значение коэффициента поглощения и его зависимость от длины волны поглощаемого излучения определяются конкретным физическим механизмом взаимодействия фотона и полупроводника. В реальном полупроводнике возможны различные механизмы поглощения. Перечислим и кратко охарактеризуем их, а затем рассмотрим более подробно.

1. **Собственное (фундаментальное) поглощение.** Этот вид поглощения обусловлен взаимодействием квантов света с атомами основного вещества, например, кремния или германия, и потому называется собственным. Он имеет место при энергии кванта, равной или больше ширины запрещенной зоны полупроводника: $\hbar\omega \geq \Delta E_g$.

2. **Примесное поглощение.** Данный вид поглощения имеет место, если энергия кванта равна или несколько превышает энергетические зазоры между зоной разрешенных энергий и уровнями примеси донорного (E_d) или акцепторного (E_a) типа:

$\hbar\omega \geq E_c - E_d, E_a - E_v$. При этом будет совершаться переход электронов с примесного донорного уровня в зону проводимости или из валентной зоны на примесный акцепторный уровень, т.е. энергия кванта излучения идет не на возбуждение атомов основного вещества, а на ионизацию примесей в нем.

3. **Поглощение свободными носителями заряда.** В этом виде поглощения квант сталкивается со свободным электроном или дыркой и, как в случае соударения с атомами, передает подвижной частице свою энергию и квазиимпульс. Приобретенный электроном при соударении квазиимпульс ускоряет носители заряда в направлении падающего света. В результате действия этого вида поглощения создается направленный поток свободных носителей заряда. Этот тип поглощения называют еще внутризонным. Оно наблюдается только в полупроводниках со сложной структурой валентной зоны или зоны проводимости и связано с переходом электронов (дырок) с одной зоны $E(\vec{k})$ в другую зону, отличающуюся кривизной указанной зависимости.

4. **Решеточное поглощение.** В этом случае квант света взаимодействует с колебаниями решетки; результатом этого является увеличение количества акустических и оптических фононов кристаллической решетки.

5. **Экситонное поглощение** света характерно тем, что энергия поглощаемого излучения идет на образование специфического комплекса – экситона: связанной пары «электрон – дырка», когда рожденные светом частицы не разлетаются в разные стороны, а остаются рядом до момента их рекомбинации.

6. **Плазменное поглощение** света. Этот вид поглощения создается совокупностью электронов и дырок ($n_{\text{электр}} = n_{\text{дырок}}$). Энергия поглощаемых фотонов идет на возбуждение плазменных колебаний.

Из приведенного рассмотрения физических процессов, определяющих взаимодействие оптического излучения с твердым телом, следует, что спектр поглощения $R(\lambda)$ должен зависеть от внешних воздействий на полупроводник, которые способны изменить состояние основных и примесных атомов полупроводника, спектр колебаний кристаллической решетки. Такими воздействиями являются температура тела, давление, электрическое и магнитное поле и т.д.

3.3 Собственное поглощение. Прямые и непрямые переходы

Рассмотрим подробно феноменологическую модель процесса поглощения кванта, поскольку детальное описание этого процесса в учебниках зачастую отсутствует.

Поглощение кванта наступает в случае, если он попадает в площадь поперечного сечения атома (как пуля в мишень), в любое его место: в центр, в периферию и т.д. Размеры ядра атома малы и поэтому основная его площадь занята электронами, вращающимися по своим орбитам. Поэтому взаимодействие кванта осуществляется преимущественно с электронами. Первыми на пути кванта встречаются внешние (валентные) электроны – они вращаются по внешним орбитам и защищают внутренние орбиты, проникнуть к которым может только высоко энергетичный квант (это особый случай, и мы его не рассматриваем). Поэтому взаимодействие оптического излучения с веществом определяется валентными

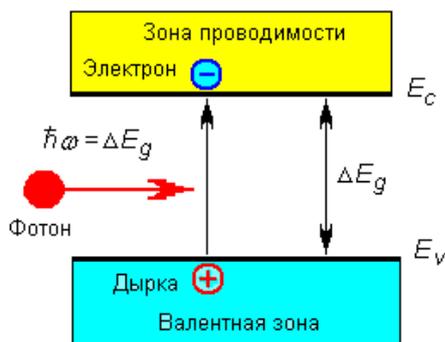


Рисунок 15 – Схема поглощения кванта света при собственном механизме взаимодействия фотона и атома

электронами. Если частота кванта находится в резонансе с частотой перехода валентных электронов ($\hbar\omega = \Delta E_g$), то столкновение кванта и атома, различающихся по массе на несколько порядков, носит характер неупругого соударения. Отдав атому всю свою энергию и весь квазиимпульс ($m_{\text{фот}} \cdot \nu$), фотон исчезает, т.к. в состоянии покоя ($m_{\text{фот}} \cdot \nu = 0$) он имеет нулевую массу ($m_{\text{фот}} = 0$). Характеризуя это соударение, говорят, что квант «поглотился» атомом. Получив энергию кванта, валентный электрон переходит в новое энергетическое состояние: он отрывается от атома и переходит в межатомное пространство, но при этом имеет нулевую скорость. Если энергия кванта больше ширины запрещенной зоны, то избыток энергии над ΔE_g идет на приобретение свободным электроном кинетической энергии.

Что же происходит с квазиимпульсом атома? Полученный от кванта квазиимпульс по своей величине обычно пренебрежимо мал по сравнению с собственным квазиимпульсом атома и потому поглощение кванта практически никак не сказывается на кинетических параметрах теплового движения атома. Хотя «удар» от кванта света есть, но он становится заметным при поглощении излучения легкой подвижной частицей, например, электроном, или при поглощении атомом высокоэнергетической частицы.

Суммируя описанное выше, на энергетической диаграмме процесс поглощения кванта излучения и выход электрона в межатомное пространство изображают так (рис. 16): исчезновение кванта и переход электрона из валентной зоны в зону проводимости. При этом создается электронно-дырочная пара.

При собственном поглощении излучения необходимо различать два случая переходов электронов из валентной зоны в зону проводимости:

- переход электрона осуществляется только под действием фотона. Такой переход называется прямым;
- переход в зону проводимости совершается под действием фотона и частично под действием колебаний решетки – фононов. Такой переход называется непрямым.

Первый тип переходов реализуется в полупроводниках, у которых минимум энергии в зоне проводимости совпадает с максимумом энергии электрона в валентной зоне (рис. 9): при $k = 0$ имеет место совпадение экстремумов зон. Второй тип переходов имеет место, если эти экстремумы не совпадают (рис. 10). Рассмотрим подробнее оба типа переходов.

Прямые переходы. При прямых переходах импульс электрона фактически не изменяется. Из закона сохранения квазиимпульса следует, что разность импульсов электрона в разных зонах должна равняться импульсу фотона, т.к. при них электрон взаимодействует только со светом. Действительно: $\bar{k}_c - \bar{k}_v = \bar{k}_{\text{фотон}} = \frac{\hbar\omega}{c} \approx 0$, т.к. c велико, т.е. $k_v \approx k_c$ – импульсы в зонах примерно равны. Значит, на энергетической диаграмме прямые переходы можно показать так: переход из максимума валентной зоны в минимум (см. рис. 9) зоны проводимости без изменения квазиимпульса. Для получения прямого перехода необходима энергия фотона не менее ΔE_g . Поэтому граничную длину волны поглощаемого излучения данным способом можно найти так: $\hbar\omega_{кр} = \Delta E_g$ или $hc / \lambda_{кр} = \Delta E_g$, откуда находим: $\lambda_{кр} = hc / \Delta E_g$. Здесь $\omega_{кр}$ и $\lambda_{кр}$ – круговая частота и длина волны излучения, соответствующего красной границе фотоэффекта. Подставив численные значения входящих сюда констант, получим:

$$\lambda_{кр} (\text{мкм}) = \frac{1,24}{\Delta E_g (\text{эВ})}.$$

Так, в кремнии при комнатной температуре $\Delta E_g = 1,12$ эВ, поэтому $\lambda_{кр} \cong 1,1$ мкм. Это означает, что если на образец кремния будет падать излучение с длиной волны меньше $\lambda_{кр}$,

то оно будет поглощать за счет собственного поглощения. Но при длине волны падающего излучения большей $\lambda_{кр}$ поглощение прекратится – энергии кванта окажется недостаточно для перевода электрона из валентной зоны в зону проводимости.

Непрямые переходы. Если переход не прямой, то разность квазиимпульсов начального и конечного состояний электрона (рис. 9) должна быть скомпенсирована участием в процессе поглощения дополнительной частицы, вероятнее всего фонона. Фонон – это квант теплового колебания решетки. С его помощью тепловое движение атомов решетки представляют, как результат столкновения регулярно расположенных в узлах решетки атомов с микрочастицами коими и являются фононы. Фононы имеют разные энергии и разные квазиимпульсы, что позволяет корректно описать хаотическое движение атомов. Значит

$$\bar{k}_c - \bar{k}_v = \bar{k}_{\text{фотон}} \pm \bar{k}_{\text{фонон}}.$$

Но при поглощении кванта света, кроме квазиимпульса, электрон должен получить (или отдать) от фонона и энергию – $E_{\text{фонон}}$ и закон сохранения энергии будет выглядеть так:

$$\hbar\omega \pm E_{\text{фонон}} = \Delta E,$$

Вероятность таких переходов с поглощением или излучением фонона будет меньше, чем вероятность прямых переходов, т.к. при непрямых переходах необходима встреча трех частиц: электрона, фотона и фонона. В связи с этим непрямые переходы имеют место, когда прямые переходы по каким-либо причинам невозможны. На энергетической диаграмме непрямые переходы показываются стрелкой, выходящей из максимума валентной зоны и заканчивающийся в абсолютном минимуме зоны проводимости (рис. 10). Если на такой полупроводник падает свет с энергией кванта $\hbar\omega = \Delta E_g$, то прямые переходы в нем невозможны, но будут происходить непрямые переходы, при которых электрон должен изменить свой квазиимпульс k от 0 до k_1 , а также увеличить энергию E на величину ΔE_g .

Непрямые переходы могут иметь место и в прямозонном полупроводнике, если его зона проводимости в значительной степени занята электронами, что возможно в полупроводниках с высокой концентрацией свободных частиц в зоне разрешенных энергий.

3.4 Примесное поглощение излучения

Этот тип поглощения света реализуется в полупроводниках, в которых есть донорные (E_d) или акцепторные (E_a) примеси, имеющие энергетические уровни в запрещенной зоне, локализованные вблизи ее краев. Если будет падать свет с $\hbar\omega \geq E_c - E_d$, $\hbar\omega \geq E_a - E_v$, то электроны с E_d будут переходить в E_c , а дырки из валентной зоны – на E_a (рис. 16). Как следствие этого положения дел, при поглощении света атомами примеси в полупроводнике

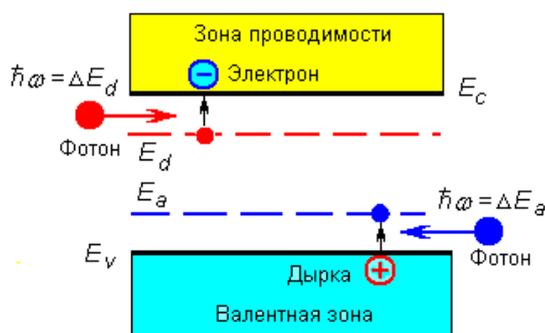


Рисунок 16 - Схема поглощения кванта света при примесном механизме взаимодействия фотона и атома

будут образовываться неподвижные ионы с фотоиндуцированным положительным или отрицательным зарядом. Заметим, что под действием света возможны не только переходы с образованием ионов, но и обратные: электронов из E_C на E_d дырок с E_a в валентную зону. Это поглощение будет наблюдаться при низких температурах, что скажется на степени ионизации примесных уровней – акцепторные уровни не должны быть полностью заполнены электронами, а донорные – полностью ионизованными. В остальном процесс примесного поглощения протекает аналогично описанному выше для собственного поглощения, но с заменой атомов основного вещества на атомы примеси. Более того, в результате примесного поглощения в полупроводнике образуется всего лишь один тип свободного носителя: электрон при поглощении донорной примесью или дырка – при поглощении кванта акцепторной примесью, тогда как при собственном поглощении генерируется пара «электрон – дырка». Этот тип поглощения имеет непрерывный спектр, т.е. при изменении длины волны света от значения, соответствующего переходу «зона – примесный уровень» и выше, коэффициент примесного поглощения остается постоянным.

Существует важное количественное различие между собственным и примесным видами поглощения: так как концентрация примесей (от 10^{12} см⁻³) значительно меньше концентрации атомов полупроводника вещества (10^{22} см⁻³), вероятность столкновения кванта с атомом примеси во много раз ниже вероятности столкновения кванта с атомом основного вещества. Значит, коэффициент примесного поглощения обычно имеет величину ($10^1 - 10^2$ см⁻¹), что на несколько порядков меньше коэффициента собственного поглощения. Значит, примесное поглощение наблюдаем при исключении собственного поглощения.

3.5 Поглощение свободными носителями (внутризонное поглощение)

Поглощение свободными носителями заряда – это один из основных механизмов поглощения оптического излучения в полупроводниках. В данном случае взаимодействие кванта света осуществляется с носителем, который свободно движется внутри зоны и реагирует на внешние воздействия. Поглощение свободными носителями характеризуется монотонным, часто бесструктурным спектром поглощения $\alpha(\lambda)$, описываемым законом $\alpha(\lambda) \sim \lambda^p$, где λ – длина волны фотона, а показатель p может принимать значения в пределах от 1,5 до 3,5. Здесь при поглощении фотона электрон совершает переход в состояние с большей энергией в пределах той же энергетической долины. Для выполнения закона сохранения квазиимпульса при таком переходе требуется дополнительная частица, в качестве которой выступает фонон или ионизованная примесь.

Рассеяние света на акустических фононах приводит к поглощению, меняющемуся по закону $\alpha_A(\lambda) \sim \lambda^{1.5}$. Рассеяние на оптических фононах дает зависимость $\alpha_\Phi(\lambda) \sim \lambda^{2.5}$, тогда как рассеяние на ионизованных примесях может дать зависимость $\alpha_{II}(\lambda) \sim \lambda^{3.0 \div 3.5}$. В общем случае реализуются все типы рассеяния и результирующий коэффициент поглощения свободными носителями представляет собой сумму трех слагаемых:

$$\alpha_\Sigma(\lambda) = \alpha_A(\lambda) + \alpha_\Phi(\lambda) + \alpha_{II}(\lambda) = A\lambda^{1.5} + B\lambda^{2.5} + C(N)\lambda^{3.5},$$

где A , B , C – константы, характеризующие вклад того или иного механизма рассеяния избытка энергии в суммарный коэффициент поглощения $\alpha_\Sigma(\lambda)$.

Таким образом, в зависимости от типа доминирующего механизма рассеяния избыточной энергии кванта характер спектральной зависимости коэффициента поглощения на свободных носителях заряда будет различен: от $\lambda^{1.5}$ до $\lambda^{3.5}$.

Спектр поглощения света в полупроводнике при поглощении свободными носителями и различной концентрации примеси (N_1 , N_2 , N_3), а также действии механизма поглощения на свободных носителях заряда имеет вид, показанный на рис. 17. В дополнение к этому заметим, что поглощение кванта света и изменение энергетического состояния свободной частицы происходит в соответствии с энергетической диаграммой той зоны разрешенных энергий, какая

свободная частица поглощает квант: электрон или дырка. При поглощении кванта света свободным электроном он получает от фотона квазиимпульс, равный k_1 . Для его поглощения электрон должен также получить дополнительную энергию, определяемую долинным спектром полупроводника. Эту энергию он получает в виде фонона.

Поглощение дыркой может произойти с участием одной из трех валентных зон: $E_V^{(1)}$, $E_V^{(2)}$, $E_V^{(3)}$. Варианты изменения энергетического состояния дырки показаны на рис.18 стрелками с номерами 2 и 3. В этом случае переход дырки с одной параболы $E(\vec{k})$ на другую проявится как пик величины коэффициента поглощения. Эти переходы носителей заряда называют еще внутризонными переходами.

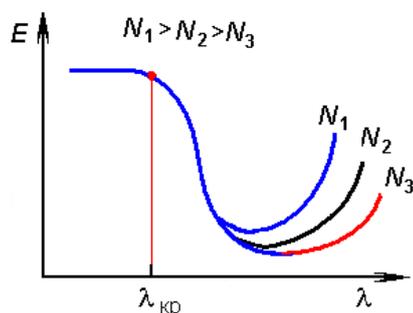


Рисунок 17 – Спектр поглощения свободными носителями заряда

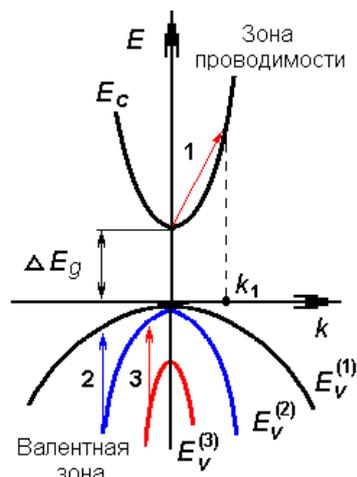


Рисунок 18 – Спектр внутризонного поглощения дырками

3.6 Решеточное поглощение

Этот тип поглощения реализуется при взаимодействии кванта света с тепловыми колебаниями решетки твердого тела, в результате которого энергия кванта переходит в колебательную энергию атомов или ионов. Механизмом взаимодействия кванта света и ионов решетки является кулоновское взаимодействие заряженной частицы (иона решетки) и электрического поля кванта с силой, равной $F = eE$, где E – напряженность электрического поля в световой волне, e – заряд иона. Если частоты колебания в кванте света и иона в узле решетки совпадают, то энергия кванта полностью передается решетке, увеличивая амплитуду колебания атома соответствующей частоты. В случае же атомной решетки взаимодействие кванта света и атомов осуществляется за счет взаимодействия электрического поля кванта с ионной компонентой связи между атомами решетки.

ВОПРОСЫ ПО ГЛАВЕ 3

1. Какими параметрами описывается взаимодействие излучения с веществом?
2. Какими характеристиками описывается взаимодействие излучения с веществом?
3. Каков физический смысл коэффициента поглощения излучения?
4. Что описывает закон Бугера – Ламберта?
5. Как происходит собственное поглощение при прямых переходах?
6. Как происходит собственное поглощение при непрямах переходах?
7. Как происходит примесное поглощение излучения?
8. В чём различия между собственным и примесным поглощениями?
9. Как происходит поглощение на свободных электронах?
10. Как происходит решёточное поглощение излучения?

4 ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

4.1. Основные понятия и параметры

В полупроводниках, в отличие от металлов, под влиянием внешних воздействий (освещения, электрического поля и т.д.) концентрации электронов и дырок могут изменяться во много раз. Это приводит к ряду специфических явлений, которые лежат в основе действия многих полупроводниковых приборов.

При нарушении термодинамического равновесия концентрации электронов и дырок в зонах (n и p) изменяются по сравнению с их равновесными значениями n_0 и p_0 за счет появления в зонах неравновесных носителей заряда с концентрациями $\delta n = n - n_0$ и $\delta p = p - p_0$. Установление концентраций в зонах определяется действием двух противоположно направленных процессов: генерации электронов и дырок, а также их рекомбинации. Существует несколько разновидностей процессов генерации (световая,

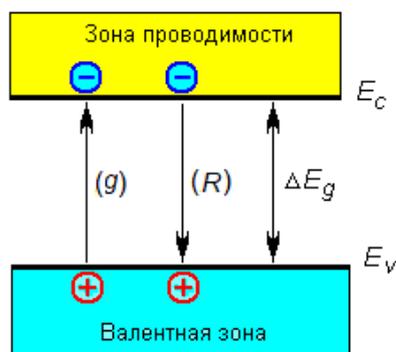


Рисунок 19 – Межзонные генерация и рекомбинация носителей заряда

тепловая и т.д.) и рекомбинации (тепловой заброс носителей заряда «зона-зона», тепловой заброс носителей заряда с участием примесного уровня «зона – уровень – зона» и т.д.). Суммарные скорости этих процессов (количество генерируемых или рекомбинирующих частиц в единичном объеме в единицу времени) принято обозначать как g и R , соответственно (рис. 19). Это два противоположно направленных процесса, равенство скоростей протекания этих противоположно направленных процессов приводит к установлению термодинамического равновесия носителей заряда в зонах разрешенных энергий полупроводника.

Время жизни неравновесных носителей. Пусть под влиянием внешнего воздействия в единице объема полупроводника в единицу времени возникает g_n электронов проводимости и соответственно g_p дырок в валентной зоне. Скорости генерации g_n и g_p будем считать постоянными по всему объему полупроводника, хотя и не обязательно равными друг другу. Пусть далее R_n и R_p – скорости исчезновения электронов и дырок в результате их рекомбинации. Если в полупроводнике нет электрического тока, то суммарная скорость изменения неравновесных концентраций электронов и дырок в зонах будет определяться скоростями их генерации и рекомбинации (рис. 21):

$$\frac{\partial \delta n}{\partial t} = g_n - R_n; \quad \frac{\partial \delta p}{\partial t} = g_p - R_p$$

Здесь функции g_n и g_p описывают генерацию за счет внешнего воздействия и не учитывают переходы, вызванные тепловым движением, которые учитываются в R_n и R_p .

Для описания кинетики (развития процессов во времени) неравновесных электронных процессов вводят понятие среднего времени жизни неравновесных электронов в зоне

проводимости τ_n и дырок в валентной зоне τ_p , которые определяются через скорости рекомбинации электронов и дырок:

$$R_n = \frac{n - n_0}{\tau_n}, \quad R_p = \frac{p - p_0}{\tau_p}$$

или иначе: $1/\tau_n$ – это вероятность исчезновения одного избыточного электрона из зоны проводимости в единицу времени в результате рекомбинации с дыркой. Аналогично: $1/\tau_p$ – вероятность рекомбинации одной дырки в единицу времени.

Пользуясь понятиями времени жизни носителей заряда, уравнения кинетики неравновесных концентраций электронов и дырок в однородном образце, в котором нет тока проводимости, можно переписать в виде:

$$\frac{\partial \delta n}{\partial t} = g_n - \frac{\delta n}{\tau_n}; \quad \frac{\partial \delta p}{\partial t} = g_p - \frac{\delta p}{\tau_p}. \quad (4.1)$$

Стационарные концентрации неравновесных носителей заряда $(\delta n)_s$ и $(\delta p)_s$, устанавливающиеся после длительного воздействия внешней генерации, можно найти, если в (5.1) приравнять к нулю левые части. Действительно, стационарное состояние характеризуется неизменностью концентраций свободных носителей в зонах, что требует в выражениях (4.1) положить все производные по времени обратить в нуль. В результате этого упрощения из (4.1) можно найти выражения для стационарных концентраций свободных носителей заряда в зонах:

$$(\delta n)_s = g_n \tau_n; \quad (\delta p)_s = g_p \tau_p. \quad (4.2)$$

Теперь рассмотрим кинетику изменения концентраций носителей заряда. В простейшем случае, когда τ_n и τ_p не зависят от n и p , интегрирование кинетических уравнений (4.1) с учетом выражений (4.2) дает:

$$\delta n(t) = g_n \tau_n - C_n \exp\left(-\frac{t}{\tau_n}\right); \quad \delta p(t) = g_p \tau_p - C_p \exp\left(-\frac{t}{\tau_p}\right). \quad (4.3)$$

Здесь C_n, C_p – постоянные интегрирования, определяемые из начальных условий: если перед освещением полупроводник, находился в термодинамическом равновесии и затем в момент времени $t = 0$ включено внешнее воздействие, создающее генерацию носителей заряда, то при $t = 0$ будем иметь $\delta n = 0$. Использование данного начального условия дает возможность найти постоянные интегрирования:

$$C_n = g_n \tau_n = (\delta n)_s; \quad \delta n(t) = (\delta n)_s \cdot \left[1 - \exp\left(-\frac{t}{\tau_n}\right)\right];$$

$$C_p = g_p \tau_p = (\delta p)_s; \quad \delta p(t) = (\delta p)_s \cdot \left[1 - \exp\left(-\frac{t}{\tau_p}\right)\right]. \quad (4.4)$$

Ход зависимостей вида (5.4) показан на рис. 20 сплошной линией. Если в некоторый момент времени $t = t_1$ генерация выключается, то для времен $t \geq t_1$, $g_n = 0$ и начальное условие для определения константы принимает вид: $t = t_1$, $\delta n = (\delta n)_1$. Тогда

$$C_n = -(\delta n)_1 \cdot \exp\left(-\frac{t_1}{\tau_n}\right) \quad \text{и} \quad \delta n(t) = (\delta n)_1 \cdot \exp\left(-\frac{(t-t_1)}{\tau_n}\right). \quad (4.5)$$

По аналогии можно записать:

$$\delta p(t) = (\delta p)_1 \cdot \exp\left(-\frac{(t-t_1)}{\tau_p}\right).$$

Ход этой зависимости показан на рис. 20 пунктирной линией. Как следует из рисунка,

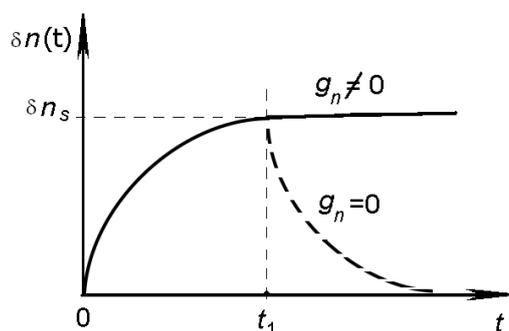


Рисунок 20 – Кинетика фотопроводимости при непрерывном освещении

выключение генерации носителей заряда приводит к тому, что избыточные концентрации электронов и дырок за счет рекомбинации будут стремиться своим равновесным значениям. Согласно выражениям (4.4), и (4.5) при постоянных значениях параметров τ_n и τ_p установление избыточных концентраций электронов и дырок в зонах разрешенных энергий, а также их исчезновение описываются экспоненциальным законом. Поэтому по аналогии с определением физического смысла коэффициента поглощения можно определить физический смысл констант τ_n и τ_p как: промежуток времени, в течение которого неравновесная концентрация электронов (дырок) при включении или выключении источника генерации носителей заряда увеличивается или уменьшается в e - раз. В большинстве практических случаев $\tau_n \approx \tau_p$.

4.2. Фотопроводимость полупроводников: собственная и примесная

Простейший способ создания неравновесных носителей заряда состоит в освещении полупроводника электромагнитным излучением. Возникновение неравновесных носителей проявляется в изменении электропроводности кристалла – в таких случаях говорят о возникновении фотопроводимости. Под фотопроводимостью полупроводника понимают изменение его проводимости, вызванное освещением полупроводника.

Электронные переходы при оптической генерации могут быть различными. Если энергия фотонов $\hbar\omega > \Delta E_g$, то неравновесные электроны и дырки образуются в результате возбуждения электронов, переходящих из валентной зоны в зону проводимости, а также дырок, остающихся в валентной зоне. Это так называемая собственная оптическая генерация и, соответственно, собственная фотопроводимость. Обратный процесс есть прямая рекомбинация «свободный электрон – свободная дырка» (рис. 22). В результате оптической генерации в области собственного поглощения в полупроводнике образуется пара свободных носителей заряда противоположных знаков.

Однако при наличии достаточного количества примесей в полупроводнике фотопроводимость может возникать также при $\hbar\omega < \Delta E_g$. На рис. 21 показана энергетическая диаграмма полупроводника, содержащего глубокую акцепторную примесь. При падении света с энергией кванта $\hbar\omega \geq E_c - E_d$ происходит возбуждение электронов, находящихся на валентной оболочке донорной примеси, в зону проводимости. Это случай примесного поглощения и, соответственно, примесной фотопроводимости. В случае, когда энергия кванта падающего светового излучения $\hbar\omega \geq E_a - E_v$, возможна генерация дырок в валентной зоне. Оба случая –

случаи монополярной проводимости. Поглощение в собственной полосе длин волн на несколько порядков больше поглощения в примесной области.

Скорость оптической генерации связана с коэффициентом поглощения света. Пусть есть монохроматический световой поток, рассчитанный на единицу поверхности, на расстоянии x от освещаемой поверхности полупроводника, а α – коэффициент поглощения излучения. Тогда число фотонов, поглощаемых в единицу времени единичной площадкой в слое,

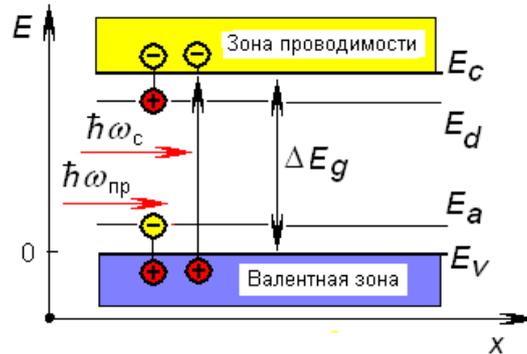


Рисунок 21 – Энергетическая диаграмма при примесном поглощении излучения

расположенном между координатами x и $x + dx$, будет равно:

$$\begin{aligned} -dI &= I(x + dx) - I(x) = I_0 \exp(-\alpha(x + dx)) - I_0 \exp(-\alpha x) = \\ &= I_0 \exp(-\alpha x) \exp(-\alpha \cdot dx) - I_0 \exp(-\alpha x) = I_0 \exp(-\alpha x) \cdot [\exp(-\alpha \cdot dx) - 1] = \\ &= -I(x) \cdot [1 - \alpha \cdot dx - 1] = -I(x) \cdot \alpha \cdot dx, \end{aligned}$$

где произведено разложение экспоненты в ряд Тейлора по степеням dx вблизи точки x . Как следует из полученного выражения, число поглощаемых фотонов, рассчитанное на единицу времени в единичном объеме, равно $I(x, \lambda) \cdot \alpha$. Поэтому скорость оптической генерации g можно представить в виде:

$$g = \eta(\lambda) \cdot \alpha(\lambda) \cdot I(x, \lambda). \quad (4.6)$$

Здесь $\eta(\lambda)$ – квантовый выход внутреннего фотоэффекта, который равен числу пар (для собственного поглощения) свободных носителей, рождаемых в среднем одним поглощенным фотоном. Квантовый выход может быть больше единицы, если при поглощении одного фотона высокой энергии возникает две или более электронно-дырочные пары. Обычно квантовый выход меньше единицы, поскольку часть фотонов поглощается либо решеткой (фононами), либо свободными носителями в зонах без образования неравновесных электронов и дырок.

Выражение (4.6) можно переписать, заменив в нем коэффициент поглощения света обратной длиной пробега фотонов $\ell_\phi = 1/\alpha$:

$$g = \eta(\lambda) \cdot \frac{I(x, \lambda)}{\ell_\phi}. \quad (4.7)$$

Данное выражение показывает, что в результате поглощения света все неравновесные фотоны сосредоточены на длине поглощения света. В действительности за счет диффузии неравновесные носители, созданные в слое поглощения, за время жизни разбегаются от места генерации на расстояние, равное диффузионной длине $L_{n,p}^{(d)}$. Это расстояние обычно составляет $(10^2 - 10^3)$ мкм, что на несколько порядков больше длины поглощения света. Тогда выражение для скорости генерации приобретет вид:

$$g_n = \eta(\lambda) \frac{I(x, \lambda)}{L_n^{(d)}}, \quad g_p = \eta(\lambda) \frac{I(x, \lambda)}{L_p^{(d)}}.$$

Для того чтобы определиться во всем многообразии возможных фотоэффектов в полупроводнике при его освещении излучением из собственной полосы поглощения, обратимся к выражению для проводимости полупроводника, которая в общем случае представляет собой сумму проводимостей по зоне проводимости, создаваемой свободными электронами – σ_n , и проводимости по валентной зоне, создаваемой дырками – σ_p :

$$\sigma = \sigma_n + \sigma_p = q\mu_n n + q\mu_p p. \quad (4.8)$$

Экспериментально установлено, что при падении на полупроводник излучения с интенсивностью I проводимость полупроводника изменяется на величину $\delta\sigma(I)$. Так как рассматриваются только слабые световые потоки, то, согласно правилам дифференциального исчисления, без каких-либо дополнительных предположений, можно вычислить полный дифференциал выражения (4.8):

$$\delta\sigma(I) = qn \cdot \delta\mu_n + qp \cdot \delta\mu_p + q\mu_n \cdot \delta n + q\mu_p \cdot \delta p, \quad (4.9)$$

где $\delta\mu_n$, $\delta\mu_p$, δn , δp представляют собой полные дифференциалы от соответствующих величин. Таким образом, согласно выражению (4.9), освещение полупроводника может изменить его проводимость двумя путями: изменив концентрации свободных носителей в зонах, а также изменив их подвижности. Однако относительное влияние этих двух механизмов фотопроводимости в реальности оказывается различным. Действительно, возникающая при поглощении фотона пара «электрон – дырка» получает некоторый квазиимпульс и энергию, равную величине $(\hbar\omega - \Delta E_g)$, которая идет на кинетическую энергию электрона. Если кинетическая энергия велика, то электрон быстро перемещается в междоузлии кристалла и испытывает множественные столкновения, происходящие чаще, чем у свободного электрона, находящегося внизу зоны проводимости. Поэтому его избыточная энергия быстро растрачивается. В результате через некоторое время, именуемое временем релаксации энергии – $\tau_E \sim 10^{-12} \text{ с}$ и которое много меньше τ_n , τ_p , средняя энергия фотоэлектронов, имевших в момент своего рождения большую кинетическую энергию, принимает значение, соответствующее температуре решетки. В таком случае говорят, что электрон термолизировался – его кинетическая энергия стала соответствовать температуре кристалла.

Таким образом, за время своей жизни фотоэлектроны в зоне успевают быстро термолизироваться (прийти в состояние термодинамического равновесия с решеткой), так что все их параметры в течение практически всего времени жизни будут такими же, как и у равновесных носителей. В этом случае подвижности фотоэлектронов равны подвижности темновых (равновесных) электронов. Значит, даже если освещать полупроводник светом с $\hbar\omega \gg \Delta E_g$, то избыток энергии вызывает изменение и концентрации носителей, и изменение их подвижностей, но через очень короткое время $\tau_E \ll \tau_n$, τ_p подвижности неравновесных носителей становятся такими же, что и у равновесных электронов и дырок. Следовательно, с высокой степенью точности можно считать, что в результате поглощения света в полупроводниках изменяются только концентрации свободных носителей заряда: электронов и (или) дырок. Тогда выражение (5.9) для фотопроводимости освещаемого полупроводника с высокой степенью точности можно записать так:

$$\delta\sigma = q(\mu_p \cdot \delta p + \mu_n \cdot \delta n).$$

Будем считать, что происходит однородная генерация пар ($g_n = g_p = g = \text{const}$) и что в полупроводнике нет тока проводимости. Тогда, умножая кинетические уравнения для концентраций носителей в зонах на $q\mu_n$, $q\mu_p$ и складывая их, получим:

$$\frac{\partial \delta \sigma}{\partial t} = q(\mu_p + \mu_n)g - \frac{\delta \sigma}{\tau_{\phi n}}; \quad (4.10)$$

где обозначено:

$$\tau_{\phi n} = \frac{\mu_n \cdot \delta n + \mu_p \cdot \delta p}{\frac{\mu_n \cdot \delta n}{\tau_n} + \frac{\mu_p \cdot \delta p}{\tau_p}}. \quad (4.11)$$

Из уравнения (4.10) следует, что характеристическое время $\tau_{\phi n}$ – время релаксации фотопроводимости определяет время установления стационарного значения фотопроводимости $\delta \sigma$ при изменении условий освещения. В стационарном состоянии фотопроводимость равна

$$\delta \sigma(t \gg \tau_{\phi n}) = \delta \sigma_s = q(\mu_n + \mu_p) \cdot g \cdot \tau_{\phi n}. \quad (4.12)$$

Здесь под $\tau_{\phi n}$ понимают его значение в стационарном состоянии, т.е. при установившихся значениях δp и δn . Отсюда видно, что чем больше $\tau_{\phi n}$, тем больше $\delta \sigma_s$, т.е. тем выше чувствительность полупроводника к оптическому излучению, т.к. выше уровень стационарного значения фотопроводимости (4.12). Однако при этом будет больше и время установления (или затухания) фотопроводимости, т.е. будет больше инерционность фоторезистора, и им нельзя будет регистрировать быстрые процессы. С этим противоречием приходится считаться при решении технических задач.

Вследствие изменения проводимости полупроводника в цепи возникает переменный ток, а на сопротивлении r – переменное напряжение. Оно усиливается широкополосным усилителем Y , а затем регистрируется осциллографом O . Зная параметры входной цепи,

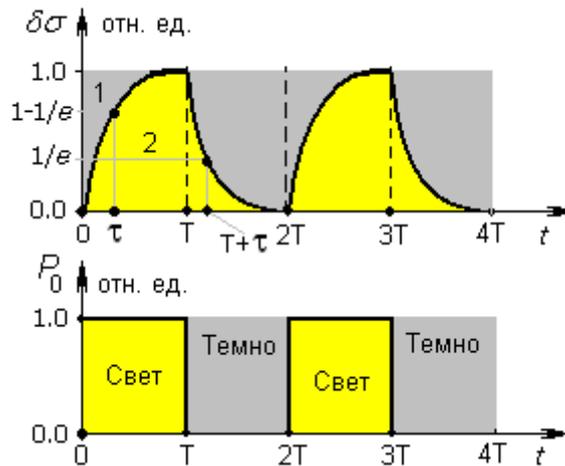


Рисунок. 22 – Кинетика фотопроводимости при импульсном освещении

можно рассчитать величину фотопроводимости и ее изменение во времени. Если $\tau_{\phi n}$ постоянно во времени, то при импульсном освещении фоторезистора со скважностью равной единице $\delta \sigma$ изменяется во времени так, как это показано на рис. 22. Если T - длительность светового импульса, за которым следует затемнение фоторезистора такой же продолжительности, то общий период модуляции светового потока составит $2T$. Будем считать, что импульсы света достаточно длинные: $T \gg \tau_{\phi n}$ и потому за время его действия фотопроводимость достигает своего максимального значения $\delta \sigma_s$. На отрезке времени $(T; 2T)$ генерация неравновесных носителей заряда отсутствует, а доминирует рекомбинация ранее созданных фотоэлектронов и фотодырок. Потому фотопроводимость полупроводника уменьшается от $\delta \sigma_s$ по экспоненциальному закону:

$$\delta\sigma(t) = \delta\sigma_s \cdot \exp\left(-\frac{t}{\tau_{\phi n}}\right).$$

Кривые релаксации фотопроводимости могут быть использованы для определения параметров полупроводника. На начальном участке нарастания $\delta\sigma$ в уравнении (5.10) можно пренебречь вторым слагаемым в правой части ($\delta\sigma$ – мало). Тогда получим:

$$\delta\sigma(t) \approx q(\mu_n + \mu_p)gt.$$

Линейное увеличение фотопроводимости со временем освещения характеризуется наклоном, тангенс угла которого равен

$$\operatorname{tg}(\alpha) = q(\mu_n + \mu_p)\eta \cdot \alpha \cdot I,$$

где η – квантовый выход, α – коэффициент поглощения. Этот участок роста фотопроводимости не зависит от процессов рекомбинации, а определяется только процессами генерации. Поэтому исследование начальной стадии нарастания фотопроводимости используют для экспериментального определения квантового выхода η .

При выключении света спад фотопроводимости используют для определения времени релаксации фотопроводимости как время, за которое фотопроводимость уменьшается от своего максимального значения $\delta\sigma_s$ в e -раз.

Коэффициент усиления фотопроводимости. Будем считать, что рассматриваемый полупроводниковый образец с омическими контактами на концах имеет длину L . В цепи с нагрузочным сопротивлением R (рис.23) возникающий фототок при действии электрического поля E будет равен:

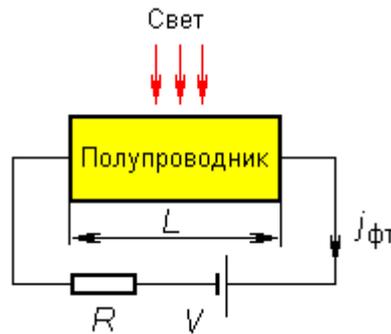


Рисунок.23 – Электрическая схема включения фоторезистора

$$j_{\phi} = \delta\sigma_s \cdot E = q(\mu_n + \mu_p) \cdot g \cdot \tau_{\phi n} \cdot E.$$

Если, то можно записать (предполагая также $\delta n = \delta p$): $j_{\phi} \approx q(\mu_n \tau_n + \mu_p \tau_p) \cdot g \cdot E$.

Введем в рассмотрение время пролета образца электроном и дыркой соответственно:

$$t_n = \frac{L}{v_n} = \frac{L}{\mu_n \cdot E}; \quad t_p = \frac{L}{v_p} = \frac{L}{\mu_p \cdot E},$$

где v_n, v_p – скорости электронов и дырок. Тогда фототок в цепи будет

$$j_{\phi} \approx q \cdot \left(\frac{\tau_n}{t_n} + \frac{\tau_p}{t_p} \right) \cdot L \cdot g.$$

Обозначим: $K = \frac{\tau_n}{t_n} + \frac{\tau_p}{t_p}$. Эта величина показывает, сколько раз за время своей жизни

электронно-дырочная пара может пролететь через весь образец, т.е. участвовать в фотопроводимости. С учетом этого запишем

$$j_{\phi} \approx q \cdot K \cdot L \cdot g.$$

Если $K < 1$, то электрон и дырка успевают прорекомбинировать, не пролетев полностью образец. Увеличив прикладываемое электрическое поле, можно уменьшить t_n и t_p , а значит, увеличить $j\phi$. Параметр K называют коэффициентом усиления фотопроводимости. При решении практических задач необходимо стремиться к тому, чтобы K было максимальным. Зависимость величины фототока от приложенного поля обычно имеет вид, показанный на рис. 26. Величину электрического поля E_0 , при котором достигается предельное значение коэффициента усиления, можно найти из условия: $t_n \approx \tau_n$, ($t_p \approx \tau_p$)

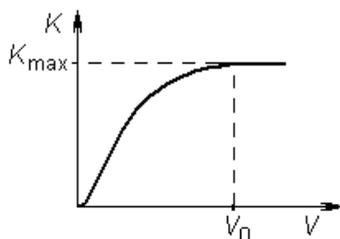


Рисунок 24 – Зависимость коэффициента усиления фотопроводимости от напряжения смещения

$$E_0 \approx \frac{L}{\mu_n \cdot \tau_n} \quad \text{или} \quad V_0 \approx \frac{L^2}{\mu_n \cdot \tau_n}.$$

Увеличивать напряжение на образце V выше V_0 нецелесообразно, т.к. при $V = V_0$ через контакт вытягиваются из образца оба неравновесных носителя, так что K_{\max} достигает максимального значения.

4.3. Полевые свойства фотопроводимости

Для измерения фотопроводимости и ее кинетики разработано много различных методик. Рассмотрим одну из них, блок - схема которой показана на рис. 25. Здесь фоторезистор, показанный сопротивлением R и включенный последовательно с источником напряжения V и нагрузочным сопротивлением r , освещается прерывистым светом. Модуляция интенсивности света может быть получена, например, с помощью вращающегося непрозрачного диска с прорезями.

В отсутствии освещения через фоторезистор течет ток

$$j_0 = \frac{V}{R_0 + r},$$

который создает на входе усилителя напряжение

$$U_{\text{вх}}^0 = r \cdot j = V \frac{r}{R_0 + r}.$$

При освещении фоторезистора его сопротивление изменяется и становится равным

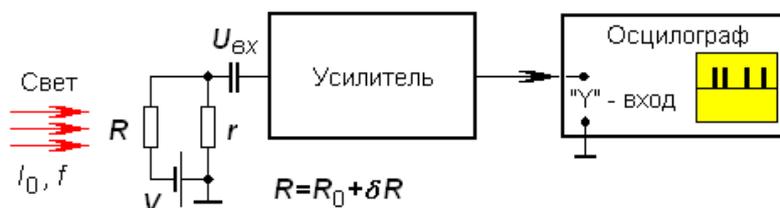


Рисунок 25 – Блок – схема установки для регистрации фотопроводимости.

$R = R_0 + \delta R$. Изменяется и ток во входной цепи и потому на входе усилителя появится дополнительное напряжение, связанное с освещением фоторезистора. Следовательно, полное напряжение на входе усилителя будет

$$U_{ex} = r \cdot j = V \frac{r}{R+r} = V \frac{r}{R_0 + \delta R + r} = V \frac{r}{(R_0 + r) + \delta R} = -V \frac{r}{R_0 + r} \frac{1}{1 + (\delta R / (R_0 + r))} \approx U_{ex}^0 \left[1 - \frac{\delta R}{R_0 + r} \right].$$

Таким образом, при освещении полупроводника на входе усилителя появляется компонента во входном напряжении, пропорциональная интенсивности освещения и равная:

$$\delta U_{ex} \approx U_{ex}^0 \frac{\delta R}{R_0 + r}.$$

Зная параметры входной цепи, можно рассчитать величину фотопроводимости и ее изменение во времени. Анализируя проведенный расчет, можно сделать ряд полезных выводов. Во-первых, данное выражение показывает, что величина сигнала фотопроводимости тем больше, чем больше значение U_{ex}^0 , т.е. чем больше постоянное напряжение V , приложенное к фоторезистору через сопротивление нагрузки. Во-вторых, можно показать, что максимальный сигнал на входе усилителя достигается тогда, когда сопротивление сопротивления нагрузки и темновое сопротивление фоторезистора равны: $R_0 = r$. Действительно, сигнал на входе усилителя можно переписать следующим образом:

$$\delta U_{ex} \approx U_{ex}^0 \frac{\delta R}{R_0 + r} = V \frac{r}{(R_0 + r)^2} \delta R = V \frac{(r/R_0)}{(1 + r/R_0)^2} \frac{\delta R}{R_0}.$$

Обозначим $x = r/R_0$. Максимум $\delta U_{ex}(x)$ найдем дифференцированием этой функции и приравняв к нулю производной, что дает одно физически реализуемое решение: $x = 1$.

4.4. Частотные свойства фотопроводимости

Если интенсивность падающего оптического излучения модулирована по гармоническому закону (так, излучение ламп накаливания модулировано по амплитуде частотой $2 \times 50 = 100$ Гц), то функции генерации электронов и дырок будут изменяться во времени по закону модуляции излучения

$$g_n = g_0 \exp(i\omega t), \quad g_p = g_0 \exp(i\omega t). \quad (4.13)$$

Значит, изменения концентраций носителей заряда также подчиняются гармоническому закону изменения во времени на той же частоте, и поэтому решения кинетического уравнения для концентраций носителей заряда будем отыскивать в виде

$$\Delta n = \Delta n_0 \cdot \exp(i\omega t), \quad \Delta p = \Delta p_0 \cdot \exp(i\omega t). \quad (4.14)$$

Подставив выражения (4.13) и (4.14) в кинетические уравнения, найдем, что максимальные значения концентраций неравновесных носителей зависят от частоты модуляции света:

$$\Delta n_0 = \frac{g_0 \cdot \tau_n}{1 + i\omega \cdot \tau_n}, \quad \Delta p_0 = \frac{g_0 \cdot \tau_p}{1 + i\omega \cdot \tau_p}. \quad (4.15)$$

Однако физический смысл изменения концентраций имеют только реальные части выражений (4.15). Отыскивая их, получим:

$$\Delta n_0 = \frac{g_0 \cdot \tau_n}{1 + \omega^2 \cdot \tau_n^2}, \quad \Delta p_0 = \frac{g_0 \cdot \tau_p}{1 + \omega^2 \cdot \tau_p^2}. \quad (4.16)$$

Выражения (4.16) показывают, что амплитуды изменений концентраций носителей в зонах зависят от соотношения частоты модуляции светового потока и времени жизни носителей. Если частота модуляции ω низкая ($\omega \cdot \tau_n \ll 1$, $\omega \cdot \tau_p \ll 1$), то изменения концентраций носителей заряда максимальны и равны (обозначим этот случай верхним индексом НЧ)

$$\Delta n_0^{\text{НЧ}} = g_0 \cdot \tau_n, \quad \Delta p_0^{\text{НЧ}} = g_0 \cdot \tau_p. \quad (4.17)$$

На высокой частоте ($\omega \cdot \tau_n \gg 1$, $\omega \cdot \tau_p \gg 1$) изменения концентраций уменьшаются с ростом частоты (этот случай обозначим индексом ВЧ):

$$\Delta n_0^{\text{ВЧ}} = \frac{g_0 \cdot \tau_n}{\omega^2 \cdot \tau_n^2}, \quad \Delta p_0^{\text{ВЧ}} = \frac{g_0 \cdot \tau_p}{\omega^2 \cdot \tau_p^2}. \quad (4.18)$$

В большинстве практически важных случаев $\tau_n = \tau_p$, поэтому обозначим $\tau_{n,p} = \tau_n = \tau_p$. Частота ω_0 , на которой начинается спад амплитуды избыточных концентраций носителей заряда, определяется выражением (16)

$$\omega_0 \cdot \tau_{n,p} = 1 \quad (4.19)$$

и называется линейной **частотой среза**:

$$f_0 = \frac{1}{\tau_{n,p}}. \quad (4.20)$$

На этой частоте амплитуда неравновесной концентрации по сравнению с максимальным значением уменьшается в два раза.

Из выражения (4.20) следует важный вывод: измеряя экспериментально зависимость амплитуд концентраций неравновесных носителей заряда от частоты модуляции светового потока, по частоте среза можно определить время жизни неравновесных носителей заряда. Тогда на низкой частоте модуляции, зная значения времен жизни неравновесных электронов и дырок $\tau_{n,p}$, по выражениям (4.17) можно найти скорость их генерации, а из неё вычислить квантовую эффективность полупроводника, если известно значение интенсивности I_0 .

Способом, аналогичным рассмотренному выше для вычисления частотной зависимости концентрации носителей заряда, можно найти закон изменения величины фотопроводимости $\Delta\sigma$ от частоты модуляции светового потока и получить следующее выражение:

$$\Delta\sigma = \frac{q(\mu_n + \mu_p) \cdot g \cdot \tau_{\phi n}}{1 + \omega^2 \cdot \tau_{\phi n}^2}, \quad (4.21)$$

где g - скорость генерации электронно-дырочных пар в полупроводнике за счет поглощения падающего излучения, а $\tau_{\phi n}$ - время релаксации фотопроводимости ($\delta n \approx \delta p$)

$$\tau_{\phi n} = \frac{\mu_n \cdot \delta n + \mu_p \cdot \delta p}{\frac{\mu_n \cdot \delta n}{\tau_n} + \frac{\mu_p \cdot \delta p}{\tau_p}} \approx \frac{\mu_n + \mu_p}{\frac{\mu_n}{\tau_n} + \frac{\mu_p}{\tau_p}} = \frac{(\mu_n + \mu_p) \cdot \tau_n \tau_p}{\mu_n \tau_p + \mu_p \tau_n}. \quad (4.22)$$

Из выражения (4.21) следует, что фотопроводимость полупроводника на низких частотах модуляции остается постоянной величиной, а на высоких уменьшается пропорционально квадрату частоты модуляции оптического излучения. При этом **частота среза фотопроводимости** определяется как величина, обратная времени релаксации фотопроводимости:

$$\tau_{\phi n} = \frac{1}{f_0}$$

Таким образом, исследуя частотные свойства фотопроводимости, можно по приведенным выражениям вычислить кинетические параметры полупроводника.

В заключение этого раздела добавим, что применение квадратичного детектора для детектирования напряжения фотопроводимости на частоте модуляции света ω изменяет зависимость (21) до вида

$$U_{\phi n} = V \frac{(r/R_0)}{(1+r/R_0)^2} \frac{1}{\sigma_0} \frac{q(\mu_n + \mu_p) \cdot g \cdot \tau_{\phi n}}{\sqrt{1 + \omega^2 \cdot \tau_{\phi n}^2}}. \quad (4.23)$$

Частотные характеристики несут также важную информацию о механизмах рекомбинации носителей заряда в фоторезисторах. Рассмотрим способ извлечения этой информации.

Экспериментальные частотные зависимости часто представляют собой наложение зависимостей с различными частотами среза и различными наклонами высокочастотной ветви. Каждую из частотных зависимостей, составляющих суммарную зависимость, можно описать зависимостью следующего вида

$$U_{\phi n} = V \frac{(r/R_0)}{(1+r/R_0)^2} \frac{1}{\sigma_0} \frac{q(\mu_n + \mu_p) \cdot g \cdot \tau_{\phi n}}{\sqrt{1 + \omega^n \cdot \tau_{\phi n}^n}} \quad (4.24)$$

Здесь n - коэффициент, определяющий угол наклона высокочастотной ветви, может принимать значение от 0.5 до 3.5 в зависимости от механизма рекомбинации неравновесных носителей заряда. По этой причине его необходимо знать. Покажем, как его найти из экспериментальной кривой.

Перепишем для удобства выражение (24) в краткой форме

$$U_{\phi n}(\omega) = \frac{U_{\phi n}(0)}{\sqrt{1 + \omega^n \cdot \tau_{\phi n}^n}}, \quad (4.25)$$

Здесь числитель дроби есть величина напряжения фотопроводимости на низкой частоте, условно обозначенной нулем. Выразим из (4.25) неизвестный параметр n . Тогда получим

$$n = \frac{\lg \left[\left(U_{\phi n}(0) / U_{\phi n}(\omega) \right)^2 - 1 \right]}{\lg(\omega \tau_{\phi n})}. \quad (4.26)$$

Вычисляя значение параметра n для каждого участка частотной характеристики, получим информацию о механизмах рекомбинации неравновесных носителей заряда в данном фоторезисторе и последовательности их взаимного преобразования.

4.5. Фотовольтаические эффекты в полупроводниках

Рассмотренная выше фотопроводимость полупроводников характеризуется положительными и отрицательными сторонами. К положительным сторонам фотопроводимости можно отнести простоту изготовления полупроводникового образца, предназначенного для регистрации фотопроводимости – прямоугольный брусок полупроводника с омическими контактами на противоположных торцах, а также простота интерпретации физических процессов, формирующих сигнал фотопроводимости. К отрицательным сторонам фотопроводимости относятся следующие: необходимость в схеме фотоприемника иметь периодически обновляемый источник постоянного напряжения, что ограничивает область применения данного фотоэффекта (исключаются системы космического базирования и т.п.), протекание постоянного тока через полупроводник приводит к дополнительным шумам (дробовым, контактными), что снижает обнаружительную способность фотоприемника. В дополнение к этому заметим, что схемотехнически регистрация фотопроводимости, как правило, осуществляется путем ее

перевода в напряжение – падение напряжения на сопротивлении нагрузки радиотехнического тракта фотоприемника.

Эти обстоятельства побуждают искать фотоэффекты, не нуждающиеся в источниках постоянного напряжения, а также сами вырабатывающие напряжение, пропорциональное интенсивности принимаемого оптического излучения.

При освещении полупроводника изменяется не только его электропроводность, но также возникают электродвижущие силы. Общая причина возникновения фотоэлектродвижущей силы (или фотоэдс), по крайней мере в наиболее важных случаях, одна и та же и заключается в пространственном разделении неравновесных носителей заряда. Это приводит к тому, что разделенные фотоэлектроны и фотодырки образуют дополнительное электрическое поле, которое приводит к появлению на торцах полупроводника фотоэдс. Процесс пространственного разделения неравновесных носителей заряда может происходить по следующим основным причинам:

- за счет различия коэффициентов диффузии электронов и дырок в полупроводнике;
- за счет действия в полупроводнике внутреннего электрического поля, которое может быть создано искусственно (как это сделано в $p-n$ переходе) или случайно (за счет неконтролируемого появления электрически заряженных дефектов или неоднородностей легирования полупроводника).

Несмотря на достаточно общую картину происхождения фотоэдс, удобно говорить о разных ее типах в зависимости от особенностей полупроводниковой структуры и условий проведения регистрации излучения.

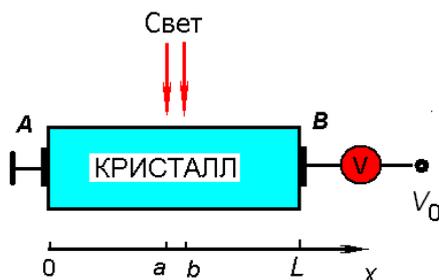


Рисунок 26 – Возникновение фотоэдс при освещении полупроводника

Однако сначала остановимся на общих условиях, необходимых для возникновения фотоэдс. Чтобы сделать рассуждения более наглядными, рассмотрим полупроводник в виде прямоугольного параллелепипеда, часть которого « ab » освещается излучением из полосы собственного поглощения. Образец ограничен слева и справа одинаковыми металлическими электродами A и B для измерения фотоэдс (рис. 26). Полупроводник может быть неоднородным, т.е. примеси в нем могут быть распределены неравномерно по длине образца, а отдельные его части могут иметь даже различный тип проводимости (электронный или дырочный). Будем считать, что подвижности электронов и дырок при освещении не меняются. Концентрации δn и δp будем предполагать зависящими только от координаты x вдоль образца (толщина образца мала по сравнению с длинами диффузии носителей).

При освещении в образце потекут токи с плотностями, содержащими дрейфовую и диффузионную компоненты:

$$j_p = \sigma_p E - qD_p \frac{dp}{dx}; \quad j_n = \sigma_n E + qD_n \frac{dn}{dx}.$$

Здесь $D_n = \sqrt{\frac{kT}{q} \mu_n}$, $D_p = \sqrt{\frac{kT}{q} \mu_p}$ - коэффициенты диффузии электронов и дырок. Полная плотность тока будет равна

$$j = j_n + j_p = \sigma \cdot \left(E + q \frac{D_n \frac{dn}{dx} - D_p \frac{dp}{dx}}{\sigma} \right),$$

где $\sigma = q(\mu_n n + \mu_p p)$ – полная электропроводность полупроводникового материала. С другой стороны, согласно закону Ома, можно записать:

$$j = \sigma \cdot (E + E^*),$$

где E^* – напряженность поля сторонних сил. Сравнивая оба выражения для j , найдем

Тогда полное фотонапряжение, измеренное на торцах образца A и B , будет равно:

$$V^* = \int_0^L E^* dx = \int_0^L \frac{D_n \frac{dn}{dx} - D_p \frac{dp}{dx}}{\mu_n n + \mu_p p} dx, \quad (4.27)$$

где интегрирование проводится вдоль всего образца по направлению к контактам – направлению регистрации фотоэдс.

Из выражения (4.27) следует, что возникающая на торцах полупроводника фотоэдс тем больше, чем значительнее различие в величинах диффузионных потоков электронов и дырок, а также чем меньше темновая проводимость полупроводника. При этом не имеет значение: вызван диффузионный поток свободных носителей заряда за счет неравномерного распределения равновесных электронов и дырок или неравновесных. Далее также ясно, создав освещением полупроводника одинаковые концентрации основных и неосновных носителей заряда роль неосновных носителей в возникновении фотоэдс должна быть значительнее. Поясним этот вывод простым примером. Пусть в полупроводнике имеется 1000 основных носителей, например, электронов и 100 неосновных – дырок. Освещением полупроводника мы создаем одинаковое количество тех и других: например, по 10 штук. Однако, такое освещение изменило число основных носителей лишь на 1%, а число неосновных – на 10%. Поэтому и эффект от изменения числа неосновных носителей следует ожидать больше, чем от основных.

Чтобы выяснить общие свойства фотоиндуцированного напряжения, исследуем более подробно полученное выражение.

Однородный полупроводник. Предположим, что $n = n_0 + \delta n$, $p = p_0 + \delta p$. Поскольку полупроводник считается однородным, то $\frac{dn_0}{dx} = \frac{dp_0}{dx} = 0$. В случае одинакового градиента неравновесных электронов и дырок

$$D_n \frac{d\delta n}{dx} - D_p \frac{d\delta p}{dx} = \frac{kT}{q} \left[\mu_n \frac{d\delta n}{dx} - \mu_p \frac{d\delta p}{dx} \right] = \frac{kT}{q} [\mu_n - \mu_p] \frac{d\delta p}{dx},$$

поскольку градиенты избыточных концентраций носителей заряда обычно одинаковы. Тогда по выражению (4.27)

$$V^* = \frac{kT}{q} (\mu_n - \mu_p) \int_{\delta p(0)}^{\delta p(L)} \frac{d(\delta p)}{\mu_n n_0 + \mu_p p_0 + (\mu_n + \mu_p) \delta p}.$$

Здесь подынтегральное выражение представляет собой полный дифференциал функции

$$\ln[\mu_n n_0 + \mu_p p_0 + (\mu_n + \mu_p) \delta p] \approx \ln[\mu_n (n_0 + \delta n) + \mu_p (p_0 + \delta p)],$$

и потому, вычислив ее значение в точках A и B , получим разность двух значений логарифмической функции, отличающихся друг от друга только небольшим значением логарифма от $\delta p(x)$. Следовательно, в этом случае фотоэдс $V^* \cong 0$.

Таким образом, в однородном полупроводнике с одинаковыми градиентами неравновесных электронов и дырок фотоэдс на освещаемом полупроводнике мала.

Неоднородный полупроводник. Пусть свет создает только основные носители. Будем для определенности рассматривать полупроводник n -типа. Для исключения эдс на торцах, связанной с неоднородным легированием будем считать, что $\frac{dn_0}{dx} = \frac{dp_0}{dx} = 0$.

Для неравновесных концентраций носителей заряда справедливы следующие неравенства:

$$D_n \frac{dn}{dx} \gg D_p \frac{dp}{dx}, \quad \mu_n n_0 \gg \mu_p p_0.$$

Поэтому выражение (4.27) дает:

$$V^* = \frac{kT}{q} \int_{\delta n(0)}^{\delta n(L)} \frac{dn}{n} = \frac{kT}{q} \ln[(n_0 + \delta n(L))] - \frac{kT}{q} \ln[(n_0 + \delta n(0))] = \frac{kT}{q} \left[\frac{\delta n(L)}{n_0} - \frac{\delta n(0)}{n_0} \right] \approx 0,$$

где при разложении логарифмической функции использовано условие малости концентрации неравновесных электронов по сравнению с концентрацией равновесных. По этой причине рассматриваемый интеграл дает малое значение.

Следовательно, освещение полупроводника, в результате которого создаются только основные неравновесные носители заряда, не создает на концах полупроводника заметной по величине фотоэдс.

Пусть свет создает только неосновные носители заряда - дырки. Для этого случая справедливы следующие неравенства:

$$\frac{dn_0}{dx} = \frac{dp_0}{dx} = 0, \quad D_n \frac{dn}{dx} \ll D_p \frac{dp}{dx}, \quad \mu_n n_0 \gg \mu_p p_0.$$

Поэтому выражение (4.27) дает:

$$V^* = -\frac{kT}{q} \int_{\delta p(0)}^{\delta p(L)} \frac{dp}{p} \approx 0,$$

т.е. опять приходим к полному дифференциалу, но теперь уже функции $\ln(p)$.

Следовательно, освещение полупроводника, в результате которого создаются только неосновные неравновесные носители заряда, создает на концах полупроводника большую по величине фотоэдс.

Если продолжить этот анализ, то можно показать (приводится без полного математического доказательства), что для возникновения значительной по величине фотоэдс необходимо, чтобы полупроводник должен быть неоднородным по уровню легирования вдоль направления регистрации фотонапряжения, а также свет должен создавать заметное количество неосновных носителей заряда (знак их заряда противоположен знаку основных носителей заряда).

4.6. Квазиуровни Ферми в полупроводнике

Существует достаточно простой и наглядный способ оценки величины фотоэдс, возникающей в полупроводниковом приборе при его освещении. Для его рассмотрения обратимся к оценке вероятности заполнения уровней энергии. В термодинамически равновесном состоянии (например, при отсутствии освещения полупроводника) вычислить концентрации электронов и дырок в зонах можно, если воспользоваться распределением Ферми-Дирака. При этом необходимо знать положение уровня Ферми F . Если концентрация электронов в полупроводнике (например, за счет его легирования) будет увеличиваться, то уровень Ферми, оставаясь в запрещенной зоне, будет смещаться к дну зоны проводимости (см. рис. 28, а). В случае увеличения концентрации дырок он смещается к потолку валентной зоны. Поэтому уровни энергий, лежащие вблизи уровня Ферми, при изменении его энергетического положения меняют вероятность заполнения электронами. Такой подход к описанию изменений заполнения энергетических состояний в равновесии прост и удобен.

Однако часто необходимо знать концентрации носителей в валентной зоне и зоне проводимости в неравновесных условиях, например, при освещении полупроводника. Попробуем применить для этой цели равновесную функцию распределения Ферми-Дирака. Использовать ее будет возможно, если удастся определить положение уровня Ферми F в неравновесных условиях. Займемся этим.

При отсутствии освещения уровень Ферми в полупроводнике определяется концентрациями равновесных носителей заряда в зонах разрешенных энергий, т.е. в конечном счете – концентрацией легирующей донорной или акцепторной примеси (см. рис.27, а). Освещение полупроводника увеличивает концентрацию электронов, и потому в неравновесных условиях уровень Ферми должен сместиться вверх относительно своего равновесного положения. Одновременно с освещением увеличивается и концентрация дырок

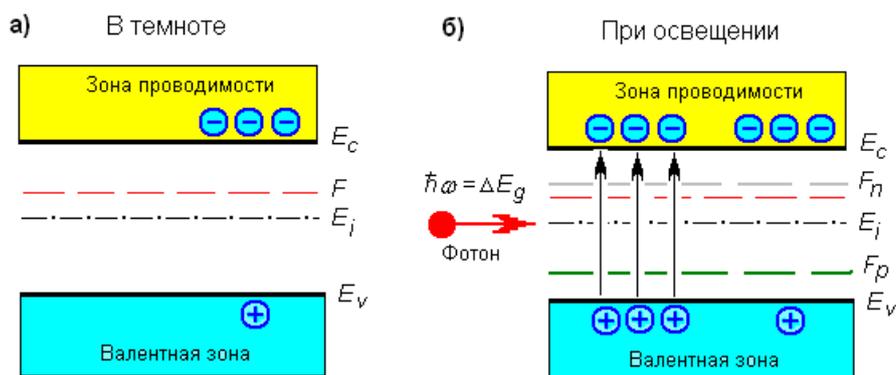


Рисунок 27 – Введение квазиуровня Ферми при освещении полупроводника

и тогда, опять-таки согласно рассуждениями первого абзаца этого параграфа, уровень Ферми должен переместиться вниз (см. рис.27, б). Но он не может при освещении одновременно двигать вверх и вниз. Значит, введение в рассмотрение единичного уровня Ферми в этих условия невозможно. Поэтому распределение свободных носителей по энергии в неравновесных условиях нельзя описать равновесной функцией Ферми с одним уровнем.

Однако можно обобщить равновесную статистику на неравновесную ситуацию, если вместо одного уровня Ферми формально ввести два квазиуровня Ферми: один - для описания распределения электронов по энергии в неравновесных условиях, а другой – для описания распределения дырок. Тогда вероятность заполнения электроном состояния с энергией E в зоне проводимости можно представить в виде функции Ферми-Дирака:

$$f_n = \left(1 + \exp \frac{E - F_n}{kT} \right)^{-1}.$$

Здесь F_n есть квазиуровень Ферми для электронов. Аналогично, для вероятности обнаружения дырки на уровне с энергией E в валентной зоне положим

$$f_p = \left(1 + \exp \frac{F_p - E}{kT} \right)^{-1},$$

где по определению F_p есть квазиуровень Ферми для дырок. Введение квазиуровней для неравновесных состояний полупроводника допустимо, т.к. созданные светом носители заряда, как уже указывалось выше, через очень короткое время ($\sim 10^{-12}$ с) термализуются.

Таким образом, появление в зонах неравновесных электронов и дырок можно описать как расщепление первоначального уровня Ферми F на два квазиуровня F_n и F_p , которые смещаются по направлению к своей зоне: F_n – к зоне проводимости, а F_p – к валентной зоне.

Можно показать, что разность квазиуровней Ферми на концах полупроводника непосредственно связана с разностью потенциалов на них. Следовательно, разность квазиуровней Ферми для электронов и для дырок есть та фотоэдс, которую можно измерить вольтметром на концах полупроводникового образца при его освещении.

4.9. Барьерная фотоэдс в полупроводниковом $p-n$ переходах

4.9.1. Электронно-дырочные переходы

Из выражения для величины фотовольтаического эффекта (4.12) следует, что она тем больше, чем меньше проводимость полупроводника, т.е. чем более высокоомный полупроводник. Наибольшим сопротивлением обладает полупроводник с концентрациями носителей заряда, равными собственной концентрации $n = p = n_i$. Однако изготовить массивный образец с собственной концентрацией технологически сложно. Поэтому целесообразно использовать для регистрации оптического излучения образец, в котором имеется отдельная область собственной проводимости. Такую область можно получить, изготовив полупроводниковый $p-n$ переход. Как известно, он образуется при соединении двух полупроводников, обладающих разными типами проводимости: электронной (n -тип) и дырочной (p -тип). Чтобы избежать неконтролируемого влияния поверхностей на свойства $p-n$ -перехода, его изготавливают из одного монокристалла легированием торцов образца донорной и акцепторной примесями, а затем вжиганием примесей в объем высокотемпературной обработкой.

Требуемое распределение доноров и акцепторов осуществляется с помощью специальных технологических приемов. Распределение концентраций электронов и дырок в $p-n$ -переходе показано на рис. 28 сплошными линиями. Причина формирования именно такого хода зависимостей концентраций равновесных носителей кроется в большом различии концентраций подвижных носителей заряда в n - и p -областях. В результате этого электроны из n -области устремляются туда, где их мало, т.е. в p -область, а дырки из p -области будут диффундировать в n -область. Эти два встречных потока будут продолжаться до тех пор, пока оставляемые без подвижного заряда положительные ионы в n -области и отрицательные в p -области не создадут электрическое поле $-\bar{E}_K$ (рис. 28), которое затормозит диффузию равновесных носителей. Это электрическое поле направлено от n -к p -области.

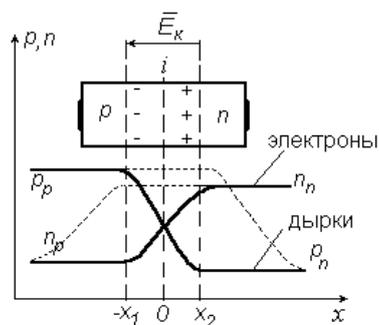


Рисунок 28 – Распределение носителей заряда в $p-n$ переходе

Одним из важных параметров $p-n$ -перехода является его высота $-\varphi_0$. Для того чтобы ее оценить, вспомним, что в равновесном состоянии положение уровня Ферми в n - и p -областях должно быть одинаковым (см. рис. 29):

$$q \cdot \varphi_0 + (F_p - E_v) = F_n - E_v.$$

Известно, что уровень Ферми в n -области перехода относительно середины запрещенной зоны E_i определяется концентрацией донорной примеси в ней:

$$F_n - E_i = kT \cdot \ln(N_d / n_i).$$

Положение уровня Ферми в p -области перехода определяется концентрацией акцепторов:

$$F_p - E_i = -kT \cdot \ln(N_a / n_i).$$

Поэтому будем иметь:

$$\cdot q\varphi_0 = kT \cdot \left\{ \ln\left(\frac{N_d}{n_i}\right) + \ln\left(\frac{N_a}{n_i}\right) \right\} \quad \text{или} \quad \varphi_0 = \frac{kT}{q} \cdot \ln\left(\frac{N_a N_d}{n_i^2}\right).$$

Такова высота потенциального барьера между n - и p -областью в отсутствие освещения, электрического поля и других факторов, нарушающих состояние термодинамического равновесия в полупроводнике. Она оказывается тем выше, чем выше уровень легирования областей и ниже собственная концентрация полупроводника. Величина барьера p - n -перехода не может превысить ширину запрещенной зоны полупроводника, т.к. в уровень Ферми в n -

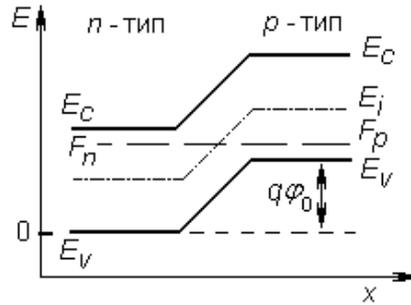


Рисунок 29 – Энергетическая диаграмма p - n перехода в отсутствие освещения

области, приблизившись к дну зоны проводимости E_C , оставит уровень донорной примеси ниже себя. Это вызовет заполнение данного уровня электронами, что означает снижение легирующего действия донорной примеси – уровень Ферми начнет опускаться, не достигнув E_C , и в итоге остановится при $F \cong E_C - E_d$. Аналогичная картина имеет место в p - области при увеличении концентрации акцепторной примеси: уровень Ферми остановится вблизи E_V .

Важной характеристикой p - n -перехода является его вольтамперная характеристика. Ее аналитическое выражение таково (приводится без вывода):

$$j(V) = j_s \cdot \left[\exp\left(\frac{qV}{kT}\right) - 1 \right].$$

Данная характеристика дает экспоненциальный рост тока при приложении положительного напряжений («+» на p области) и постоянство тока в области отрицательных напряжений: $j(V) = -j_s$, где ток j_s называется обратным током или током насыщения.

Пусть теперь к переходу приложено внешнее напряжение так, что на p -области находится «+» источника (рис. 29). Это приложенное напряжение частично компенсирует поле \bar{E}_K . В результате равновесие между диффузионным потоком носителей и током проводимости поля \bar{E}_K будет нарушено в пользу первого. В результате дырки из p -области устремятся в n -область, где они станут неосновными носителями заряда. Так как $p_p \gg p_n$, то эти дырки в n -области будут рекомбинировать с электронами. Однако вследствие конечного значения времени жизни дырок τ_p , рекомбинация произойдет не сразу, и поэтому в некоторой области за пределами перехода концентрация дырок будет больше p_n . По этим же причинам увеличится концентрация электронов в n -области, так как дополнительные электроны войдут из электрода для компенсации объемного заряда пришедших дырок: основные носители в n -области – электроны будут переходить в p -область, становясь там неосновными носителями, и постепенно рекомбинировать с основными носителями – дырками. Поэтому, слева от перехода концентрация электронов увеличивается, а также увеличивается концентрация дырок, которые войдут из левого электрода для компенсации объемного заряда электронов. Это явление получило название инжекции неосновных носителей. Распределение концентраций неосновных носителей при инжекции показано на рис. 33 штриховыми линиями.

В результате приложения напряжения V указанной полярности величина тока через p - n -переход будет расти по экспоненте с ростом величины V .

4.9.2. Фотоэлектрические свойства перехода. Барьерная фотоэдс.

Вернемся к вопросу о влиянии освещения на свойства полупроводников. При освещении $p-n$ -перехода высота барьера понижается на величину, пропорциональную уровню освещения. Изменение высоты барьера на границе раздела двух типов проводимости называют барьерной фотоэдс. Она возникает в том случае, когда свет генерирует носители заряда вблизи $p-n$ -перехода. Здесь следует различать две принципиально различные схемы включения освещаемого перехода. В первой из них $p-n$ -переход замкнут на внешнюю нагрузку без внешнего источника напряжения и является преобразователем световой энергии в электрическую (рис. 30, а). Такое включение реализует так называемый фотопентильный режим работы $p-n$ -перехода.

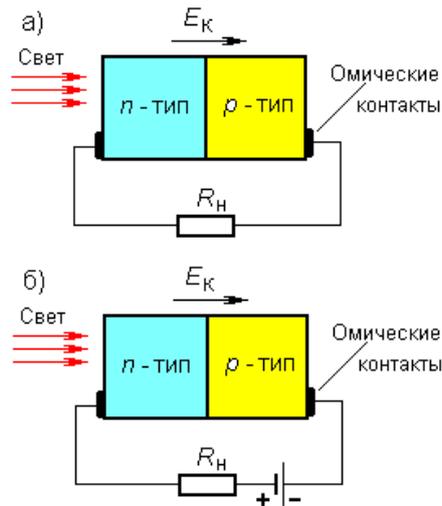


Рисунок 30 – Схемы включения $p-n$ перехода для регистрации барьерной фотоэдс.

Другая схема во внешней цепи содержит источник напряжения, включенный так, чтобы на $p-n$ -переходе было запирающее напряжение (рис. 30, б) – это так называемый фотодиодный режим. В схеме б существующий в цепи ток сильно изменяется с освещением. Поэтому изменяется падение напряжения на сопротивлении R . При правильном выборе напряжения источника и внешнего сопротивления величина сигнала в фотодиодном режиме может быть сделана больше, чем в фотопентильном. Величина вентильной фотоэдс на несколько порядков больше, чем эдс Дембера и объемной фотоэдс. Физическая причина ее возникновения заключается в том, что электрическое поле $p-n$ -перехода разделяет потоки фотодырок и фотоэлектронов, подошедших к переходу. Поясним это, предположив, что сильно поглощенное излучение падает на одну из граней кристалла, параллельную плоскости $p-n$ -перехода. Для определенности будем считать, что освещается n -область,

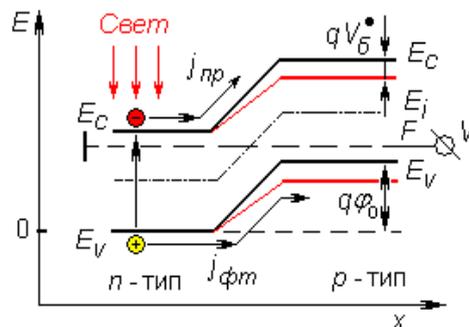


Рисунок 31 – Энергетическая диаграмма $p-n$ перехода при его освещении; формирование барьерной фотоэдс

излучение полностью поглощается в n -слое. Толщина этого слоя выбирается меньше длины диффузии неосновных носителей в нем (дырок). На концах прибора измеряется эдс холостого хода.

Фотоэлектроны и фотодырки в n -слое будут диффундировать вглубь слоя, и некоторая их доля, не успев прорекомбинировать, достигнет p - n -перехода. Однако для основных носителей (электронов) в p - n -переходе существует потенциальный барьер, и поэтому они не пройдут в p -область. Для неосновных носителей (дырок) потенциального барьера нет, иносителей (электронов) в p - n -переходе существует потенциальный барьер, и поэтому они не пройдут в p -область. Для неосновных носителей (дырок) потенциального барьера нет, и потому все достигшие переход дырки перенесутся полем перехода в p -область. При этом они создадут фототок $j_{\phi m}$. Если g – скорость световой генерации электронно-дырочных пар в n -области, а β – их доля, дошедшая до p - n -перехода, то $j_{\phi m} = qg\beta$. Вследствие появления фототока $j_{\phi m}$ p -область будет заряжаться положительно, а n -область – отрицательно, и между n - и p -областями появится разность потенциалов – барьерная фотоэдс V_{ϕ}^* . Снижение высоты барьера между n - и p - областями перехода на величину фотоэдс согласно вольтамперной характеристике приводит к увеличению темнового тока, созданного основными носителями заряда – j_{np} :

$$j_{np} = j_s \left[\exp\left(\frac{qV_{\phi}^*}{kT}\right) - 1 \right].$$

Он направлен обратно току $j_{\phi m}$. Его стационарное значение может быть найдено из условия равновесия, при котором полный ток через переход отсутствует: $j_{\phi m} = j_{np}$. Это дает выражение для вольтамперной характеристики p - n -перехода при его освещении:

$$j_{\phi m} = j_s \cdot \left[\exp\left(\frac{qV_{\phi}^*}{kT}\right) - 1 \right],$$

где j_s – ток насыщения (обратный ток) p - n -перехода. Отсюда можно найти величину вентильной фотоэдс:

$$V^* = \frac{kT}{q} \cdot \ln \left[1 + \frac{j_{\phi m}}{j_s} \right]. \quad (5.18)$$

Как следует из данного выражения, вентильная фотоэдс тем больше, чем меньше обратный ток перехода и выше уровень освещенности. Значит, для получения высокочувствительного фотодиода необходимо сильное легирование n - и p -областей p - n перехода.

Согласно выражению (5.18), при малых уровнях освещения барьерная фотоэдс равна

$$V_{\phi}^* = \frac{kT}{q} \cdot \ln \left[1 + \frac{j_{\phi m}}{j_s} \right] \approx \frac{kT}{q} \cdot \left(\frac{j_{\phi m}}{j_s} \right),$$

т.е. она прямо пропорциональна интенсивности падающего излучения:

$$j_{\phi m} = q\beta g = q\beta\eta\alpha I_0(1 - R)$$

Однако рост барьерной фотоэдс с увеличением уровня освещения перехода не беспределен. Поскольку V_{ϕ}^* направлено обратно к ϕ_0 – высоте барьера, то она не может быть больше высоты барьера. Следовательно, величина фотоэдс не может превышать контактную разность потенциалов между n - и p -областями, а ее максимальная величина сопоставима с шириной запрещенной зоны ΔE_g :

$$V_{\max}^* \approx \frac{\Delta E_g}{q}.$$

Так, для кремния $V_{\max}^* \approx 1 \text{ В}$, а для германия $V_{\max}^* \approx 0.6 \text{ В}$ при быстродействии, определяемом временем жизни неравновесных носителей заряда в $p-n$ -переходе. В зависимости от физико-химического состава полупроводникового материала оно колеблется в диапазоне от 10^{-3} до 10^{-6} с.

В настоящее время $p-n$ -переход – один из наиболее широко используемых фотодатчиков, т.к. он обладает высокой чувствительностью и приемлемым быстродействием, не требует посторонних источников напряжения (как фоторезистор), технологически совместим со схемой обработки сигналов, т.е. фотоприемник на основе барьерной фотоэдс можно в одном технологическом цикле изготовить со схемой усиления фотоэдс, детектирования и т.д. на одном полупроводниковом кристалле.

4.10. Лавинные фотодиоды

При описании фотодиодного режима указывалось, что в этом режиме можно получить более чувствительные фотодатчики, чем в фотовентильном. Если к фотодиоду приложить достаточно большое обратное напряжение, то напряженность электрического поля ε в $p-n$ -переходе может оказаться настолько высокой ($5 \cdot 10^5 - 1 \cdot 10^6$) В/см, что будут идти процессы ударной ионизации атомов ускоренными электронами. Это приведёт к росту числа неравновесных носителей (обратному току $p-n$ -перехода) при данном световом потоке. Рисунок 37 поясняет процесс рождения носителей заряда. Как темновые, так и фотоэлектроны приобретают в поле энергию большую, чем они теряют при соударении с атомами решетки. Если полученная электроном энергия превышает энергию ионизации – E_{ion} (обычно $\Delta E_g \ll E_{ion} < 1,5 \Delta E_g$), то электрон может создать новую электронно-дырочную пару. При протяженной области локализации сильного электрического поля возникшие электрон и дырка сами могут ускориться до энергии E_{ion} и совершить новые ионизации и так далее. В итоге в области $p-n$ -перехода будет лавинное размножение числа носителей заряда (рис.32).

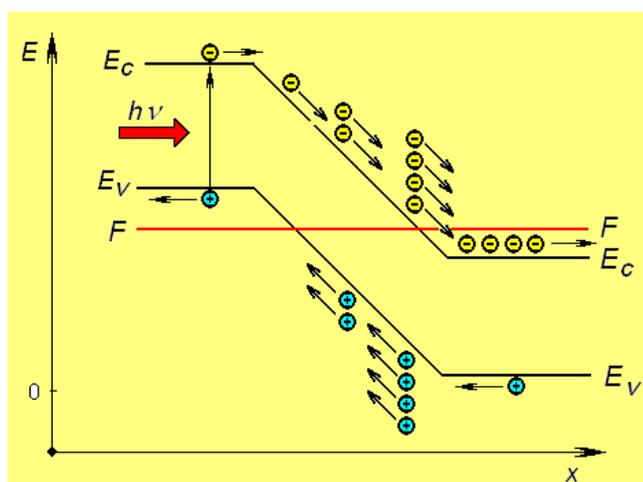


Рисунок 32 – Размножение неравновесных носителей заряда в лавинном фотодиоде

Увеличение тока в результате ионизации при больших ε характеризуется коэффициентом лавинного умножения носителей, определяя его следующим образом:

$$M = \frac{j(\varepsilon)}{j(0)},$$

где $j(\varepsilon)$ – плотность ток при обратных электрических полях ε , вызывающих лавинное размножение, а $j(0)$ – в отсутствии внешнего электрического поля, когда умножения носителей нет. Иначе говоря, $M = n/n_0$, где n_0 – число электронов, вошедших в область

поля, а n – число вышедших. Величина M заключена в диапазоне значений от 1 до ∞ и отображает общее число ионизаций, совершенных как первичными, так и вторичными электронами и дырками. Однако использование параметра, имеющего бесконечное значение, порой оказывается неудобным. Гораздо более удобно пользоваться параметром, имеющим ограниченный диапазон значений, но такой же физический смысл. Таким параметром является квантовый выход процесса ионизации, определяемый как:

$$N = \frac{n - n_0}{n}.$$

Диапазон значений N ограничен: $0 \leq N \leq 1$. Случай $N = 0$ соответствует отсутствию умножения в p - n - переходе, а случай $N > 0$ соответствует лавинному умножению носителей. Таким образом, $M = 1/(1 - N)$, где N – число актов ионизаций, приходящихся на один электрон, прошедший всю область локализации электрического поля.

Квантовый выход ионизации зависит от коэффициентов ударной ионизации (числа пар, созданных электроном или дыркой на единице пути) для электронов (α_i) и дырок (β_i), а также от ширины d области, в которой сосредоточено электрическое поле. При $\alpha_i = \beta_i$ и однородном электрическом поле шириной d имеем $N = \alpha_i d$.

В настоящее время разработаны кремниевые лавинные диоды, имеющие коэффициент умножения до 10^3 , поэтому лавинный диод пригоден для регистрации очень слабых световых сигналов. Его инерционность достигает 10^{-9} с, что позволяет регистрировать быстропротекающие процессы. Однако ударная ионизация в области p - n -перехода создает дополнительные шумы такого диода. Поэтому у лавинного фотодиода существует оптимальное значение прикладываемого напряжения, при котором обеспечивается коэффициент лавинного умножения $M > 1$, но уровень шумов еще невысок.

ВОПРОСЫ ПО ГЛАВЕ 4

1. Что такое время жизни неравновесного носителя заряда?
2. Какова кинетика избыточной концентрации неравновесных носителей заряда при непрерывном освещении полупроводника?
3. Что такое фотопроводимость полупроводника?
4. Какова кинетика фотопроводимости при импульсном освещении полупроводника?
5. Какие свойства фоторезистора определяет время релаксации фотопроводимости?
6. Что описывают полевые зависимости фоторезистора?
7. От чего зависят частотные свойства фотопроводимости и какие параметры можно найти из частотных зависимостей фоторезистора?
8. Что такое барьерная фотоэдс?
9. Какова основная физическая причина возникновения барьерной фотоэдс?
10. Каков принцип работы лавинного фотодиода?

5 ЭМИССИЯ ИЗЛУЧЕНИЯ ИЗ ТВЕРДЫХ ТЕЛ

5.1. Описание излучательных процессов в полупроводниках

Термин «эмиссия излучения» означает испускание света, например, твердым телом в результате какого-либо воздействия. Излучение, связанное с рекомбинацией электронов и дырок в полупроводниках, известно с начала XX века. По своей физической сущности излучательная рекомбинация представляет собой процесс, обратный процессу фотовозбуждения. В простейшем случае эмиссия излучения возникает, когда электроны из зоны проводимости рекомбинируют с дырками, находящимися в валентной зоне. Рекомбинация может сопровождаться излучением высвобождающегося избытка энергии в

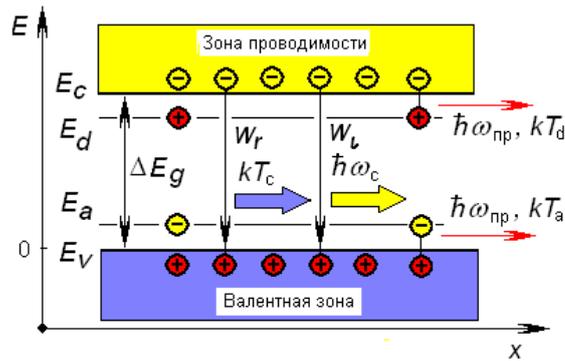


Рисунок 33 - Механизмы излучения света в полупроводниках

виде фотона с энергией $\hbar\omega_l = \Delta E_g$, когда атом «выстреливает» частичку света, либо кванта тепла (фонона) с энергией kT_c и передаваемого кристаллической решетке (рис.33). В последнем случае рекомбинация заканчивается «вздрагиванием» атома, на валентную оболочку которого сел электрон.

Аналогично может происходить излучательная рекомбинация через примесные уровни, когда квант света будет иметь энергию $\hbar\omega_{пр} = \Delta E_{d,a}$. В этом случае энергия эмитируемых фотонов будет меньше ширины запрещенной зоны и определяется энергетическим положением примесного уровня в запрещенной зоне полупроводника. Примесные переходы могут заканчиваться излучением фотона с энергией $\hbar\omega_{пр}$ или фонона с энергией kT_d или kT_a (рис.33).

Одним из основных условий наблюдения эмиссии излучения из твердого тела является преобладание процесса рекомбинации свободных носителей заряда над их генерацией (она может происходить за счет поглощения приповерхностными атомами фотонов, испущенных атомами объема полупроводника). Поэтому преобладание рекомбинации над генерацией может иметь место только в случае, если в зонах имеются неравновесные носители заряда, т.е. создана неравновесная ситуация. Для того чтобы рекомбинация неравновесных носителей заряда шла преимущественно с испусканием фотона, а не фонона, необходимо, чтобы в данном полупроводнике (или, в общем случае, твердом теле) вероятность излучательной рекомбинации $W_l = 1/\tau_l$ была бы больше вероятности безызлучательной – $W_r = 1/\tau_r$ (рис.33). Полная вероятность рекомбинации электронов и дырок $W_s = W_l + W_r$. Поэтому эффективность испускания светового излучения зависит от соотношения времен жизни неравновесных носителей, оканчивающихся излучательным - τ_l или безызлучательным - τ_r переходом электрона в валентную зону.

Для численной оценки эффективности излучательной способности твердого тела вводят два параметра. Первый из них называется внутренней квантовой эффективностью η . Он определяет число генерируемых фотонов на полное число прорекомбинировавших пар «электрон-дырка», характеризует способность атомов полупроводника излучать фотоны при

рекомбинации электронно-дырочных пар и определяется следующим образом:

$$\eta = \frac{1/\tau_l}{1/\tau_l + 1/\tau_r} = \frac{\tau_r}{\tau_r + \tau_l} . \quad 0 \leq \eta \leq 1.$$

Удобство данного параметра заключается в возможности сравнения излучающих свойств различных материалов без учета особенностей их частных свойств. У какого материала этот параметр больше, тот и предпочтительнее для создания твердотельных источников света.

Второй параметр, называемый внешней квантовой эффективностью и обозначаемый как ξ , характеризует долю сгенерированных в объеме полупроводника фотонов, вышедших из полупроводника через его поверхность в направлении наблюдателя. Действительно, для того чтобы наблюдать эмиссию излучения из полупроводника, необходимо, чтобы сгенерированные фотоны вышли из полупроводника – ведь распространяясь в кристалле на пути к излучающей поверхности фотоны могут быть поглощены другими атомами, находящимися в основном энергетическом состоянии. Физически этот процесс поглощения происходит по закону Бугера - Ламберта, который в этой ситуации будет записан в виде:

$$I(0) = (1 - R) \int_0^L I(x) \exp\left(-\frac{x}{\ell_\phi}\right) dx = I_{ген}(1 - R) \cdot \ell_\phi \left[\exp\left(-\frac{L}{\ell_\phi}\right) - 1 \right] \approx I_{ген}(1 - R) \cdot \ell_\phi \cdot \exp\left(-\frac{L}{\ell_\phi}\right),$$

где интегрирование вызвано необходимостью суммирования интенсивностей от всех атомов с длины кристалла. Здесь $I(0)$ – интенсивность вышедшего из полупроводника излучения; $I(x)$ – интенсивность излучения, сгенерированного атомами, находящимися на удалении x от излучающей поверхности в слое толщиной dx ; R – коэффициент отражения излучения от границы раздела «полупроводник – внешняя среда», L – толщина полупроводника в направлении на излучающую поверхность; ℓ_ϕ – длина свободного пробега фотона в полупроводнике. Из полученного результата следует, что для получения большой интенсивности $I(0)$ необходимо брать кристаллы малой толщины: $L < \ell_\phi$. По этой причине определение внешней квантовой эффективности в упрощенной форме имеет вид:

$$\xi = \frac{\ell_\phi}{\ell_\phi + L} . \quad 0 < \xi \leq 1$$

Значит, условие получения высокого выхода светового излучения заключается в том, чтобы сгенерированные фотоны могли выйти из полупроводника, не будучи поглощенными другими атомами на пути от излучающего атома до поверхности. Для этого необходимо, чтобы длина свободного пробега фотона была больше толщины образца: $\ell_\phi \gg L$.

На практике приходится встречаться с многими механизмами, ограничивающими мощность выходного излучения, в частности, с эффектом полного внутреннего отражения от границы раздела «твердое тело - вакуум» при распространении света из оптически более плотной среды в менее плотную. В настоящее время с данным ограничением успешно борются созданием специальной геометрии излучающей поверхности, ее просветлением и т.д. Рассмотрению этого и других подобных частных вопросов будет посвящен отдельный параграф в главе, посвященной конструкции и параметрам светодиодов.

Способы возбуждения полупроводника. Эмиссия излучения может быть + различными воздействиями на твердое тело. Неравновесные носители заряда можно инжектировать в однородный полупроводник путем оптического возбуждения или бомбардировкой заряженными частицами. Возникающее в результате этих операций излучение называется фотолюминесценцией или катодолюминесценцией соответственно. Возможно также создание неравновесных носителей за счет ударной ионизации в сильном электрическом поле. Этот способ возбуждения приводит к эмиссии, называемой электролюминесценцией.

Наиболее же эффективный и изящный метод создания больших концентраций неравновесных носителей заряда в полупроводниках заключается в использовании *p-n*-

перехода, включенного в прямом направлении. Именно на этом принципе основано действие полупроводниковых лазеров и светодиодов.

Другим видом эмиссии в полупроводниках является тепловое излучение. Этот излучательный процесс имеет место, когда между образцом полупроводника и окружающей средой существует разность температур. Тогда полупроводник будет излучать во всем спектре частот в соответствии со спектром черного тела. В результате установления термодинамического равновесия имеющийся избыток энергии полупроводник будет излучать в окружающее пространство в виде фотонов с длиной волны, задаваемой шириной запрещенной зоны и энергетическим спектром примесей. Ввиду конечных размеров образца часть эмитируемого излучения будет выходить из кристалла и не будет уравниваться тепловым потоком, входящим в образец, ввиду того, что полупроводник не находится в тепловом равновесии с окружающей средой. Поэтому эмиссионные свойства полупроводника отличаются от свойств абсолютно черного тела, и изменение наблюдаемой излучательной способности определяется величиной и структурой коэффициента поглощения. Это утверждение справедливо независимо от природы процесса поглощения. Такую люминесценцию называют термолюминесценцией.

Флюоресценция – это люминесценция, которая происходит только во время действия источника возбуждения неравновесных пар. Фосфоресценция – это люминесценция, которая продолжается после действия возбуждения.

В общем случае определение люминесценции таково. Люминесценция - есть избыток излучения твердого тела над тепловым излучением, если это избыточное излучение обладает длительностью, значительно превышающей период испускаемых световых волн. Данное определение отделяет люминесценцию от теплового излучения в равновесных условиях и дает право отнести ее к разделу неравновесных явлений. Это определение выделяет люминесценцию из всех других видов неравновесного излучения, таких как отражение и рассеяние света, излучение Вавилова-Черенкова и т.д. Как следует из определения в отличие от перечисленных излучений, в случае люминесценции акты поглощения и испускания света разделены промежуточными процессами, что и приводит к продолжительному существованию свечения после прекращения возбуждения.

5.2. Зависимость эмиссии от уровня легирования

Закономерным является вопрос о том, есть ли у разработчиков оптоэлектронных приборов возможность управлять скоростью излучательной рекомбинации, например, изменением технологических режимов выращивания полупроводника или, например, внесением специальных добавок в состав полупроводникового материала. Наиболее простым и доступным способом управления физико-химическим составом полупроводника является операция его легирования.

В этой связи проанализируем зависимость времени излучательной рекомбинации в полупроводнике от его параметров. Для этого рассмотрим прямые переходы «зона проводимости – валентная зона» в полупроводнике, выведенном из равновесного состояния. Пусть n_i – собственная концентрация носителей. Из физических соображений следует, что скорость излучательной рекомбинации R_l в полупроводнике с произвольной концентрацией электронов пропорциональна концентрации электронов n и концентрации пустых состояний (дырок) p : чем больше электронов, тем выше скорость (есть чему рекомбинировать); но скорость рекомбинации растет и с увеличением концентрации дырок (есть с чем рекомбинировать). Для упрощения дальнейших физических рассуждений и математических выкладок, будем считать, что избытки концентраций электронов Δn и дырок Δp малы по сравнению с равновесными значениями n_0 и p_0 . Для того чтобы в первом приближении математически описать прямо пропорциональную зависимость скорости излучательной рекомбинации R_l от n и p , необходимо взять произведение полных концентраций

носителей заряда: $n_0 + \Delta n$ и $p_0 + \Delta p$. Таким простым способом в процесс излучательной рекомбинации включены не только неравновесные носители заряда, но и равновесные. Тем самым учтено то обстоятельство, что рекомбинация неравновесной дырки возможна с равновесным электроном, а неравновесного электрона – с равновесной дыркой.

Пронормируем указанное произведение концентраций на квадрат собственной концентрации носителей заряда. Это позволит упростить физическую интерпретацию коэффициента пропорциональности в обсуждаемой зависимости. Тогда скорость излучательной рекомбинации можно записать в виде:

$$R_l = \frac{n \cdot p}{n_i^2} R_i,$$

где R_i – скорость излучательной рекомбинации в собственном полупроводнике. Поскольку, по определению, скорость любой рекомбинации обратно пропорциональна времени жизни неравновесных носителей заряда, то из данного выражения следует, что время жизни носителей в зоне проводимости, заканчивающееся излучением фотона, уменьшается при возрастании концентраций и электронов, и дырок в соответствующих зонах.

Используем введенные ранее обозначения концентраций и их отклонений от равновесных значений. Выражения для полных концентраций электронов и дырок при наличии неравновесных концентраций: $n = n_0 + \Delta n$, $p = p_0 + \Delta p$. Тогда скорость рекомбинации также может быть представлена как сумма скоростей в равновесии и ее отклонения:

$$R_l^0 + \Delta R_l = \frac{(n_0 + \Delta n)(p_0 + \Delta p)}{n_i^2} R_i \cong \frac{n_0 p_0 + (n_0 + p_0) \Delta n}{n_i^2} R_i,$$

т.к. считаем $\Delta n = \Delta p$. Поэтому имеем:

$$R_l^0 = \frac{n_0 p_0}{n_i^2} R_i, \quad \Delta R_l \cong \frac{(n_0 + p_0) \cdot \Delta n}{n_i^2} R_i.$$

По определению скорость рекомбинации есть $\Delta n / \tau_l$, где τ_l – время жизни электрона в зоне проводимости, заканчивающееся излучением света:

$$\Delta R_l = \frac{\Delta n}{\tau_l} = \frac{(n_0 + p_0) \Delta n}{n_i^2} R_i.$$

Отсюда находим выражение для определения времени излучательной рекомбинации:

$$\tau_l = \frac{\Delta n}{\Delta R_l}.$$

Тогда найдём, что

$$\tau_l = \frac{1}{R_i} \cdot \frac{n_i^2}{n_0 + p_0}$$

Время излучательной рекомбинации в собственном полупроводнике $\tau_l^{(i)}$ является фундаментальным параметром для любого полупроводникового материала. Оно определяется скоростью излучательной рекомбинации и концентрацией носителей заряда в собственном полупроводнике: $\tau_l^{(i)} = \frac{n_i}{2R_i}$. Тогда, объединив полученные выражения,

найдем окончательное выражение для времени излучательной рекомбинации в легированном полупроводнике:

$$\frac{\tau_l}{\tau_l^{(i)}} = \frac{2n_i}{n_0 + p_0}. \quad (5.1)$$

Численные значения времени излучательной рекомбинации в различных беспримесных полупроводниках сильно отличаются друг от друга. Экспериментальные исследования показывают, что, например, для Si $\tau_l^{(i)} \approx 4,6$ часа, для Ge $\tau_l^{(i)} \approx 0,6$ с, InSb $\tau_l^{(i)} \approx 6 \cdot 10^{-7}$ с.

Согласно выражению (5.1), зависимость времени излучательной рекомбинации «зона – зона» от концентрации носителей заряда в полупроводнике имеет вид, оказанный на рис. 34. По оси абсцисс отложено отношение концентрации электронов n_0 к концентрации дырок p_0 . Поэтому правая половина рисунка соответствует полупроводнику n -типа проводимости, а левая – p -типу. Из рисунка видно, что с увеличением концентрации основных носителей время излучательной рекомбинации быстро уменьшается по величине, а скорость излучательной рекомбинации растет. Такое поведение времени излучательной рекомбинации объясняется увеличением вероятности рекомбинации неравновесного неосновного носителя за счет роста концентрации основных носителей заряда. Дополнительные расчеты показывают, что уменьшение ширины запрещенной зоны также снижает время излучательной рекомбинации.

Таким образом, если для изготовления источника излучения использовать объемный монокристалл, то для повышения интенсивности его эмиссии излучения необходимо брать его с высокой концентрацией основных носителей и малой шириной запрещенной зоны.

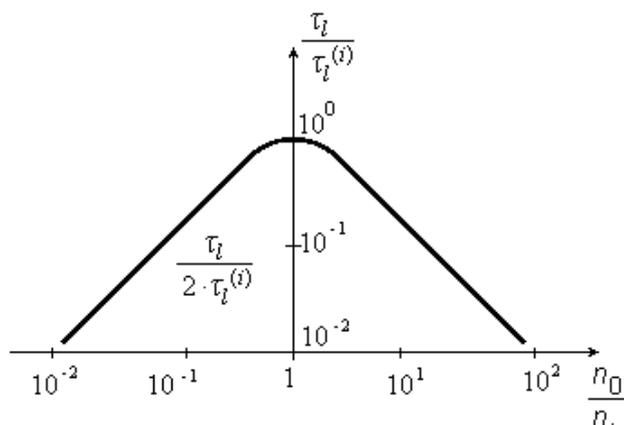


Рисунок 34 – Зависимость времени излучательной рекомбинации от концентрации легирующей примеси по модели «пуля - мишень»

5.3. Спонтанное и вынужденное излучение атома

Для понимания условий генерации оптического излучения в твердых телах рассмотрим процесс поглощения и излучения света в атоме. Из всех квантовых состояний, в которых может находиться атом, рассмотрим два состояния с энергиями E_1 и E_2 . Уровень E_1 описывает основное (невозбужденное) состояние, а E_2 первое возбужденное (рис. 35). Пусть теперь на такой атом падает излучение, имеющее спектральную плотность энергии U_ν (размерность Вт/Гц). Невозбужденный атом, поглощая энергию излучения $h\nu$, перейдет в возбужденное состояние E_2 . Переходы с поглощением кванта света совершаются тем чаще, чем больше величина U_ν . Если N_1 – число атомов в невозбужденном состоянии, а $U_\nu B_{12}$ – вероятность перехода с поглощением кванта при единичной спектральной плотности падающего излучения, то число переходов с поглощением фотона в единицу времени в единице объема будет

$$\partial N_{12} = B_{12} U_\nu N_1.$$

Спонтанное излучение. Атом, находясь в возбужденном состоянии в результате поглощения кванта света, через некоторый промежуток времени, подчиняясь закону

Здесь использовано выражение для определения среднего значения случайной величины по известной плотности вероятности случайного процесса. Из приведенного вычисления следует, что средняя продолжительность жизни атома в возбужденном состоянии равна

$$\tau = \frac{T}{N_{20}} = \frac{1}{A_{21}} . \quad (3.1)$$

Поэтому релаксацию интенсивности свечения можно переписать в виде:

$$I = I_0 \cdot \exp\left(-\frac{t}{\tau}\right).$$

Таким образом, после прекращения возбуждения интенсивность спонтанного излучения во времени уменьшается по экспоненциальному закону с постоянной времени $\tau = 1/A_{21}$. Поскольку спонтанные переходы являются случайными, то спонтанное излучение имеет статистический характер, оно не когерентно, т.е. генерируются фотоны с разной поляризацией, распространяющиеся в произвольных направлениях и имеющие произвольные фазы, хотя частоты излучаемых квантов света близки. Это обусловлено тем, что в силу случайного характера акты спонтанного излучения происходят независимо друг от друга.

Вынужденное излучение. Однако возбужденный атом может перейти в невозбужденное состояние и под действием внешнего излучения. Такой переход называется вынужденным, а излучение, создаваемое атомами, носит название вынужденного, стимулированного или индуцированного. Индуцированные переходы будут совершаться тем чаще, чем больше плотность энергии возбуждающего света U_ν . Поэтому вероятность этих переходов в единицу времени будет $U_\nu B_{21}$. Здесь B_{21} – вероятность индуцированного перехода атома с уровня E_2 на уровень E_1 при единичной интенсивности падающего излучения. Если N_2 – число атомов в возбужденном состоянии, то число индуцированных переходов в единицу времени в единице объема составит

$$dN_2'' = U_\nu B_{21} N_2 .$$

Следовательно, внешнее излучение с частотой ν вызывает не только поглощение, но и добавочное, индуцированное излучение.

Среднее число переходов с излучением кванта света в единицу времени в единице объема определяется спонтанными и вынужденными переходами и равно:

$$dN_{21} = dN_{21}' + dN_{21}'' = (A_{21} + U_\nu B_{21}) N_2 .$$

В стационарном состоянии на основании принципа детального равновесия числа прямых и обратных переходов в единицу времени равны. Следовательно, $dN_{21} = dN_{12}$ или $(A_{21} + U_\nu B_{21}) N_2 = U_\nu B_{12} N_1$. Здесь B_{12} – вероятность перехода атома с уровня E_1 на уровень E_2 под действием излучения единичной интенсивности. Отсюда следует, что:

$$\frac{N_2}{N_1} = \frac{U_\nu B_{12}}{A_{21} + U_\nu B_{21}} . \quad (5.2)$$

Поскольку равновесие имеет статистический характер, то отношение N_2/N_1 можно выразить при помощи распределения Больцмана, согласно которому для системы, имеющей состояния с энергией E_1, E_2, \dots, E_m , вероятность пребывания в состоянии E_m равна

$$W_m = C g_m \cdot \exp\left(-\frac{E_m}{kT}\right) ,$$

где g_m – фактор вырождения данного уровня, показывающий, сколько атомом находятся в данном состоянии. Поэтому можно записать:

$$\frac{N_2}{N_1} = \frac{g_2 \cdot \exp\left(-\frac{E_2}{kT}\right)}{g_1 \cdot \exp\left(-\frac{E_1}{kT}\right)} = \frac{g_2}{g_1} \cdot \exp\left(-\frac{E_2 - E_1}{kT}\right). \quad (5.3)$$

Из сравнения равенств (3.2) и (3.3) следует, что:

$$\frac{U_\nu B_{12}}{A_{21} + U_\nu B_{21}} = \frac{g_2}{g_1} \cdot \exp\left(-\frac{E_2 - E_1}{kT}\right).$$

С учетом того, что $E_2 - E_1 = h\nu$, получим:

$$U_\nu = \frac{g_2 A_{21}}{g_1 B_{12} \cdot \exp\left(\frac{E_2 - E_1}{kT}\right) - g_2 B_{21}}. \quad (5.4)$$

Для определения коэффициентов A_{21} и B_{21} используем предельное условие. При $T \rightarrow \infty$ плотность энергии U_ν должна стремиться к бесконечности (при сильном нагревании твердое тело излучает больше, чем при слабом), а это значит, что знаменатель в (3.4) стремится к нулю или $g_1 B_{12} = g_2 B_{21}$. С учетом этого выражение (3.4) переписывается в виде:

$$U_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \left/ \left(\exp\left(\frac{h\nu}{kT}\right) - 1 \right) \right.$$

При $h\nu \ll kT$ в первом приближении имеем:

$$\exp\left(\frac{h\nu}{kT}\right) \approx 1 + \frac{h\nu}{kT} \quad \text{и} \quad U_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \cdot \frac{kT}{h\nu}. \quad (5.5)$$

С другой стороны, при условии $h\nu \ll kT$ справедлива формула Релея-Джинса: $U_\nu = 8\pi\nu^2 kT / c^3$. Из ее сопоставления с выражением (5.5) найдем:

$$\frac{g_2 A_{21}}{g_1 B_{12}} = \frac{8\pi\nu^2}{c^3} h\nu.$$

Отсюда следует: $A_{21} = \frac{8\pi\nu^2}{c^3} h\nu \frac{g_1}{g_2} B_{12}$. Тогда окончательно получим:

$$A_{21} = \frac{8\pi\nu^2}{c^3} h\nu \cdot B_{21}. \quad (5.6)$$

Данное выражение показывает, что вероятность спонтанного перехода однозначным образом связана с вероятностью вынужденного излучения, что указывает на наличие у атома единого механизма по излучению (вынужденному или спонтанному) и поглощению кванта света. Важно отметить, что если у данного вещества вероятность вынужденного перехода больше, чем у других веществ, то и вероятность спонтанного излучения в аналогичных условиях измерений также должна быть больше. При вынужденном переходе генерируется фотон, имеющий ту же частоту, направление и фазу, что и фотон, вызвавший этот переход. Излучение, обладающее указанными свойствами, называют когерентным.

5.4. Стимулированное излучение твердых тел

Рассмотрим условия, при которых будет наблюдаться стимулированное излучение большого числа атомов. Пусть пучок параллельного монохроматического света интенсивностью I падает на объем, в котором сосредоточены атомы. Направление распространения света свяжем с осью x . При прохождении слоя вещества толщиной Δx интенсивность света уменьшится на величину $-\Delta I = \alpha \cdot I \cdot \Delta x$. Это уменьшение происходит

из-за переходов в количестве $U_\nu B_{12} \cdot h\nu \cdot N_1 \Delta x$ с поглощением энергии и индуцированных переходов с излучением энергии в количестве $U_\nu B_{21} \cdot h\nu \cdot N_2 \Delta x$. В результате имеем:

$$-\Delta I = \alpha \cdot I \cdot \Delta x = (B_{12}N_1 - B_{21}N_2) \cdot U_\nu \cdot h\nu \cdot \Delta x.$$

Если учесть, что плотность излучения U_ν в параллельном пучке сечением 1 см^2 , распространяющемся со скоростью c , равна: $U_\nu = I/c$, то коэффициент поглощения приобретает вид:

$$\alpha = \frac{h\nu}{c} (B_{12}N_1 - B_{21}N_2) = \frac{h\nu}{c} \left(1 - \frac{g_1 N_2}{g_2 N_1} \right) B_{12} N_1.$$

Учитывая выражение (5.3), получим ($E_2 - E_1 = h\nu$):

$$\alpha = B_{12} N_1 \left[1 - \exp\left(-\frac{h\nu}{kT}\right) \right] \frac{h\nu}{c}, \quad (5.7)$$

Поскольку в состоянии термодинамического равновесия показатель экспоненты всегда отрицателен, то квадратные скобки в (5.7) дают положительную величину. Это означает, что в случае термодинамического равновесия при любой конечной температуре суммарный коэффициент поглощения положителен, т.е. имеет место поглощение света, а индуцированные переходы лишь уменьшают его. Причина этого заключается в том, что в равновесном состоянии число невозбужденных атомов всегда больше числа возбужденных ($N_1 \gg N_2$). Поэтому количество переходов с уровня E_1 будет больше, чем с E_2 . Если же мы хотим определить условия усиления света атомной средой, то надо сделать коэффициент поглощения света отрицательным. Тогда по закону Бугера – Ламберта интенсивность света будет увеличиваться по мере его распространения по объему с возбужденными атомами.

Как же добиться отрицательного значения коэффициента поглощения α ? Из (5.7) следует, что это возможно, если показатель экспоненты положителен, а для этого необходимо, чтобы абсолютное значение температуры кристалла было отрицательным: $T < 0$. Данное условие выглядит невозможным, но это только на первый взгляд. Продолжим поиск путей создания усилителя света квантовой системой.

Из распределения Больцмана (5.3) следует, что температура тела связана с заполнением уровней энергии основного E_1 и возбужденного состояния E_2 :

$$T = -\frac{E_2 - E_1}{k \cdot \ln(g_1 N_2 / g_2 N_1)}.$$

Поскольку $E_2 > E_1$, а в равновесии заполнение верхнего уровня всегда меньше, чем нижнего $g_1 N_2 < g_2 N_1$, то абсолютное значение температуры твердого тела T всегда положительно. Однако если выполняется условие:

$$g_1 N_2 > g_2 N_1, \text{ тогда } T < 0. \quad (5.8)$$

Таким образом, согласно выражению (5.8), для того чтобы создать среду с отрицательным поглощением (с усилительными свойствами), необходимо получить неравновесное состояние вещества, при котором число возбужденных атомов было бы больше числа невозбужденных. Заселенность энергетических состояний атомов, удовлетворяющих условию (3.8), носит название «инверсной заселенности», т.е. обратной заселенности по отношению к равновесному распределению атомов по уровням энергии. Реализовав инверсию населенности за счет накачки, получим кристалл, способный усиливать свет с энергией кванта, равной расстоянию между первым возбужденным состоянием квантовой системы и основным ее состоянием.

5.5. Принцип создания генератора света

Из радиотехники известно, что любой усилитель радиосигналов с коэффициентом усиления $G \gg 1$ можно превратить в генератор синусоидальных сигналов. Для этого

необходимо выполнить несколько условий по преобразованию усилителя (рис.36):

1. применить в нем положительную обратную связь, т.е. подать с выхода усилителя на его вход напряжение в фазе с входным напряжением;

2. коэффициент обратной связи K_{OC} должен быть достаточно велик, чтобы выполнялось условие $K_{OC}^{(+)}(f_0) \cdot G(f_0) > 1$. Тогда после каждого прохода переменного напряжения через усилитель и цепь обратной связи на выходе устройства будет напряжение, превосходящее напряжение при предыдущем проходе. Увеличение выходного напряжения будет происходить до тех пор, пока амплитуда переменного напряжения на выходе усилителя не сравнится с напряжением его питания;

3. частота получаемого выходного сигнала f_0 будет определять частотными свойствами, как усилителя, так и цепи обратной связи. Прежде всего, усилитель должен иметь частотную характеристику, включающую частоту f_0 ;

4. для получения синусоидального напряжения на выходе в цепь обратной связи необходимо включить частотный фильтр с требуемым коэффициентом обратной связи.

Используя данный принцип в случае оптических колебаний можно создать усилитель света. Для превращения его в генератор необходима положительная обратная связь. Элементом, способным с выходного торца кристалла вернуть свет на его входной торец, является зеркало. Значит, расположив кристалл между зеркалами на его переднем и заднем торце объема с рабочим веществом, получим обратную связь. С тем, чтобы генерируемое в кристалле излучение не выходило из кристалла в окружающее пространство, зеркала следует расположить параллельно друг другу. Положительный характер этой связи обеспечивает настройка расстояния между зеркалами так, чтобы на нем укладывалось целое число полуволен генерируемого излучения.

Максимальное увеличение интенсивности излучения будет происходить при многократном прохождении световой волны по кристаллу в результате отражения излучения от торцов кристалла. Многократно отражаться будут лишь те лучи, которые падают перпендикулярно на торцевые поверхности кристалла. В силу этого излучение, вызванное индуцированными переходами, будет сосредоточено в малом телесном угле.

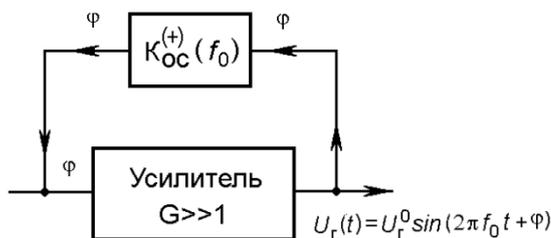


Рисунок 36 – Блок-схема автогенератора

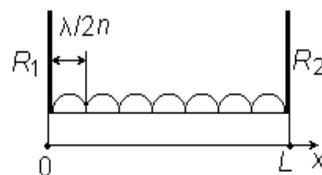


Рисунок 37 – Образование положительной обратной связи в электромагнетизме

$\Delta\Omega \approx l^2 / L$ Здесь l^2 - площадь зеркала, L - расстояние между зеркалами. В результате возникает узконаправленный поток излучения большой интенсивности, так как яркость его, в сравнении с всенаправленным излучением, возрастает в $4\pi/\Delta\Omega$ - раз. Спонтанное же излучение распределяется равномерно во всех направлениях в пределах телесного угла 4π . Поэтому оно уходит из кристалла через боковые поверхности, равномерно окрашивая кристалл в цвет генерируемого излучения. Индуцированное излучение является когерентным. Оно распространяется в том же направлении и имеет ту же фазу, что и первичное. Поэтому лазер можно рассматривать как оптический резонатор, внутри которого образованы стоячие волны. Их число кратно длине кристалла L (рис.37):

$$m \frac{\lambda}{2n} = L,$$

где m – целое число; λ/n – длина волны в кристалле, n – его показатель преломления.

Поскольку на максимум спектрального распределения спонтанного излучения приходится больше фотонов, на этой частоте происходит больше переходов, чем на краях спектра. Это приводит к значительному сужению спектра излучения и, следовательно, усилению интенсивности излучения в его максимуме. Если интенсивность пика излучения растет сверхлинейно с уровнем возбуждения, то процесс называется «сверхсвечением».

Определим условия возникновения светового возбуждения в кристалле, помещенном в резонатор. Пусть резонатор имеет длину L , а R_1, R_2 – коэффициенты отражения на гранях. Если α – коэффициент поглощения лазерного излучения посторонними механизмами, то излучение, имевшее на входе в резонатор интенсивность I_0 , прошедшее через резонатор и вернувшееся в ту же точку, будет иметь интенсивность (рис.38)

$$I = I_0 \cdot R_1 R_2 \cdot \exp(2gL - 2\alpha L).$$

Следовательно, условие возникновения лазерного эффекта, предполагающее нарастание интенсивности излучения после однократного прохождения активного вещества, заключается в превышении коэффициента усиления над коэффициентом потерь, будет

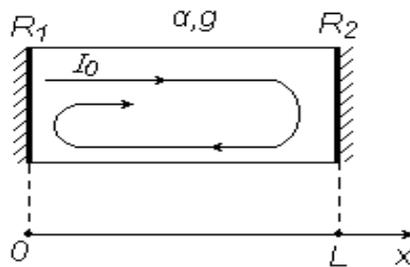


Рисунок 38 – Однократное прохождение луча в резонаторе лазера

$$g > \alpha - \frac{1}{2L} \cdot \ln(R_1 \cdot R_2). \quad (5.9)$$

Если не принять специальных мер, то ввиду низкой эффективности излучательной рекомбинации только малая часть энергии, затраченной на создание инверсии населенности, выделяется в виде лазерного излучения. Остальная часть идет на нагревание кристалла, что в ряде случаев заставляет использовать импульсный режим накачки.

5.6. Линия излучения твердого тела

Кроме двух параметров, с помощью которых можно численно охарактеризовать эффективность эмиссии излучения твердым телом, вышедшее излучение характеризуют зависимостями его параметров от независимых переменных. Наиболее используемая из них – зависимость интенсивности свечения от энергии изученного фотона или его длины волны,

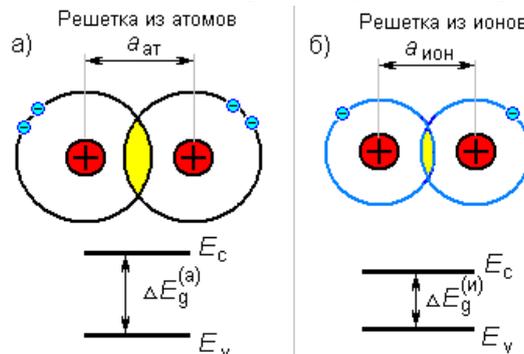


Рисунок 39 – Изменение в перекрытии волновых функций валентных электронов у атомов и ионов

называемая спектром излучения твердого тела.

Сдвиг Франка – Кордона. Поскольку поглощение света твердым телом и излучения света можно рассматривать как два взаимно обратных процесса, то в первом приближении можно ожидать, что спектры поглощения и излучения должны быть одинаковыми. Однако, спектры излучения твердого тела отличаются от соответствующих спектров изолированных атомов примеси двумя основными чертами. Во-первых, излучение твердого тела обычно представляет собой набор широких пиков, тогда как спектр поглощения состоит преимущественно из «полочек» так, что спектр излучения напоминает собой продифференцированный спектр поглощения (рис. 39). Во-вторых, спектр излучения, как правило, смещен относительно спектра поглощения в сторону длинных волн и, как следствие этого излучаемая энергия всегда меньше поглощаемой на той же длине волны. Разницу между этими энергиями называют сдвигом Франка-Кордона. Эта особенность обязана различием взаимодействия центра поглощения и центра излучения с решеткой.

Поясним сказанное (рис.39). Излучение кванта света производится атомом, который, ввиду полного заполнения электронных оболочек, имеет наибольший диаметр. По этой причине внешние оболочки соседних атомов сильно перекрываются, что по влиянию на ширину запрещенной зоны полупроводника аналогично сближению атомов, т.е. уменьшению межатомного расстояния. Излучение же кванта света производится ионом, диаметр которого меньше из-за частичного заполнения внешних оболочек. Значит, перекрытие внешних оболочек соседних ионов меньше, чем в случае атомов. Тогда, согласно модели расщепления уровней энергии атомов при их сближении (рис. 5) и рис. 39 условная постоянная решетки, составленной из атомов $a_{ат}$, приводит к меньшей ширине запрещенной зоны кристалла (на рис.5 движение вправо по оси абсцисс) по сравнению с решеткой, составленной из ионов $a_{ион}$. Если же для этих случаев ширины запрещенных зон различны

$\Delta E_g^{(ат)} > \Delta E_g^{(и)}$, то и энергии фотонов для переходов «зона – зона» также будут различны, что и приводит к сдвигу Франка-Кордона: $\Delta\lambda_{фк} = hc \left(1/\Delta E_g^{(и)} - 1/\Delta E_g^{(ат)} \right)$.

Те же механизмы взаимодействия света и твердого тела, рассмотренные при изучении спектра поглощения излучения, имеют место при рассмотрении спектра излучения. Это прямые и непрямые переходы «зона – зона», переходы с участием экситонов, примесных состояний и т.д. Наиболее важными являются переходы «зона – зона», которые в спектре излучения выглядят как пики при энергиях, равных ширине запрещенной зоны или энергетического зазора между зоной и примесным уровнем.

Лоренцева линии излучения. Важными характеристиками спектра излучения являются форма линии излучения и ее ширина. Для определения формы линии излучения твердого тела вспомним, что потери интенсивности вынужденных переходов обусловлены наличием спонтанных переходов. Тогда выходная интенсивность излучения будет

$$\frac{\partial I}{\partial t} + \frac{I}{\tau} = 0 \quad . \quad (5.16)$$

Отсюда находим закон изменения интенсивности со временем

$$I(t) = I_0 \cdot \exp\left(-\frac{t}{\tau}\right) . \quad (5.17)$$

Это выражение совпадает с выражением для интенсивности спонтанного излучения атомом, если считать, что $\tau = 1/A_{21}$, т.е. что время релаксации спонтанного излучения атома обратно пропорционально вероятности спонтанного перехода с верхнего уровня на нижний.

Выражение (5.17) указывает на то, что генерируемая интенсивность излучения не является чисто монохроматической, а содержит экспоненциально убывающую амплитуду. Значит, спектр спонтанного излучения атома содержит более одной составляющей на частоте энергетического перехода между уровнями энергии ω_0 . Чтобы найти спектр спонтанного излучения, необходимо найти Фурье – образ выражения (5.17), т.е. представить

его в виде суммы (или интеграла – как получится) гармонических составляющих вида

$$A(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(t) \cdot \exp(-i\omega t) dt. \quad (5.18)$$

Гармоническое изменение амплитуды колебания представим в экспоненциальной форме:

$$A(t) = 0.5A_0 \cdot \exp(-\gamma t/2) \cdot \exp(i\omega_0 t).$$

Подставляя данное выражение в интеграл Фурье (5.18) и интегрируя, получим:

$$A(\omega) = \frac{A_0}{2\pi} \cdot \frac{1}{i(\omega - \omega_0) + \gamma/2}.$$

Нас интересует спектр мощности излучения, который связан со спектром амплитуды

$$|A(\omega)|^2 = A^*(\omega) \cdot A(\omega) = \frac{A_0^2}{4\pi^2} \cdot \frac{1}{(\omega - \omega_0)^2 + (\gamma/2)^2}.$$

Зависимость вида

$$g_L(\omega) = \frac{g_0}{(\omega - \omega_0)^2 + (\gamma/2)^2} \quad (5.19)$$

называется **кривой Лоренца**, где g_0 - нормировочный коэффициент, определяемый из условия нормировки

$$\int_{-\infty}^{\infty} g_L(\omega) d\omega = 1.$$

Это дает $g_0 = \gamma/2\pi$. Зависимость вида (5.19) имеет вид колоколообразной кривой (перевернутой параболы) с максимумом, равным $g(\omega_0) = 2/\pi\gamma$ на частоте ω_0 . На частотах, удовлетворяющих условию: $(\omega - \omega_0) = \pm \tau/2$, спектральная плотность уменьшается в два раза. По этой причине интервал частот $\Delta\omega = 2(\omega - \omega_0) = \gamma$ называют шириной линии излучения. Она представляет собой естественную ширину спектральной линии. Отсюда следует фундаментальное соотношение спектрального анализа

$$\Delta\omega \cdot \tau = 1. \quad (5.20)$$

Приведенные выше результаты получены из рассмотрения излучающих свойств колеблющегося электрического диполя в приближении классической электродинамики. В действительности атом является квантовой системой, и потому корректным было бы рассматривать его свойства с позиций квантовой механики. Такое рассмотрение было проведено и показано, что получаемые в этом случае основные результаты полностью совпадают с выводами классической модели излучающего диполя. Причина совпадения результатов кроется в том, что в квантовом описании в качестве электрического диполя фактически выступает электронное облако той или иной конфигурации в зависимости от значения квантовых чисел. Переход с одного состояния на другое сопровождается сменой конфигурации, а значит и смещения всего электронного облака, что эквивалентно осцилляции диполя, а значит и излучению электромагнитной энергии.

Наиболее важным отличием квантового подхода является установление правил отбора переходов между различными уровнями энергии спектра атома. Эти правила таковы:

- изменение главного квантового числа может быть любым, т.е. Δn - любое;
- орбитальное квантовое число l может изменяться только на единицу: $\Delta l = \pm 1$;
- магнитное квантовое число m_l может изменяться на ноль и ± 1 , т.е. $\Delta m_l = 0, \pm 1$.

Это означает, что переход между уровнями E_n и E_m будет разрешен, если уровни с квантовыми числами n и m подчиняются приведенным правилам отбора.

На основе полученных выражений для вероятностей переходов квантово-механический подход к описанию излучающих свойств атома позволяет оценить их значения. Для

спонтанного перехода получено, что имеет величину порядка: $A_{nm} \approx 10^8 \text{ с}^{-1}$. Следовательно, время релаксации атомной системы составляет $\tau \approx 10^{-8} \text{ с}$, если релаксация происходит только за счет спонтанных переходов. Если действуют еще другие механизмы релаксации, например, безизлучательные переходы, то время релаксации будет меньше указанного.

Уровни энергии с малыми временами релаксации энергии называют лабильными. Уровни энергии, переходы с которых на более низкие уровни энергии запрещены при дипольных взаимодействиях, называют метастабильными. Время жизни атомов в этих состояниях порядка 10^{-3} с и больше.

5.7. Светодиоды на основе гомогенного *p-n* перехода

Инжекция неосновных носителей в *p-n*-переходе, включенном в прямом направлении, может привести к высокой эффективности преобразования электронов и дырок в фотоны даже в отсутствии лазерного эффекта. Процесс спонтанного излучения может происходить в полупроводниках, как с прямыми, так и непрямыми переходами посредством одного из механизмов излучательной рекомбинации, о которых говорилось ранее. Хотя внутренняя квантовая эффективность может быть близка к единице, внешняя эффективность обычно мала из-за действия эффекта полного внутреннего отражения (рис.40). Суть этого действия такова: согласно закону Снелиуса при падении света из оптически менее плотной среды в более плотную ($n_2 > n_1$) угол падения будет больше угла преломления (рис.40, а). Их взаимосвязь описывается выражением

$$\frac{\sin \alpha}{\sin \beta} = \frac{n_2}{n_1}.$$

Если сменить направление распространения луча, направив его из оптически более плотной среды в менее плотную (рис.40,б), то соотношение углов сохранится:

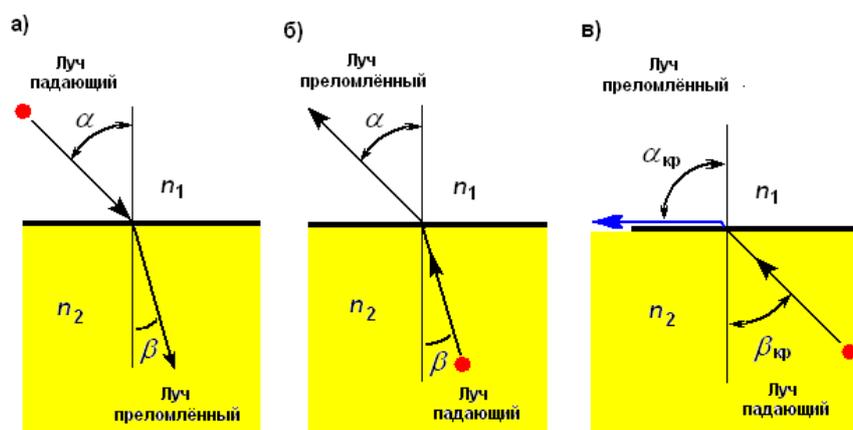


Рисунок 40 – Ограничение выхода излучения из полупроводника за счёт действия полного внутреннего отражения

$$\frac{\sin \beta}{\sin \alpha} = \frac{n_1}{n_2}.$$

При этом угол выхода преломленного луча, отсчитанный от нормали к поверхности, будет больше, чем угол падения. Если теперь увеличивать угол падения, то преломлённый луч будет опускаться вниз, приближаясь к границе раздела двух сред. Значит, существует такой угол $\beta = \beta_{кр}$, при котором преломлённый луч ляжет на границу раздела (рис.40, в). Это соответствует углу преломления $\alpha = \alpha_{кр} = 90^\circ$, т.е. преломлённый луч не выедет за пределы

оптически более плотной среды, останется в ней, где после многочисленных отражений от её границ поглотится полупроводником. Найти критический угол падения $\beta = \beta_{кр}$ можно, приравняв в законе Снелиуса угол $\alpha = \alpha_{кр} = 90^\circ$. Тогда получим

$$\beta_{кр} = \arcsin\left(\frac{n_1}{n_2}\right)$$

В типичной ситуации излучение выходит в окружающее пространство, где показатель преломления $n_1 = 1$. Типичное значение показателя преломления полупроводников $n_2 \approx 4$. Тогда значение критического угла в градусах составит

$$\beta_{кр} \approx \arcsin\left(\frac{1}{4}\right) \approx \arcsin\left(\frac{57^\circ}{4}\right) \approx 15^\circ.$$

Полученный результат указывает на то, что лишь небольшая часть сгенерированного полупроводником излучения, падающая на границу раздела под углом $\beta < \beta_{кр}$, выйдет во

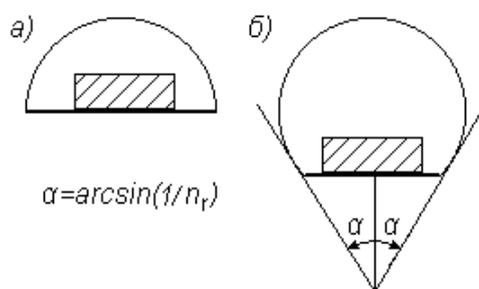


Рисунок 41 – Формы выходной линзы у светодиода

внешнюю среду, тогда как её основная вернётся в полупроводник, где поглотится.

С данным эффектом ограничения выхода излучения борются нанесением на излучающую поверхность специальных линз. Так, использование полусферы (рис. 41, а) или сферы Вейерштрасса (рис. 41, б) с апертурой $\alpha = \beta_{кр}$ позволяет в несколько раз увеличить внешнюю квантовую эффективность светодиодов.

5.8. Светоизлучающие структуры с высоким квантовым выходом

Светодиодные структуры, выполненные по диффузионной технологии изготовления *p-n* переходов на 3D-полупроводниках, являются источниками излучения с низким КПД. Причин для этого довольно много, но наиболее важные из них следующие:

- размытость области рекомбинации электронов и дырок, имеющая протяженность больше диффузионных длин обоих типов носителей заряда;
- излучательная рекомбинация носителей в области перехода на встречных курсах имеет низкую эффективность из-за малого времени их взаимодействия и большой протяженности областей с избытками носителей заряда;
- невозможность управления спектром излучения, задаваемого шириной запрещенной зоны материала *n*- и *p*- областей;
- высокая скорость выгорания зеркала, отражающего свет от тыловой части светодиода, из-за его нахождения в высокотемпературной области рекомбинации носителей заряда;
- высокие потери на перепоглощение сгенерированного излучения при его распространении к излучающей грани из-за одинакового значения ширины запрещенной зоны активной области и эмиттеров светодиода.

Устранить указанные недостатки и тем самым увеличить КПД диффузионных

светодиодов удаётся использованием светодиодных гетероструктур, в которых области излучательной рекомбинации выполнены в виде наноразмерных объектов: квантовых ям или квантовых точек.

5.9. Лазеры на полупроводниковых гетероструктурах

Лазеры на основе полупроводниковых гомопереходов, т.е. изготовленные из одного материала, обладают рядом недостатков, которые снижают выходную мощность таких лазеров, увеличивают пороговый ток до значений 10^4 А/см², ускоряют деградацию прибора. Перечислим эти недостатки:

- неопределенность размеров области генерации оптического излучения, которая определяется диффузионной длиной носителей заряда (обычно несколько микрон), встречающихся после снятия потенциального барьера в приборе, что приводит к формированию широкого луча лазера;
- неполная рекомбинация электронов и дырок в области перехода из-за размытости области рекомбинации и случайного характера рекомбинации в течение времени излучательной рекомбинации, что увеличивает пороговый ток лазера;
- возможность перепоглощения сгенерированных фотонов атомами *n*-и *p*- областей при распространении лазерного луча между зеркалами, что увеличивает пороговый ток;
- высокая скорость деградации зеркал из-за близкого их расположения к области генерации света.

Еще в 70-х годах прошлого столетия установлено, что твердотельные лазеры на двойных гетеропереходах (DHL) обладают на порядок меньшим пороговым током – около 1000 А/см². В этих гетеропереходах размеры области между барьерами достаточно велики – от 0.1 до 1 мкм = 1000 нм и потому квантование состояний электронов отсутствует. Поэтому лазеры на двойных гетеропереходах имеют на порядок меньшие пороговые токи (10^2 А/см²).

На рис.42 показана энергетическая диаграмма двойной гетероструктуры, которая является основным излучающим элементом в лазерах нового типа. Как следует из него,

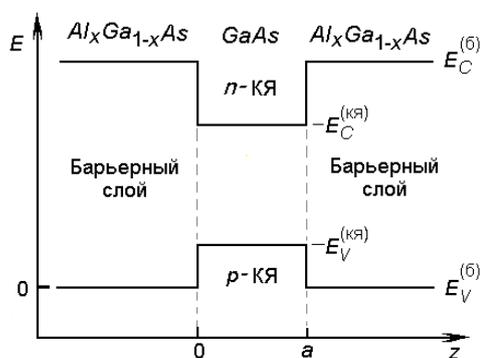


Рисунок 42 – Энергетическая диаграмма двойной гетероструктуры на основе арсенида галлия и алюмоната галлия

гетеропереходы позволяют формировать потенциальные ямы для электронов и дырок, повышая их концентрацию в активной области, но и увеличивать инверсию населенности электронов и дырок. Размеры активной области в лазерах на двойных гетероструктурах составляют порядка 0.1 мкм, что недостаточно для квантования энергии в квантовых ямах.

Следующее преимущество лазеров типа DHL перед лазерами на основе гомопереходов связано с тем, что у первых показатель преломления GaAs (область генерации излучения) примерно на 5% выше, чем у окружающего материала AlGaAs. Этой разницы оказывается достаточно, для уменьшения области оптической локализации излучения с 2.0 мкм в гомопереходах до 0.5 мкм.

Более того, преимущество лазеров на двойных гетеропереходах заключается в их

топологии: ввиду того, что излучение генерируется в GaAs (ширина запрещенной зоны 1.41 эВ), а окружающие области из AlGaAs (ширина запрещенной зоны 2.0 эВ), свет может распространяться без поглощения через окружающие области. Такой подход позволяет создавать лазеры с так называемой «полосковой геометрией», позволяющей легко присоединять их к оптическим устройствам, волоконным линиям связи. Это позволяет дополнительно снизить пороговый ток до 10^{-2} А за счет уменьшения длины резонатора Фабри-Перо, а значит уменьшения числа генерируемых мод в пользу низших мод.

Для повышения эффективности лазеров на двойных гетероструктурах в большинстве случаев используют полосковую геометрию, в которой размеры активной области в поперечном или горизонтальном измерении значительно уменьшены, что приводит к уменьшению порогового тока (рис.42). Благодаря такой форме активной области полосковые лазеры легко присоединяются к другим устройствам типа волокон, волноводов и т.д.

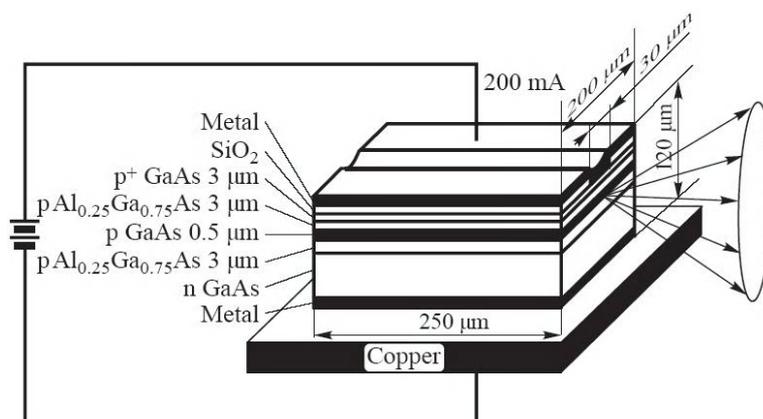


Рисунок 43 – Полосковый лазер на основе алюмината галлия и арсенида галлия

Одним из важных недостатков лазеров на основе двойных гетеропереходах является то, что удержание носителей заряда, генерация излучения и волновое распространение света происходят в одной области, что приводит к перепоглощению сгенерированных фотонов атомами вещества области распространения света. Более функциональные лазеры можно создать на основе квантовых ям, используя особенности зонного спектра (см.рис.42).

Гетероструктуры с квантовой ямой имеют более высокий коэффициент оптического усиления - g . Это достигается тем, что g пропорционален функции плотности состояний, которая в 3D полупроводниках зависит от энергии E как $\sqrt{(E - E_c)}$, а в 2D – эта функция скачкообразно возрастает при энергии, равной энергии первого состояния в квантовой яме E_1 . Различие в крутизне этих зависимостей приводит к различию их коэффициентов усиления оптического излучения при $E \geq E_1$ в пользу двумерных систем. Поэтому лазеры на основе гетероструктур с КЯ должны иметь более низкое значение порогового тока.

Следующее преимущество лазеров на основе гетероструктур с квантовыми ямами состоит в том, что они более надежны, чем лазеры на основе двойного гетероперехода, т.к. в них отсутствует деградация оптических зеркал, которая в случае двойного гетероперехода вызывается высокой скоростью рекомбинации носителей заряда в активной области лазера, сопровождаемой сильным нагревом зеркал. Более того, по сравнению с ними лазеры на основе квантовых ям обладают более высокой эффективностью и имеют меньшие потери.

ВОПРОСЫ ПО ГЛАВЕ 5

1. Какими параметрами описываются эмиссионные свойства твердых тел?
2. Какие типы люминесценции в зависимости от способа возбуждения существуют?
3. Что такое спонтанное излучение атома и как оно связано с другими переходами?
4. Что такое вынужденное излучение атома и как оно связано с поглощением?

5. В чем заключается принцип создания лазера?
6. Что такое линия излучения атома?
7. В каких случаях линия излучения описывается лоренцевой кривой?
8. Поясните принцип работы полупроводникового светодиода?
9. Какова основная причина малого вывода излучения из твердого тела?
10. Какова причина преимуществ светодиодных источников на основе гетероструктур?

6 ФЛУКТУАЦИОННЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

Под шумами понимают случайные изменения электрических сигналов на выходе электронного прибора, которые не связаны с входным сигналом, а вызваны случайными изменениями параметров и свойств самого прибора. Изучение шумов в приборах электронной техники представляет собой важную проблему в науке и технике, поскольку они определяют наименьшие уровни сигналов, которые могут быть обработаны электронными средствами, а также точность измерений величин электрических параметров. Чтобы определить эти уровни, необходимо уметь измерять основные параметры шумов, уметь минимизировать отношения «шум – сигнал» в устройствах приема и обработки и сигналов. Решение указанных вопросов невозможно без понимания природы шумов.

Физические механизмы появления шумов в полупроводниковых приборах во всех случаях связаны со случайным характером рождения-уничтожения свободных носителей заряда в активной области прибора, хаотическим движением заряженных частиц, разбросом значений его параметров, тепловыми колебаниями решетки.

6.1 Типы и физическая природа шумов

Природа шумов. Все шумы по своей физической природе имеют причиной случайный характер движения заряженных частиц в твердом теле, их случайное рождение или уничтожение, а также случайные моменты вхождения-выхода электронов и дырок. Наиболее важные для практических применений шумы в твердых телах таковы:

- тепловой шум;
- дробовой шум;
- генерационно-рекомбинационный шум;
- избыточный шум.

Тепловой шум. Тепловой шум проявляется как наличие на торцах полупроводникового образца, находящегося при температуре отличной от абсолютного нуля, знакопеременного напряжения. Его амплитуда, частота и начальная фаза меняются со временем случайным образом. Причиной появления теплового шума является броуновское движение свободных электронов (дырок) в полупроводнике и их столкновения с решеткой.

Для того чтобы понять физический механизм появления электрического напряжения на торцах полупроводника, характеризующего величину теплового шума, будем рассматривать полупроводник как совокупность двух взаимодействующих через соударения подсистем: ионную и электронную.

Тепловое движение заряженных частиц, например, k -того электрона, можно рассматривать как хаотический микроток величиной i_k (напомним, что ток – это направленное движение заряда). Протяженность этого микротока равна длине свободного пробега электрона. После его столкновения с решеткой направление движения и скорость электрона меняются случайным образом. Соударение, как это интерпретирует электродинамика, есть проявление сопротивления микроток, которое обозначим как R_k . Есть ток, есть сопротивление. Тогда согласно закону Ома можно говорить о наличии в объеме полупроводника на длине свободного пробега случайного источника микронапряжения величиной $U_k = i_k R_k$. Ясно, что от одного столкновения до следующего действует один микроисточник со своими параметрами напряжения, после столкновения – другой и т.д. Сколько столкновений совершает электрон при своем движении за время его наблюдения – столько микроисточников напряжения вводится в рассмотрение. Ясно, что число вводимых микроисточников огромно и потому проанализировать параметры отдельного микроисточника невозможно. У каждого микроисточника напряжения будет своя амплитуда, свой знак, своя ориентация в пространстве и т.д. Значит, можно рассматривать один источник: это некий источник переменного напряжения, полученный суммированием всех микроисточников от всех электронов в направлении регистрации. Поэтому если на

концах полупроводникового образца регистрировать переменное напряжение, которое будет представлять собой мгновенную сумму напряжений от всех микроисточников в направлении регистрации по длине кристалла, то окажется, что это суммарное напряжение изменяет случайным образом во времени свою величину, знак, фазу и частоту. Изменения параметров суммарного напряжения будут однозначно характеризовать тепловое движение электронов в решетке, а значит, температуру полупроводника.

Согласно теореме Найквиста, впервые описавшего этот тип шума, средняя амплитуда теплового шума резистора величиною R , измеренная с помощью прибора, имеющего полосу пропускания Δf , равна:

$$\sqrt{U_{\text{тепл}}^2} = \sqrt{4kTR \cdot \Delta f}. \quad (6.1)$$

На первый взгляд кажется удивительным, что в выражении (1) измеряемый уровень шума определяется параметрами измерительного прибора: полосой его пропускания. На самом деле ничего удивительного здесь нет. Тепловой шум описывается бесконечным числом синусоидальных источников переменного напряжения, амплитуда каждого из которых не зависит от частоты и равна

$$\sqrt{U_{\text{тепл}}^2(f)} = \sqrt{4kTR(f)}.$$

Поэтому чем шире полоса измерительного прибора, тем больше шумовых компонент анализируемого шума попадут в прибор и дадут вклад в измеряемую величину. Именно это обстоятельство отражено тем, что в выражении (6.1) фигурирует полоса пропускания измерительного прибора Δf : чем она больше, тем больше амплитуда измеряемого шума.

Если рассматривается тепловой шум двухполюсника, имеющего активное и реактивное компоненты импеданса, то тепловой шум будет создаваться только активной частью сопротивления двухполюсника $\text{Re}(Z)$, и тогда величина теплового шума в единичной полосе частот будет равна

$$\sqrt{U_{\text{тепл}}^2} = \sqrt{4kT \text{Re}(Z)}. \quad (6.2)$$

Если рассматриваемая цепь представляет собой последовательно соединенные активные сопротивления R_1, R_2, \dots, R_n , то полный шум такой цепи в единичной полосе частот будет иметь среднюю амплитуду, представляющую собой результат сложения квадратов дисперсий тепловых шумов от каждого сопротивления:

$$\begin{aligned} \sqrt{U_{\text{ш тепл}}^2} &= \sqrt{U_{\text{ш } 1}^2 + U_{\text{ш } 2}^2 + \dots + U_{\text{ш } n}^2} = \sqrt{4kTR_1 + 4kTR_2 + \dots + 4kTR_n} = \\ &= \sqrt{4kT(R_1 + R_2 + \dots + R_n)}. \end{aligned} \quad (6.3)$$

Этот же результат можно получить, если сложить все последовательно соединенные шумящие сопротивления, а только потом посчитать тепловой шум полного сопротивления.

Дробовой шум. Дробовой шум возникает в результате случайного характера изменения числа частиц, пролетающих область их регистрации. Типичный пример появления этого вида шума - вакуумный диод, с катода которого в каждый момент времени вылетают электроны. Вследствие разброса тепловой энергии электронов в катоде, приобретаемой в результате их столкновения с решеткой материала катода, в каждый момент времени вылетает разное число электронов. В среднем же за достаточно большой промежуток времени их число постоянно и определяется температурой катода, величиной работы выхода из его материала. Другой пример источник дробового шума - протекание постоянного тока через сопротивление.

Для описания дробового шума лучше вводить случайный источник тока, а не напряжения, как это сделано для теплового шума. Средняя амплитуда дробового шума, создаваемого током величиной I_0 и измеренная в полосе частот измерительного прибора Δf определяется следующим образом:

$$\sqrt{I_{\text{др}}^2} = \sqrt{2q \cdot I_0 \cdot \Delta f}, \quad (6.4)$$

где q - заряд электрона. Амплитуда напряжения полного шума в единичной полосе частот, снимаемого с активного сопротивления, по которому течет постоянный ток, определяется тепловым шумом и шумом за счет протекания по сопротивлению тока I_0 :

$$\sqrt{U_{\text{ш}}^2} = \sqrt{U_{\text{тепл}}^2 + R^2 I_0^2} = \sqrt{4kTR + 2qR^2 I_0}. \quad (6.5)$$

Генерационно-рекомбинационный шум. Этот вид шума характерен только для полупроводников, где зачастую является доминирующим. Природа его проста: за счет тепловой генерации и случайной рекомбинации в каждый момент времени в зоне проводимости и (или) валентной зоне появляются неравновесные носители заряда, число которых изменяется во времени случайным образом. Так, например, в первую секунду образовалось 100 электронов и дырок; во вторую секунду установилось 120 пар, в следующую – 80 и т.д. Таким образом, в среднем в полупроводнике в каждую секунду появляется 100 электронов и дырок. Однако около этого среднего значения возможны флуктуации числа частиц – в нашем случае около 20 пар. Именно это число неравновесных электронов и дырок приведет к случайному изменению проводимости полупроводника, что при пропускании через него постоянного тока величиной I_0 создаст на концах полупроводника шумовое напряжение

$$\delta U_{\text{гр}}(t) = I_0 \cdot \delta R(t). \quad (6.6)$$

6.2. Основные параметры и характеристики шумов

Флуктуирующие напряжения и токи являются случайными переменными, поэтому их изучение проводят, пользуясь аппаратом теории вероятностей. Основным способом статистического описания случайной переменной $x(t)$ является введение в рассмотрение ее среднего значения – $\overline{x(t)}$, которое обозначают как \overline{X} , и среднего квадрата случайной величины – $\overline{x(t)^2}$. Среднее значение случайной величины $x(t)$ находится как постоянная составляющая, вокруг которой происходят знакопеременные изменения случайной величины $x(t)$. Дисперсия случайного процесса $x(t)$ в общем случае вычисляется вычитанием из каждого значения случайного процесса его среднего значения, возведения полученного результата в квадрат и усреднения полученной величины по времени – $\overline{(x(t) - \overline{X})^2}$. Рисунок 43 иллюстрирует сказанное. Как можно видеть, дисперсия представляет собой квадрат средней амплитуды отклонения случайного процесса от своего среднего значения. Проще говоря, дисперсия – это усредненный по времени квадрат амплитуды случайного процесса.

Часто \overline{X} равно нулю, и тогда наиболее значимой величиной становится $\overline{x(t)^2}$. Таким образом, по аналогии с напряжением в электронном устройстве, где есть постоянное и переменное напряжение, например, на аноде лампы, и потому полное напряжение представляет собой их сумму, случайный процесс также представляют как сумму среднего значения шума и квадратного корня из его дисперсии.

Автокорреляционная функция. В стационарных случайных процессах важным параметром является среднее значение произведения двух значений случайного процесса, сдвинутых по времени на промежуток s : $\overline{X(t) \cdot X(t+s)}$. Оно называется **автокорреляционной функцией** и является мерой продолжительности влияния значения случайной переменной в данный момент времени на последующие ее значения, т.е. описы-

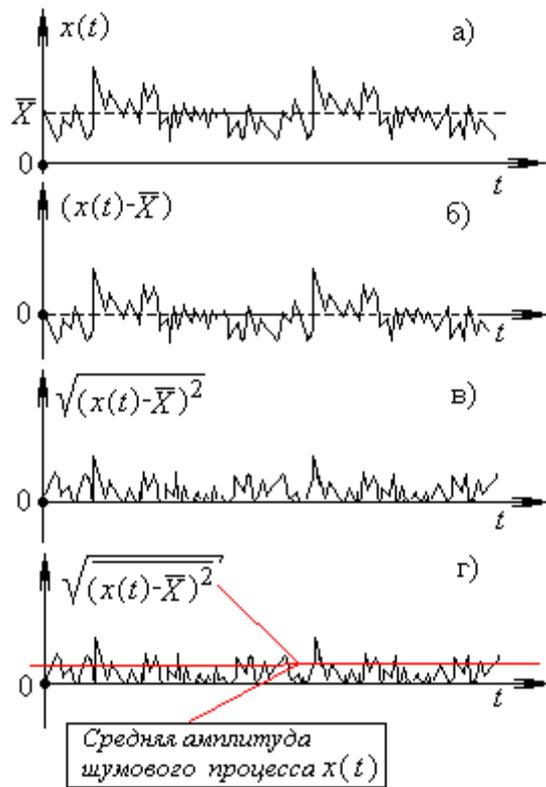


Рисунок.43 – Вычисление среднего значения и дисперсии случайной величины

вайт влияние настоящего случайного процесса на его будущее. Если $\overline{X(t) \cdot X(t+s)} = A\delta(s)$, т.е. является δ -функцией параметра запаздывания s , то шум называют белым, т.е. его дисперсия не зависит от частоты, оставаясь постоянной.

6.3. Полевые и фоновые свойства шумов в полупроводниках

Для выяснения зависимости величины регистрируемого шума полупроводникового прибора от значения постоянного напряжения на нем и наличии фоновой подсветки полупроводникового прибора мощностью P_{ϕ} рассмотрим входную цепь регистрирующего тракта установки (рис.2). Здесь обозначено: $R_{\phi p}$ - сопротивление полупроводникового прибора, например, фоторезистора, r - сопротивление нагрузки, V - напряжение смещения. В

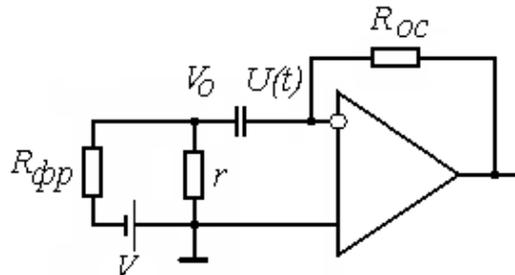


Рисунок 44 – Схема измерения шумовых свойств фоторезистора

результате протекания постоянного тока на сопротивлении нагрузки падает постоянное напряжение V_0 , которое зависит сопротивлений фоторезистора и нагрузки:

$$V_0 = r \cdot j = V \frac{r}{R_{\phi p} + r}.$$

Если сопротивление фоторезистора изменяется, например, за счет тепловых флуктуаций или фоновой засветки: $R_{\phi p} = R_{\phi p}^0 + \delta R(t, P_{\phi})$, то на входе усилителя появится напряжение

$$U(t, P_{\phi}) = V \frac{r}{R_{\phi p}^0 + \delta R + r} = V \left(\frac{r}{R_{\phi p}^0 + r} \right) \frac{1}{1 + \left(\delta R / (R_{\phi p}^0 + r) \right)},$$

переменная составляющая которого будет равна

$$U(t, P_{\phi}) \approx V \left(\frac{r}{R_{\phi p}^0 + r} \right) \frac{\delta R(t, P_{\phi})}{R_{\phi p}^0 + r}. \quad (6.7)$$

Отсюда найдем шумовую компоненту как корень из дисперсии флуктуации:

$$U_{ш}(V, P_{\phi}) = \sqrt{(U(t, P_{\phi}))^2} \approx V \left(\frac{r}{R_{\phi p}^0 + r} \right) \frac{\sqrt{(\delta R(t, P_{\phi}))^2}}{R_{\phi p}^0 + r}. \quad (6.8)$$

Выражение (8) показывает, что при приложении к фоторезистору постоянного напряжения величина шумового напряжения на входе устройства регистрации растет по линейному закону с увеличением V . Из выражения (8) также следует, что увеличение мощности фоновой подсветки, увеличивая число свободных носителей, в соответствии с законом больших чисел должно снижать уровень шума по закону $1/P_{\phi}^{1/2}$.

6.4. Частотные свойства шумов различной природы

Одним из эффективных методов анализа случайных величин является метод Фурье. Он основан на рассмотрении спектральных плотностей случайного процесса – $S_f(f)$. С помощью этой величины флуктуационную эдс $V(t)$ в узком интервале частот можно представить в виде источника синусоидальной с величиной шумового эдс $\sqrt{S_V(f)\Delta f}$. Достоинство такого подхода состоит в том, что с введением спектральной плотности для анализа шумов можно пользоваться теорией цепей переменного тока. Например, можно ввести шумовые генераторы тока и напряжения.

Важной теоремой спектрального анализа случайной величины является **теорема Винера-Хинчина**. Она позволяет вычислить спектральную плотность случайного процесса. Её формулировка такова.

Пусть $x(t)$ является стационарным случайным процессом. Разложим $x(t)$ в ряд Фурье в интервале времени $0 \leq t \leq T$, в течение которого он наблюдается:

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \cdot \exp(j\omega_n t),$$

где $n = 0, \pm 1, \pm 2, \dots$, а $\omega_n = \frac{2\pi n}{T}$ и $a_n = \frac{1}{T} \cdot \int_0^T X(t) \cdot \exp(-j\omega_n t) dt$. Тогда спектральная

плотность $S_x(f)$ случайного процесса $x(t)$ будет определяться выражением

$$S_x(f) = \lim_{T \rightarrow \infty} 2 \cdot T a_n a_n^*,$$

где знаком (*) отмечена комплексно-сопряженная величина. Другими словами, в соответствии с теоремой Винера-Хинчина спектральная плотность может быть найдена с помощью следующего выражения:

$$S_x(f) = 2 \cdot \int_{-\infty}^{\infty} \overline{X(t)} \cdot X(t + \tau) \cdot \cos \omega \tau \, d\tau .$$

При математическом рассмотрении случайного процесса обычно сначала находят автокорреляционную функцию, а потом определяют спектральную плотность $S_x(f)$. В экспериментальных исследованиях поступают наоборот: сначала измеряют $S_x(f)$, а затем находят автокорреляционную функцию. Если в выражении для спектральной плотности временной сдвиг τ равен нулю, то можно найти:

$$\overline{x^2(t)} = \int_0^{\infty} S_x(f) df , \quad (6.9)$$

Данное выражение указывает на способ экспериментального определения дисперсии случайной величины. Оказывается, что дисперсию можно определить путем интегрирования спектральной плотности случайного процесса $S_x(f)$ на конкретной частоте измерения во всем диапазоне частот. Для этого необходимо долгое время на конденсаторе, выступающем в роли усредняющего устройства, накапливать выпрямленное напряжение случайного процесса, выделенного селективным устройством

Как экспериментально измерить спектральную плотность шума? Оказывается совсем просто. Вспомним выражение для определения дисперсии: $\overline{x^2(t)} = \overline{(x - \bar{X})^2}$. Как следует из этого выражения, для вычисления дисперсии сначала надо из случайного процесса вычесть его среднее значение. На практике это делается установкой на входе измерителя дисперсии разделительного конденсатора – именно он выделяет из полного шумового напряжения $x(t)$ переменное напряжение, отделяя его от постоянного \bar{X} . Далее надо возвести в квадрат полученное знакопеременное шумовое напряжение, а затем извлечь корень, чтобы вернуться к общепринятым единицам измерения – вольтам. Операцию по переводу двуполярного напряжения в однополярное выполняет известный в радиотехнике квадратичный детектор, представляющий собой мост Уинстона, составленный из четырех диодов – диодный мостик. Такие детекторы стоят во всех выпрямителях, например, сетевого напряжения. Квадратичный детектор делает напряжение однополярным, что соответствует операциям возведения в квадрат и извлечения квадратного корня. Именно это требуется для вычисления дисперсии случайного процесса.

В результате этой операции шумовое напряжение станет однополярным и пульсирующим с удвоенной частотой изменения во времени. Такое напряжение содержит постоянную составляющую и переменную на удвоенной частоте. Поясним это на примере синусоидального напряжения: в квадрате оно имеет вид

$$(A \cos \omega t)^2 = A^2 \left(\frac{1 + \cos(2\omega \cdot t)}{2} \right) .$$

Его функция преобразования входного сигнала описывается выражением вида:

$$\sqrt{(A \cos \omega t)^2} = A \sqrt{\left(\frac{1 + \cos 2\omega \cdot t}{2} \right)} . \quad (6.10)$$

Необходимая для вычисления дисперсии операция усреднения во времени реализуется установкой на выходе диодного мостика конденсатора большой емкости. Что получим на выходе? Усреднение правой части выражения (10) с учетом извлечения корня даст константу $1/\sqrt{2}$, которая, будучи умноженной на амплитуду входного сигнала A , определит

эффективную амплитуду выпрямленного напряжения, а в применении к измерению шумов определит эффективную величину дисперсии:

$$U_{\text{ш}} = \sqrt{U_{\text{ш}}^2} = \frac{A}{\sqrt{2}} \sqrt{1 + \cos(2\omega \cdot t)} = \frac{A}{\sqrt{2}}.$$

Что получилось в итоге? Оказалось, что корень из дисперсии случайного процесса можно экспериментально измерить, если подать исследуемый шум через конденсатор на обычный квадратичный детектор, нагруженный на фильтр низких частот (ФНЧ). Такие устройства являются стандартными в любом измерителе напряжения – вольтметре и даже в выпрямителе напряжения. Единственное отличие, которое должно быть учтено в вольтметрах для измерения шума так это то, что исследуемые шумовые величины обычно имеют очень малую величину: до единиц нановольт и меньше. Значит, для изучения шума необходимы нановольтметры с квадратичными детекторами и ФНЧ на выходе.

Для изучения природы исследуемого шума весьма полезны частотные свойства шума: зависимость шумового напряжения (дисперсии случайного процесса) от частоты измерения. По виду этой зависимости можно определить тип доминирующего шума: тепловой и дробовой шум имеют независимую от частоты измерений дисперсию, генерационно-рекомбинационный шум имеет спектральную зависимость в виде «полочки» (постоянное значение на низких частотах) и падающий участок на высоких частотах. Избыточный шум имеет падающую степенную зависимость дисперсии от частоты и т.д.

Для измерения частотного состава исследуемого шумового напряжения разработаны и широко применяются так называемые селективные нановольтметры. Они представляют собой вольтметры, способные регистрировать напряжения до единиц нановольт - 10^{-9} В и имеющие очень узкую и перестраиваемую оператором полосу пропускания. Последнее достигается установкой на входе нановольтметра перестраиваемого полосового фильтра. С помощью такого прибора удастся проанализировать зависимость амплитуды шума, измеренной в полосе измерительного прибора равной 1 Гц, от частоты его измерения. Такая зависимость называется **спектральной плотностью шума**. С точки зрения физической интерпретации спектральной плотности шума (и не только шума, но и детерминированного сигнала) поведение шума во всей полосе исследованных частот можно объяснить, предположив, что он создается множеством источников синусоидального переменного напряжения. Амплитуда каждого источника, действующего на заданной частоте, равна спектральной плотности шума.

На практике обычно экспериментально наблюдаемое шумовое напряжение формируется в результате действия нескольких механизмов, например, теплового $U_{\text{тепл}}(t)$, дробового $U_{\text{дроб}}(t)$ и избыточного $U_{\text{изб}}(t)$. Поэтому часто необходимо знать, как разложить наблюдаемый шум на составляющие. Специалисты, знакомые с работой радиоэлектронных систем, знают, что при одновременном действии на общую нагрузку нескольких переменных напряжений от независимых источников максимальное напряжение в цепи равно сумме напряжений от отдельных источников. Аналогичная картина имеет место, если эти напряжения – шумы от различных механизмов. Значит, в этих ситуациях результирующее напряжение, например, в зависимости от частоты измерения f , есть сумма дисперсий от каждого механизма:

$$U_{\text{ш}}(f) = \sqrt{U_{\text{сум}}^2(f)} = \sqrt{U_{\text{тепл}}^2(f) + U_{\text{дроб}}^2(f) + U_{\text{изб}}^2(f)}. \quad (6.11)$$

ВОПРОСЫ ПО ГЛАВЕ 6

1. Какие типы шумов возможны в приборах оптоэлектроники?
2. Какова природа теплового шума и как его уменьшить?
3. Какова природа дробового шума и как его уменьшить?
4. Какова природа генерационно-рекомбинационного шума?

5. Какова природа избыточного шума и как его уменьшить?
6. Что описывает в применении к шуму среднее значение случайной величины?
7. Что описывает в применении к шуму дисперсия случайной величины?
8. Как рассчитать среднюю амплитуду шума и как её измерить экспериментально?
9. Как экспериментально измерить полевые свойства полупроводникового прибора?
10. Как измерить частотные свойства шума у полупроводникового прибора?

7 ЭЛЕКТРООПТИЧЕСКИЕ ЭФФЕКТЫ ДЛЯ УПРАВЛЕНИЯ ИЗЛУЧЕНИЕМ

7.1 Общие сведения об электрооптических эффектах

Ранее показано, что электрическое поле может существенно изменить оптические свойства полупроводниковых гетероструктур. Такие эффекты получили название электрооптических эффектов. Электрическое поле может влиять как на коэффициент преломления среды (*эффект электропреломления*), так и на её коэффициент поглощения (*эффект электропоглощения*). Если энергия фотона меньше ширины запрещенной зоны ($\hbar\omega < \Delta E_g$), то основным эффектом будет эффект электропреломления т.к. поглощение в этом случае мало. При высоких значениях энергии фотона основным эффектом может стать электропоглощение.

Электрооптические эффекты широко используются во многих оптоэлектронных устройствах для управления светом с помощью электрического поля. Так, перпендикулярное к плоскости барьера электрическое поле расширяет спектр поглощения квантовой точки («красный сдвиг») и повышает коэффициент преломления материала ниже края поглощения. Эти изменения становятся существенными и сильно зависят от энергии падающих фотонов вблизи границ запрещенной зоны.

Простейшим вариантом регулирования интенсивности света является использование электропоглощения. Из физики трехмерных полупроводников известно, что приложенное к полупроводнику электрическое поле сдвигает край собственного поглощения в сторону длинных волн и создает колебания при энергиях выше ширины запрещенной зоны (эффект Франца - Келдыша). Таким образом, для фотонов с энергией чуть ниже края собственного поглощения в отсутствие электрического поля полупроводник прозрачен. Наложение электрического поля приводит к красному сдвигу спектра поглощения, и полупроводник для этого же света становится непрозрачным из-за возросшего поглощения. В объемных полупроводниках этот эффект мал и практического применения не нашел, но в квантовых точках может стать основой для создания модуляторов оптического излучения. Когда вектор электрического поля параллелен плоскости квантовой ямы, ситуация похожа на случай объемного материала. Если же оно перпендикулярно слоям квантовых ям и барьеров, то возникающий эффект подобен эффекту Штарка в атомах (сдвиг уровней электронов пропорционально полю за счет кулоновского взаимодействия поля и заряда электрона). По этой причине этот эффект называют «квантово-размерным эффектом Штарка». Смещение электронов и дырок в квантовой яме внешним полем в противоположных направлениях приводит к изменению гамильтониана в уравнении Шредингера. Это приводит к изменению энергий стационарных состояний электронов в квантовой яме, что обнаруживается как «красный» сдвиг спектра поглощения. В экспериментах установлено, что в квантовой яме GaAs - AlGaAs толщиной 10 нм приложение электрического поля сдвигает энергетические уровни электронов на 0.015 эВ, а дырок на 0.002 эВ. В ямах толщиной 20 нм поле величиной 80 кВ/см снижает энергию края поглощения с 152 эВ до 142 эВ.

Таким образом, квантовые ямы перспективны для модуляции оптического излучения с эффективностью, превосходящей все используемые в настоящее время материалы (ниобат лития, объемные полупроводники и т.д.).

7.2 Эффекта Франца – Келдыша в оптоэлектронике

Эффект Франка - Келдыша представляет собой изменение ширины запрещенной зоны полупроводника при приложении к нему достаточно большого электрического поля. Физическое его проявление заключается в следующем: если на полупроводник направить излучение с энергией кванта меньше ширины запрещенной зоны полупроводника, то в отсутствие электрического поля падающее излучение не поглотится и на выходе из кристалла может зарегистрировано. Однако если приложить электрическое поле, то в

результате наклона энергетической диаграммы полупроводника эффективная ширина запрещённой зоны уменьшается, и ранее непоглощавшееся излучение будет поглощаться. Этот эксперимент указывает на то, что в больших электрических полях происходит расширение области межзонного поглощения полупроводника в сторону бо́льших длин волн. Тогда при регистрации излучения на просвет полупроводникового кристалла в отсутствие электрического поля падающее длинноволновое излучение будет проходить через модулятор без поглощения, а при действии электрического поля будет поглощаться в полупроводнике. В области, где энергия фотона больше ширины запрещенной зоны, возникают осцилляции коэффициента поглощения и отражения, как функции приложенного электрического поля и частоты, что позволяет оптимизировать параметры кристалла и величины прикладываемого поля для достижения максимального эффекта модуляции. По своей величине эффект Франца Келдыша невелик и в реальных ситуациях не превышает единицы процентов.

На рисунке 45 приведены энергетические диаграммы 3D полупроводника, на которых указана ширина запрещённой зоны, регистрируемая различными переходами электронов между зонами энергий в отсутствие электрического поля F и его наличии. Когда поле

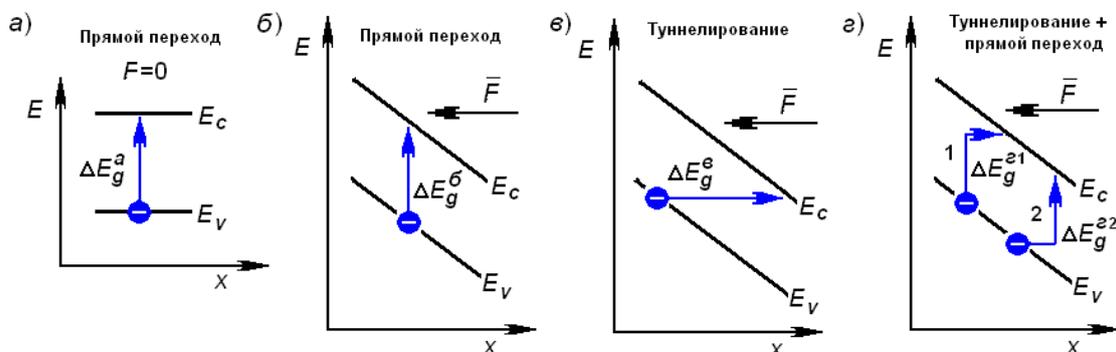


Рисунок 45 - Варианты переходов электронов при поглощении фотонов в 3D полупроводниках в отсутствие внешнего электрического поля – а) и его наличие – б), в), г).

отсутствует, падающее оптическое излучение поглощается за счёт зон-зонных переходов с энергией ΔE_g^a (рис.45, а). Приложение электрического поля приводит к тому, что зон-зонные переходы могут происходить тремя путями с различной энергией поглощаемых фотонов (рис.45, б – рис. 45, г). При малых полях, когда вызываемый им наклон диаграммы невелик, поглощается излучение с энергией $\Delta E_g^b > \Delta E_g^a$. Можно показать, что

$$\Delta E_g^b = \Delta E_g^a \sqrt{1 + q^2 F^2},$$

где обозначено: q - заряд электрона; F - электрическое поле. В слабых полях это выражение можно разложить в ряд Тейлора и получить выражение вида

$$\Delta E_g^b \approx \Delta E_g^a \left(1 + q^2 F^2 / 2\right).$$

Из данного выражения следует, что увеличение энергии поглощаемых фотонов по сравнению со случаем отсутствия поля составит малую величину, пропорциональную квадрату величины поля. Данный эффект по своему влиянию на длину волны поглощения. противоположен эффекту Франца – Келдыша

Если поле достаточно большой величины, то наряду с перехода по схеме рис.43, б появятся туннельные переходы между зонами (рис.43, в), которые не требуют изменения их энергии. Вероятность таких переходов W_B сложным образом зависит от величины поля и

исходной ширины зоны - ΔE_g^a . При этом вероятность переходов по схеме рис.43, б W_6 значительно меньше W_B .

Однако в сильных полях возможны переходы смешанного типа (рис.43, г), когда вблизи зон доминируют туннельные переходы, а в середине запрещённой зоны они заменяются вертикальными переходами с рис.43, б. Заметим, что в переходах смешанного типа наиболее вероятными являются переходы, сочетающие в себе переходы вида «1» и вида «2»: вблизи E_V вплоть до энергий $E_V^{вир}$ реализуется переход вида «1» с последующим его преобразованием в вертикальный переход. Вертикальный переход имеет место до уровня энергии $E_C^{вир}$, после которой опять доминирующим становится туннельный переход. На рис.44, а приведена энергетическая диаграмма 3D полупроводника, какой она должна быть в отсутствие электрического поля, чтобы объяснить происходящую смену механизма межзонного перехода электрона. На данном рисунке вблизи зон разрешённых значений энергии введены зоны виртуальных состояний, через которые осуществляется туннельный переход. Границы этих зон имеют энергии $E_C^{вир}$ и $E_V^{вир}$, которые определяются из условия равенства вероятностей W_B и W_6 ввиду уменьшения концентраций виртуальных состояний $N_{вир}(E)$ по мере удаления от края зоны (рис.46, а).

Введение виртуальных состояний позволяет объяснить поглощение фотонов с длиной волны λ_M больше красной границы внутреннего фотоэффекта λ_a (рис.46, б), коэффициент поглощения которых $\alpha_{вир}(\lambda, F)$ уменьшается с увеличением длины волны фотонов. Спектр поглощения фотонов в отсутствие электрического поля показан ступенчатой кривой $\alpha_a(\lambda)$.

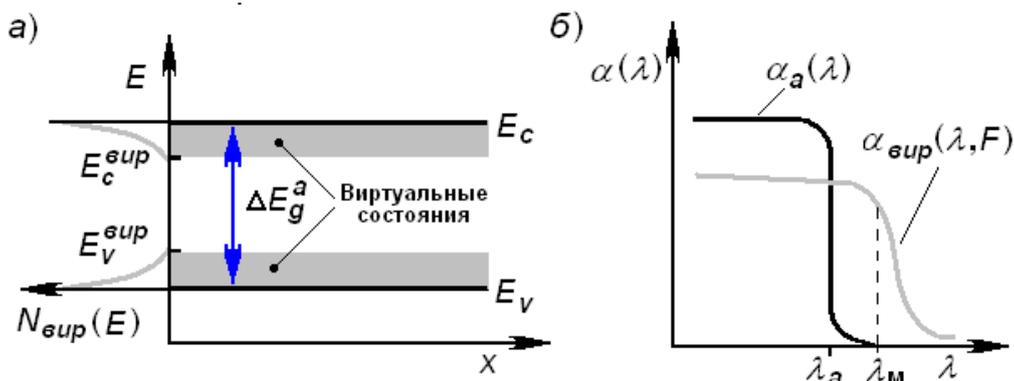


Рисунок 46 - Виртуальные состояния в эффекте Франца - Келдыша: а) положение этих состояний в запрещенной зоне полупроводника, б) – спектр поглощения полупроводника без виртуальных состояний и с ними

Таким образом, для объяснения эффекта Франца – Келдыша в рамках классических представлений о процессах в объемных полупроводниках, помещённых в сильное электрическое поле, необходимо предположить наличие (или возникновение) виртуальных электронных состояний, находящихся вблизи краёв зон разрешённых энергий, концентрация которых уменьшается с ростом их энергии.

Как использовать данный эффект для модуляции и переключения оптического излучения? В электропоглощающих модуляторах и переключателях, к которым относят модуляторы на эффекте Франца – Келдыша, модулируемый пучок света с длиной волны $\lambda_M > \lambda_a$ направляют на кристалл. В отсутствие электрического поля свет проходит через кристалл, не поглощаясь. При приложении электрического поля, изменяющегося во времени по закону модуляции, происходит расширение полосы поглощения за счёт эффекта Франца – Келдыша на столько, что длина волны модулируемого излучения попадает в хвост спектра поглощения. Часть интенсивности этого излучения поглотится в полупроводнике и на

выходе из кристалла интенсивность будет меньше падающей. Меняя величину электрического поля, будет меняться коэффициент поглощения в хвосте спектра по закону модуляции, заложенному в величине поля $F(t)$.

Модуляторы на основе эффекта Франца – Келдыша способны работать с высокой скоростью изменения коэффициента поглощения и невысоких напряжениях по сравнению с модуляторами, основанными на эффекте двулучепреломления в анизотропных кристаллах.

7.3 Двулучепреломление в анизотропных 3D кристаллах

Оптоэлектроника имеет дело с двумя объектами, используемыми для передачи информации в том или ином виде, их взаимодействием и взаимным превращением: с электронами и фотонами. Однако управление движением этих частиц совершается совершенно разными способами. Чтобы изменить траекторию движения электронов, достаточно создать требуемую конфигурацию электрического и магнитного полей, в которых они движутся. Изменить траекторию движения фотонов можно только за счет отражения света от границы раздела двух сред с разными оптическими свойствами либо использовать дифракцию на специально создаваемых неоднородностях этих свойств. Ясно, что такое управление движением фотонов требует применения.

Далее рассматриваются оптические свойства кристаллов в твёрдой и жидкой фазах, которые обладают анизотропией. Следствием этого такие вещества проявляют эффекты преобразования оптического излучения, которые могут быть использованы для управления оптическим излучением.

Рассмотрим распространение плоской электромагнитной волны через оптически прозрачный кристалл, обладающий разным значением диэлектрической проницаемости по направлениям: x , y , z . Пусть вдоль осей x , y диэлектрическая проницаемость кристалла одинакова, а вдоль оси z она имеет другое значение: $\epsilon_x = \epsilon_y \neq \epsilon_z$. Направим плоскую волну так, чтобы вектор поляризации был параллелен оси $z = x_3$, а единичный вектор нормали к волновому фронту направлен по оси $y = x_2$.

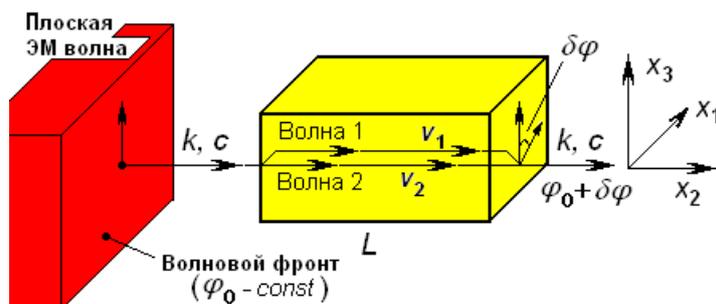


Рисунок 47 – Эффект поворота фазы плоской электромагнитной волны при её распространении в анизотропном кристалле

Рассматриваемая ситуация и результаты рассмотрения представлены на рис. 47. В анизотропных кристаллах в общем случае имеет место эффект двулучепреломления: плоская монохроматическая электромагнитная волна с определенной линейной поляризацией (направлением колебания вектора электрической индукции), вошедшая в анизотропный кристалл и прошедшая в нем некоторое расстояние не менее несколько длин волн, меняет своё положение так, что в вектор электрической индукции поворачивается вокруг

направления распространения на угол $\delta\varphi$, пропорциональный величине анизотропии диэлектрической поляризации и пройденному расстоянию. Важно, что обсуждаемый набег фазы на порядки превышает изменение фазы световой волны, получаемый только за счёт пройденного расстояния без учёта анизотропии.

Появление фазового набега можно объяснить следующим образом. Вошедшая в кристалл волна может быть представлена как сумма двух линейно поляризованных волн с разными скоростями распространения фазового фронта: v_1 , v_2 и взаимно перпендикулярными векторами \vec{D} . Имея разные фазовые скорости, расщепленные волны распространяются через кристалл длиной L и на выходе из него их вектора электрической индукции будут иметь разные фазы: $\phi_1 = \omega L n_1 / c + \omega t$ и $\phi_2 = \omega L n_2 / c + \omega t$. Выходя из кристалла, волны должны объединиться сложением взаимно перпендикулярных векторов электрической индукции со своими фазами. Если бы их фазы были одинаковы ($n_1 = n_2$), то $\Delta\varphi = \phi_1 - \phi_2 = \omega L \cdot (n_1 - n_2) / c = 0$. При разных же значениях коэффициента преломления анизотропного кристалла возникает разность фаз, что приводит к повороту суммарной световой волны на выходе из кристалла на угол $\Delta\varphi$. Поворот плоскости поляризации электромагнитной волны при прохождении через оптически активный кристалл можно обнаружить экспериментально, сравнивая с помощью поляризатора интенсивности электромагнитной волны на входе в кристалл и на его выходе. При отсутствии набега фазы или набеге, кратном 180° , интенсивности будут равны, а при наличии двулучепреломления интенсивность на выходе будет меньше входной. По мере увеличения разности фаз до 90° она будет уменьшаться по закону $\cos^2(\Delta\varphi)$ как меняется проекция вектора электрического поля прошедшей волны на направление анизотропии $z = x_3$. Если на выходе из кристалла поставить поляризатор, пропускающий ту же поляризацию, что и у входной волны, то увидим, что интенсивность выходного излучения окажется меньше входной на величину, определяемую набегом фазы.

Управлять набегом фазы можно изменением величины электрического поля, прикладываемого вдоль $z = x_3$, а выходное излучение пропускать через поляризатор, пропускающим излучение с поляризацией, перпендикулярной поляризации падающего излучения. Реально достижимо изменение интенсивности до (10-20) % при длине кристалла в один сантиметр и амплитуде модулирующего напряжения до 1 кВ.

ВОПРОСЫ ПО ГЛАВЕ 7

1. В чём различие между эффектами электропоглощения и эффектами электропреломления?
2. В чём суть эффекта Франца – Келдыша и каков механизм его действия на оптическое излучение?
3. Каковы причины появления виртуальных состояний в эффекте Франца – Келдыша?
4. Почему в эффекте Франца-Келдыша вероятность туннелирования электрона оказывается выше вероятности его прямого перехода?
5. Как использовать эффект Франца – Келдыша для модуляции оптического излучения?
6. Как действует эффект двулучепреломления в 3D полупроводниках?
7. Как должен быть ориентирована поляризация падающего излучения для достижения эффекта двулучепреломления?
8. Как использовать эффект двулучепреломления для модуляции оптического излучения?
9. Можно ли на эффекте двулучепреломления получить набег фазы излучения близкий к 90° ?
9. Какова эффективность модуляции излучения с помощью эффекта двулучепреломления?
10. Какова эффективность модуляции излучения с помощью эффекта Франца-Келдыша?

8 ЖИДКИЕ КРИСТАЛЛЫ В ОПТОЭЛЕКТРОНИКЕ

8.1. Общие сведения о жидких кристаллах

В 1888 г. австрийский ученый Рейнитцер получил новое органическое вещество – холестерилбензоат. Определяя температуру его плавления, он заметил, что вещество плавится в две стадии: вначале образуется мутный расплав, который при дальнейшем повышении температуры превращается в прозрачную изотропную жидкость. Необычными в этом процессе были появление и свойства мутного раствора. Обладая всеми свойствами жидкости, мутный расплав проявлял оптические свойства одноосного кристалла: свет, проходя через него, испытывал двойное лучепреломление. Рейнитцер решил, что перед ним смесь двух веществ, плавящихся при разных температурах, причем в расплаве одного из веществ находятся кристаллиты другого. Он не смог разделить эту смесь и послал свой препарат немецкому физическому Леману, занимавшемуся определением оптических свойств кристаллов. Леман выяснил, что присланный ему холестерилбензоат представляет собой чистое вещество, а не смесь двух, которое при нагревании переходит в ранее неизвестное агрегатное состояние: между кристаллической фазой и изотропной жидкостью образовывался мутный расплав с анизотропными оптическими свойствами кристалла. При дальнейшем повышении температуры эта фаза переходила в изотропную жидкость. На рис. 48 показано изменение условного параметра – «параметра структурного порядка» η кристалла, проявляющего указанные температурные изменения своего строения.

Вскоре выяснилось, что холестерилбензоат не единственное вещество, обладающее такими необычными свойствами. Леман нашел подобные фазы при плавлении еще целого ряда органических веществ. Всем им он дал название **жидкие кристаллы**, а оптически анизотропную фазу вскоре назвали **мезофазой** (от греческого слова «мезос» –

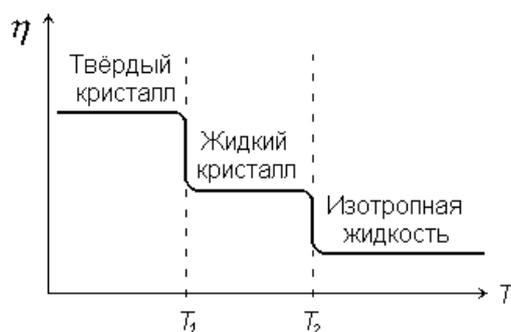


Рисунок 48 – Зависимость агрегатного состояния холестерилбензоата от температуры

промежуточный). Леман обнаружил, что у некоторых органических веществ аналогичные фазы возникают не только при нагревании, но и при растворении в спиртосодержащем растворе. Такие жидкие кристаллы называют **лиотропными** (от греческого слова «лио» – растворяю), в отличие от **тиреотропных**, которые образуются при нагреве. В настоящее время известно около 500 органических веществ, обладающих жидкокристаллической мезофазой.

8.2. Основные свойства жидких кристаллов

После открытия жидких кристаллов их свойствами заинтересовались химики и физики, и к 30-м годам XX века были выяснены все основные типы жидких кристаллов, особенности строения и их основные физические свойства, наиболее необычными из которых оказались следующие.

Двулучепреломление света. Явление двулучепреломления – это типично кристаллический эффект, состоящий в том, что скорость света в кристалле зависит от ориентации плоскости поляризации света. Она достигает максимального и минимального

значений для двух взаимно ортогональных ориентаций плоскости поляризации. Ориентации поляризации, соответствующие экстремальным значениям скорости света, определяются анизотропией свойств кристалла и однозначно задаются ориентацией кристаллографических осей относительно направления распространения света.

Линейно поляризованный свет, распространяясь в оптически активной среде, изменяет ориентацию плоскости поляризации. Угол поворота плоскости поляризации прямо пропорционален пути L , пройденному светом, т.е. выражается формулой $\varphi = \varphi_0 L$, где величина φ_0 определяет угол поворота на единичном отрезке пути и называется **удельной вращательной способностью**. Величина вращательной способности φ_0 для жидких кристаллов в сотни и тысячи раз превосходит эту величину для наиболее оптически активных кристаллов, например, кварца. Более того, угол поворота плоскости поляризации в жидких кристаллах зависит от длины волны света. В твердых телах, как, впрочем, и в обычных жидкостях, удельная вращательная способность φ_0 имеет вполне определенный знак, не зависящий от длины волны света. Это означает, что вращение плоскости поляризации света в них происходит в определенном направлении. В жидких кристаллах для коротких длин волн величина φ_0 , например, могла быть положительной, а для более длинных – отрицательной. Но в любом случае направление вращения плоскости поляризации света при прохождении им жидкокристаллической фазы изменялось в зависимости от длины волны. Такое поведение угла вращения плоскости поляризации не укладывалось в рамки существовавших представлений об оптической активности веществ.

Чувствительность к воздействиям. Удивительными оказались и другие свойства жидких кристаллов, такие как исключительно высокая чувствительность структурного состояния, оптических свойств жидкого кристалла к магнитным, электрическим полям, примесям, ионизирующему воздействию, температуре и другим воздействиям, что и определило практическое применение жидких кристаллов.

В первые годы жидкие кристаллы казались бесперспективными для практического использования. Однако в 1963 г. был запатентован метод регистрации тепловых излучений с помощью тонкой пленки жидкого кристалла, изменяющей свой цвет при нагревании. Он позволял решить одну из важных задач квантовой электроники – отображение тепловых потоков. С этого момента начинается бурное развитие исследований физических свойств жидких кристаллов и их применение в оптике, электронике, приборостроении и т.д., которое продолжается по настоящее время и только усиливается.

Итак, жидкие кристаллы – это специфическое агрегатное состояние органического вещества, в котором оно одновременно проявляет свойства жидкости и кристалла. При плавлении кристаллы этих веществ образуют жидкокристаллическую фазу, отличающуюся от обычных жидкостей: она существует в узком температурном диапазоне, вне которого имеем кристалл или обычную жидкость. Обязательным свойством жидких кристаллов является наличие порядка в пространственной ориентации молекул. Все известные в настоящее время жидкие кристаллы составлены из одного сорта молекул, которые, как правило, имеют вытянутую стержнеобразную форму либо вид вытянутой спирали. В последнем случае направление закручивания спирали может быть либо в одну сторону, либо в другую. Такие жидкие кристаллы называют **хиральными**.

8.3. Типы и структура жидких кристаллов

В случае стержнеобразных молекул в зависимости от характера упорядочения длинных осей молекул в слое жидкие кристаллы разделяют на **нематики**, **смектики** и **холестерики**. Рассмотрим их внутреннюю структуру и пространственное упорядочение молекул. Для наглядности будем представлять молекулы жидких кристаллов в виде стержней, т.к. их длина во много раз больше толщины.

Нематики. Структура нематика представляет собой «жидкость одинаково ориентированных стержней» (рис. 49, а, б). В твердом состоянии молекулы кристалла строго ориентированы в пространстве (рис. 49, а). В жидкокристаллическом состоянии молекулы могут перемещаться под действием температуры в любую сторону. Относительный порядок существует только в ориентации длинных осей молекул. Эти стержни движутся хаотически в жидкости, но их преимущественная ориентация не изменяется. Под действием температуры они совершают ориентационные (угловые) колебания, амплитуда которых возрастает при приближении к температуре фазового перехода (рис. 49, б). После фазового перехода ориентационное упорядочение исчезает, и ориентация молекул становится полностью хаотической (рис. 49, в). В отношении оптических свойств нематики подобны оптическим одноосным кристаллам.

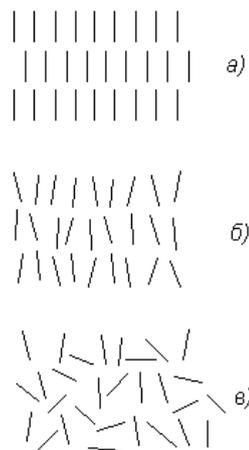


Рисунок 49 – Строение нематического жидкого кристалла с ростом температуры

Для характеристики ориентационного порядка в жидком кристалле вводится вектор единичной длины \bar{n} , называемый **директором**, направление которого совпадает с направлением усредненной ориентации длинных осей молекул. По существу, он представляет собой единичный вектор усредненного направления длинных осей молекул.

Кроме директора для оценки структурного состояния жидкого кристалла, вводится еще один параметр – скалярный **параметр порядка** Q , который характеризует угловое отклонение длинных осей молекул от директjра:

$$Q = \frac{3}{2} \left(\overline{\cos^2 \theta} - \frac{1}{3} \right),$$

где θ – угол между направлением директора и мгновенным направлением оси молекулы, а $\overline{\cos^2 \theta}$ – среднее значение по времени. Значение $Q = 1$ соответствует полной ориентации всех молекул нематика в одном направлении, а $Q = 0$ – полному беспорядку в ориентации молекул. По физическому смыслу параметр порядка есть дисперсия случайной переменной – ориентации длинных осей молекул жидкого кристалла.

До сих пор речь шла об однодоменном образце нематика, в котором ориентация директора одинакова во всех его точках. В реальности молекулы жидкого кристалла могут изменять свою ориентацию по поверхности слоя жидкого кристалла, а также с его толщиной, образуя области одинаковой ориентации длинных осей молекул жидкого кристалла – домены. Причины такого деления нематика заключены в случайном характере внешних воздействий, способных задавать ориентацию молекул. Преимущественные ориентации длинных осей молекул в различных доменах никак не связаны между собой. Поэтому ориентация директоров в разных доменах хаотична, и, следовательно, оптические индикатрисы в доменах также разориентированы – в каждом домене ориентация

индикатрисы своя. Значит и показатель преломления для света, падающего на многодоменный жидкий кристалл, в каждом домене свой. На границе доменов имеет место скачок показателя преломления, который в поляризованном свете в скрещенных поляризаторах видится как светлая нить произвольной формы на темном фоне (нить по-гречески «*нема*»), отсюда и название жидкого кристалла такого строения), которая указывает местоположение отдельных доменов в слое нематика и их размеры. Поэтому, чтобы сделать его одномерным, говорят, «однодоменным», необходимы специальные меры. К ним относятся любые способы, задающие ориентацию длинных осей молекул: механическая полировка внутренних поверхностей стекол, между которыми помещается жидкий кристалл; задание ориентации с помощью внешних электрических или магнитных полей и т.д.

Смектики. В них степень упорядочения молекул выше, чем в нематиках: помимо ориентационной упорядоченности молекул, существует частичное упорядочение центров их тяжести. Таким образом, смектики организованы в слои, расстояния между которыми фиксированы. Что же касается расположения центров тяжести молекул в пределах одного слоя, то оно хаотично, как в нематике. Поэтому для описания состояния смектика также используются понятия директора \bar{n} и параметра порядка Q .

Общим для всех смектиков является достаточно сильное взаимодействие молекул в пределах выделенного слоя и слабое взаимодействие между слоями. Поэтому слои легко скользят друг относительно друга и на ощупь все смектики мылоподобны (мыло по-гречески «*смегма*»). В отношении оптических свойств смектики ведут себя как двухосные кристаллы.

В зависимости от расположения молекул в слое, а также общего поведения директора по многим слоям смектики делятся на следующие типы: *A*, *B*, *C*, *D*, *F*, *G* и т.д. Так, в *A*-смектике оси молекул во всех его слоях перпендикулярны поверхности слоя; в *B*-смектике оси молекул находятся под небольшим углом к поверхности, одинаковым для всех слоев. В каждом слое *D*-смектика отсутствует какое-либо упорядочение осей молекул – молекулы ориентированы хаотически.

Смектики обладают двулучепреломлением, но в исходном состоянии представляют собой совокупность хаотически ориентированных доменов.

Если растопить смектический кристалл, то сначала он дает жидкий смектик. При дальнейшем повышении температуры смектик переходит в нематик, а уже тот – в изотропную жидкость. Это и понятно: сравнительно слабое взаимодействие между слоями с повышением температуры разрушается раньше, чем взаимодействие молекул внутри каждого слоя – образуется нематик. У отдельных смектиков нематическая фаза отсутствует.

Холестерики. Они устроены более сложно, чем нематики и смектики. Локально холестерический кристалл устроен так же, как и нематик: в малом объеме упорядочение молекул холестерика можно охарактеризовать директором и параметром порядка Q . Отличие холестерика от нематика проявляется в больших, по сравнению с молекулярными размерами,

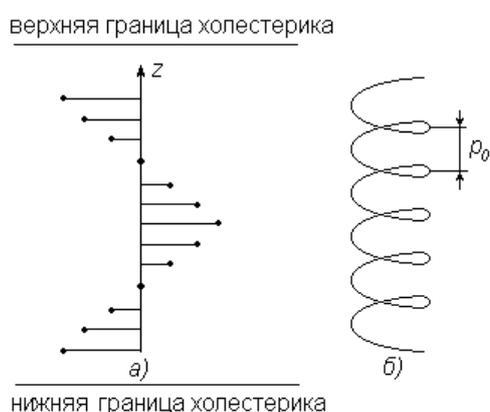


Рисунок 50 – Строение холестерического жидкого кристалла в поперечном сечении

масштабах. Эти отличия связаны с тем, то в нем молекулы образуют **холестерическую спираль**: оказывается в нематике существует направление, называемое **холестерической осью**, вдоль которого ориентация директора регулярным образом изменяется. Директор перпендикулярен этой оси и вращается вокруг нее (рис. 50).

Угол поворота линейно связан с расстоянием z вдоль холестерической оси и может быть представлен в виде: $\varphi = (2\pi / p_0) \cdot z$. Расстояние p_0 , на котором директор поворачивается на 360° , называется **шагом холестерической спирали**. Холестерическая фаза может наблюдаться сразу после плавления кристалла или после смектической фазы. При дальнейшем повышении температуры он переходит в изотропную жидкость.

8.4. Ориентационные эффекты в жидких кристаллах

Нематики. Как указывалось выше, в исходном состоянии пленки жидких кристаллов по своим свойствам разбиты на отдельные домены, в пределах которых ориентация длинных осей молекул постоянна. Для научных исследований и практических применений необходимы однодоменные пленки. Это предполагает, что за счет принятых мер все молекулы жидкокристаллического образца нематика ориентированы одинаковым образом. Существуют две основных ориентации молекул нематика: вдоль поверхности и перпендикулярно ей. В соответствии с этим используют два типа жидкокристаллических ячеек, показанных на рис.51.

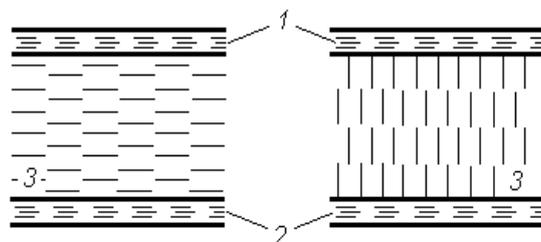


Рисунок 51 – Ориентация молекул нематика в планарной и гомеотропной оптических ячейках

Здесь 1 – верхняя опорная поверхность ячейки, 2 – нижняя опорная поверхность, которые обычно изготавливают из прозрачного стекла. Пленка нематика – 3 помещается между опорными поверхностями. Ячейку, показанную в левой части рис. 65, называют **планарной**, а показанную в правой части рис. 65 – **гомеотропной**. Для получения планарной ячейки внутренние поверхности стекол, контактирующие с жидким кристаллом, подвергают механической шлифовке в одном направлении, в результате которой на поверхности стекла остаются микроборозды. Молекулы нематика «ложатся» в эти борозды, образуя строго ориентированный поверхностный слой нематика планарной структуры. Ввиду действия сил связи молекулы соседних слоев перенимают ориентацию граничного слоя, передавая его в объем жидкого кристалла. Силы упругости между молекулами стремятся разрушить этот порядок. Поэтому существует некоторое расстояние, при удалении на которое от поверхности ориентирующее действие опорной поверхности и разориентирующее влияние упругих сил уравниваются друг друга. Это расстояние называется **когерентной длиной** и для типичных жидких кристаллов составляет величину несколько микрон. Поэтому, если толщина слоя нематика невелика – до единиц микрон, то разориентирующее действие теплового движения несущественно и во всем объеме жидкого кристалла его молекулы (на рисунке они показаны в виде коротких линий между стеклами) сориентированы вдоль поверхности ячейки.

Для получения гомеотропной ячейки внутренние поверхности стекол обрабатывают в специальном химическом веществе – поверхностно-активном веществе. Молекулы этого вещества также имеют вытянутую форму и при обработке одним концом «приклеиваются» к обрабатываемым поверхностям. Молекулы жидкого кристалла входят в пространство между

молекулами поверхностно-активного вещества, ориентируясь перпендикулярно поверхности ячейки. Следующие слои молекул нематика, стремясь минимизировать свою потенциальную энергию по отношению к молекулам соседнего слоя жидкого кристалла, выстраиваются также перпендикулярно стеклам.

Переход Фредерикса. Для технических применений жидких кристаллов важной является возможность изменять ориентацию молекул путем приложения, например, электрического поля, которое обычно прикладывают перпендикулярно поверхностям ячеек. Важной особенностью нематиков является различное значение его диэлектрической проницаемости ϵ , измеренной вдоль осей молекул и перпендикулярно им. Поэтому при помещении нематика в постоянное электрическое поле из-за анизотропии диэлектрической проницаемости его молекулы стремятся ориентироваться так, чтобы направление большего значения ϵ совпадало с направлением приложенного поля. Значит, исходная планарная структура нематика в электрическом поле может перейти в гомеотропную или наоборот. Эффект изменения структуры нематика во внешнем электрическом поле называют **переходом Фредерикса**. Общим свойством переориентации нематика под действием внешнего поля является ее пороговый характер. Критическое значение электрического поля, при котором совершается переход Фредерикса, вычисляется по следующей формуле:

$$E_c = \frac{\pi}{d} \left(\frac{4\pi K}{\Delta\epsilon} \right)^{1/2},$$

где d – толщина слоя жидкого кристалла, K – его коэффициент упругости, $\Delta\epsilon = \epsilon_{||} - \epsilon_{\perp}$ анизотропия диэлектрической проницаемости жидкого кристалла. Здесь $\epsilon_{||}$ – диэлектрическая проницаемость жидкого кристалла, когда поляризация света параллельна директору, ϵ_{\perp} – диэлектрическая проницаемость жидкого кристалла, когда поляризация света перпендикулярна директору. Из этого выражения следует, что критическое напряжение, прикладываемое к слою нематика для его переориентации, не зависит от толщины d :

$$V_c = \pi \left(\frac{4\pi K}{\Delta\epsilon} \right)^{1/2}. \quad (8.1)$$

Подставив в выражение (9.1) типичные значения: $K=10^{-6}$ дин, $\Delta\epsilon \approx 3$, получим, что $V_c \approx 2$ В. Выше этого значения напряжения на ячейке изменения структуры быстро достигают насыщения.

Ориентационное действие электрического поля величиной E проявляется на протяжении **электрической когерентной длины**, отсчитываемой от опорной поверхности,

$$\lambda_k = \frac{1}{E} \left(\frac{4\pi K}{\Delta\epsilon} \right)^{1/2}. \quad (8.2)$$

Подставив в это выражение типичные значения: $K=10^{-6}$ дин, $\Delta\epsilon \approx 3$ и $E = 10^3$ В/см, получим, что $\lambda_k \approx 10$ мкм. Следовательно, в практических приложениях пленки нематика должны иметь толщину значительно больше двух электрических когерентных длин, что соответствует нескольким десяткам микронов.

Если увеличивать напряжение на ячейке выше порога Фредерикса, то вблизи десятка вольт структура нематика претерпевает следующее скачкообразное изменение: вдоль поверхности ячейки образуются круговые потоки молекул нематика, что формирует структуру типа упорядоченного слоя вращающихся цилиндров. Это так называемые **домены Капустина-Вильямса**. Их возникновение обусловлено проявлением ионной проводимости жидкого кристалла в электрическом поле: ионы движутся к противоположно заряженной опорной поверхности, где и изменяют заряд. Затем движутся в обратном направлении и все повторяется на другой опорной поверхности. Из-за изменения ориентации молекул в

пределах домена он выступает как неоднородность диэлектрической проницаемости, изменяющейся по толщине ячейки по гармоническому закону.

Дальнейшее повышение напряжения на ячейке приводит сначала к увеличению числа цилиндров-доменов по толщине жидкого кристалла, а затем к разрушению доменов и появлению хаотических турбулентных потоков в слое нематика. В результате такого состояния структуры жидкого кристалла он рассеивает падающий на него свет равномерно во все стороны. Поэтому данное состояние нематика **называют динамическим рассеянием света**, при котором рассеяние падающего света происходит во всех направлениях, что облегчает визуальное наблюдение измененного состояния жидкого кристалла.

Перечисленные структурные состояния нематика в электрическом поле могут быть использованы в электронике: для отображения и хранения информации, преобразования световых потоков и т.д.

Холестерики. Следует заметить, что шаг холестерической спирали чувствителен к электрическому полю, температуре, химическим примесям на поверхности холестерика, радиационному излучению и т.д.: при приложении, например, электрического поля шаг спирали увеличивается с ростом величины поля. При приближении к некоторому пороговому напряжению (несколько вольт) шаг холестерической спирали начинает резко увеличиваться, а при пороговом напряжении

$$E_p = \left(\frac{\pi^2}{2\rho_0} \right) \cdot \left(\frac{4\pi K}{\Delta\varepsilon} \right)^{1/2} \quad (8.3)$$

спираль вообще распрямляется. В результате этого холестерик переходит в нематик. Здесь ρ_0 – шаг холестерической спирали в отсутствие воздействия, который обычно выбирают в десятки раз меньше толщины слоя жидкого кристалла. Поэтому значение E_p оказывается больше, чем поле, необходимое для перехода Фредерикса. Если продолжать увеличивать напряжение на пленке холестерика, то далее можно обнаружить те же структурные преобразования, о которых говорилось при обсуждении свойств нематика: домены Капустина-Вильямса, динамическое рассеяние.

Поведение холестерика в электрических полях сложнее нематика. Так, динамическое рассеяние в холестериках может обладать эффектом памяти: рассеивающее состояние холестерика сохраняется в течение длительного времени после снятия поля. Время памяти, т.е. сохранения рассеивающего состояния, зависит от свойств конкретного холестерика и может колебаться от минут до нескольких лет. Наложение переменного электрического поля возвращает холестерик в исходное состояние.

9.5. Оптические свойства жидких кристаллов

Для понимания принципов работы оптических жидкокристаллических ячеек необходимо вспомнить раздел оптики, в котором рассматриваются особенности прохождения через оптически активный кристалл линейно поляризованного излучения с заданной ориентацией поляризации и направлением распространения излучения относительно оптической оси кристалла.

Нематики. Нематик является оптически одноосным кристаллом, оптическая ось которого совпадает с направлением директора. Поэтому оптические характеристики планарной и гомеотропной структур различны. В ячейке с гомеотропной структурой свет распространяется вдоль оптической оси (вдоль директора), поэтому, как во всяком одноосном кристалле, в этом случае оптические характеристики ячейки не зависят от поляризации света: свет проходит жидкокристаллическую оптическую ячейку без двулучепреломления, а значит без изменения фазы. При падении же света на ячейку с планарной структурой прохождение света сквозь нее зависит от поляризации падающего света: без изменения поляризации в ней распространяется только свет, линейно поляризованный ортогонально директору. В случае поляризации по директору свет

расщепляется на обыкновенную волну и необыкновенную, которые имеют разные скорости распространения и взаимно перпендикулярные поляризации. В итоге на выходе из оптической ячейки эти волны складываются в одну, вектор поляризации которой оказывается повернут на некоторый угол относительно поляризации на входе. Подбирая толщину ЖК или значения его диэлектрических проницаемостей ϵ_{\parallel} и ϵ_{\perp} , можно добиться того, что на выходе поляризация будет перпендикулярна к поляризации на входе в планарную оптическую ячейку. Используя поляроидные пленки, нанесенные на опорные поверхности, можно сделать гомеотропную структуру непрозрачной для света, скрестив верхний и нижний поляроиды. Планарная структура при таком же положении поляроидов пропускает свет, если ориентация поляроида не совпадает с направлением директора. Отсюда следует принцип, на основе которого с помощью жидкокристаллической ячейки можно отображать информацию – путем прикладываемого напряжения переводить планарную текстуру в гомеотропную и обратно. В планарной структуре направление распространения света перпендикулярно оптической оси. Поэтому показатель преломления света различен для света, плоскость поляризации которого перпендикулярна направлению директора в слое, и света, поляризация которого содержит это направление. Знак оптической анизотропии может быть как положительным, так и отрицательным. В нематике с $\Delta\epsilon < 0$ фазовая скорость света, поляризованного перпендикулярно направлению директора, оказывается меньше фазовой скорости света, поляризованного в направлении директора.

Переход Фредерикса, в котором при отсутствии поля нематик с $\Delta\epsilon > 0$ образует планарную текстуру, а при приложении поля переходит в гомеотропную, называют **S-эффектом**.

В **B-эффекте** исходная текстура гомеотропная, а поле переводит ее в планарную для чего требуется нематик с отрицательной диэлектрической анизотропией $\Delta\epsilon$.

Для **T-эффекта** исходная планарная структура берется закрученной (ориентация молекул на одной поверхности составляет угол 90° относительно ориентации на другой). Приложенное поле, как и в случае S-эффекта, изменяет планарную ориентацию молекул нематика с $\Delta\epsilon > 0$ на гомеотропную. Работа ячейки, использующей T-эффект, основывается на свойстве закрученной твист-структуры вращать плоскость поляризации света.

Цветовые эффекты в НЖК. Электрооптические эффекты в ячейке могут проявляться не только в переходах из непрозрачного состояния в прозрачное или наоборот, но также в изменении окраски ячейки – это так называемый эффект «гость – хозяин». Сущность данного эффекта заключается в следующем. Если в нематик (хозяин) в небольшом количестве добавить молекулы дихроичного красителя (гость), которые ориентируются нематиком так же, как и собственные молекулы, то при переходе Фредерикса вместе с переориентацией молекул нематика происходит переориентация молекул красителя. В результате этого изменяется окраска жидкокристаллической ячейки из-за изменения ориентации молекул дихроика. Это происходит потому, что дихроичные молекулы красителя имеют линию поглощения света на строго определенной длине волны, причем величина коэффициента поглощения на этой волне зависит от взаимной ориентации поляризации падающего на него света и длинной оси молекул дихроика. Так что свет с поляризацией перпендикулярной оси дихроичной молекулы ею не поглощается, а с поляризацией параллельной оси дихроика – поглощается. В последнем случае краситель приобретает цвет, определяемый линией поглощения дихроика: это может быть и красный цвет ($\lambda=0,64$ мкм), и зеленый ($\lambda=0,55$ мкм) и т.д. В соответствующий линии поглощения цвет окрашивается оптическая ячейка.

Поскольку молекул красителя в нематик добавляют мало, все диэлектрические характеристики жидкого кристалла сохраняются. Можно ввести два красителя с различными частотами дихроичных линий поглощения и тем самым добиться двойного изменения окраски ячейки при переключении электрического поля: один цвет до действия поля и другой – после перехода Фредерикса.

Смектики. Оптические свойства смектиков находятся в прямой зависимости от их строения. Смектик *A*-типа, как и нематик, является одноосной средой в оптическом отношении – в нем существует одно направление распространения света, совпадающее с директором, для которого показатель преломления не зависит от поляризации света.

Смектики *C*-типа являются оптически двухосными, т.к. ориентационные движения молекул в смектиках отличаются от таковых в нематике. Если в нематике молекулы могут свободно вращаться вокруг своих длинных осей, то в смектиках в общем случае вращательное движение заторможено, т.е. хотя молекула и вращается, но с наибольшей вероятностью она принимает определенное ориентационное положение. Поскольку свет «видит» усредненное положение молекул, то в случае свободного вращения усредненная молекула осесимметрична, а при заторможенном вращении она выглядит как часть плоскости (блиноподобна). Из теории анизотропных сред известно, что ансамблю осесимметричных молекул соответствует оптически одноосная среда, а системе молекул, у которых отсутствует ось вращения, соответствует оптически двухосная среда.

В природе встречаются смектики, обладающие большим электрическим дипольным моментом. Происхождение дипольного момента связано с тем, что образующие смектик молекулы обладают собственным электрическим моментом: центры «тяжести» (говорят – центроиды) положительного и отрицательного зарядов в отдельной молекуле сдвинуты в пространстве относительно друг друга. Направление дипольного момента изменяется от слоя к слою, образуя спиральную структуру. Если эту спираль раскрутить электрическим полем, перпендикулярным оптической оси, то можно обнаружить дипольный момент. При этом изменяются оптические свойства смектика, что можно использовать на практике.

Холестерики. Этот тип жидких кристаллов характеризуется спиральным поведением директора. Но с изменением направления директора изменяется диэлектрическая проницаемость. Следовательно, по толщине кристалла значение диэлектрической проницаемости изменяется от ϵ_{\parallel} – диэлектрическая проницаемость жидкого кристалла, когда

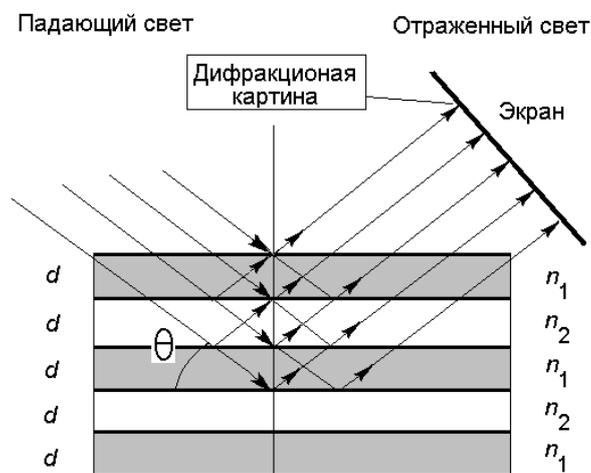


Рисунок 52 – Возникновение дифракции Вульфа – Брэгга на периодической слоистой структурой с чередованием слоёв с разным показателем преломления

поляризация света параллельна директору, до ϵ_{\perp} – диэлектрическая проницаемость жидкого кристалла, когда поляризация света перпендикулярна директору (рис.52). Диэлектрические свойства холестерика периодичны по координате с периодом $d = p_0/2$. В холестерике периодом служит не шаг спирали p_0 , на котором директор совершает поворот на 2π , а $p_0/2$, так как ориентации директора \vec{n} и $-\vec{n}$ эквивалентны, и период диэлектрических свойств холестерика оказывается вдвое меньше, чем расстояние, на котором директор совершает полный оборот. По этой причине холестерик можно представить в виде

чередующихся слоев двух типов с различными значениями проницаемости: ε_{\parallel} и ε_{\perp} . Поскольку шаг спирали выбирается в диапазоне длин волн видимого света, то свет может дифрагировать на периодической структуре холестерика.

Если на такую структуру под углом θ - углом между поверхностью кристалла и направлением падения света, падает белый свет (т.е. излучение с широким диапазоном значений длин волн), то в отраженном свете будет иметь место дифракционная картина, полученная от оптического излучения только одной длины волны λ из всего диапазона длин волн падающего излучения. Внешне это будет выглядеть как окрас слоя холестерика в цвет длиной λ . Это явление, называемое дифракцией Вульфа-Брэгга, рассматривалось нами ранее в главе 5. Условие наблюдения этого вида дифракции таково:

$$\sin \theta = \frac{m\lambda}{2d},$$

где $d = p/2$ - толщина слоя неоднородности холестерика (взаимобратная ориентация директора в холестерике на шаге спирали идентична по оптическим свойствам), p - шаг холестерической спирали, m - порядок дифракционного максимума.

Ранее отмечалось, что шаг спирали холестерика чувствителен к внешним воздействиям. Его легко изменить, нагревая холестерик, помещая в электрическое или магнитное поле и т.д. Отсюда следует один из принципов использования холестерических кристаллов в оптоэлектронике: плавное изменение цвета оптической ячейки в широких пределах при действии электрического поля, теплового или ионизирующего излучения.

8.6 Применение жидких кристаллов в оптоэлектронике

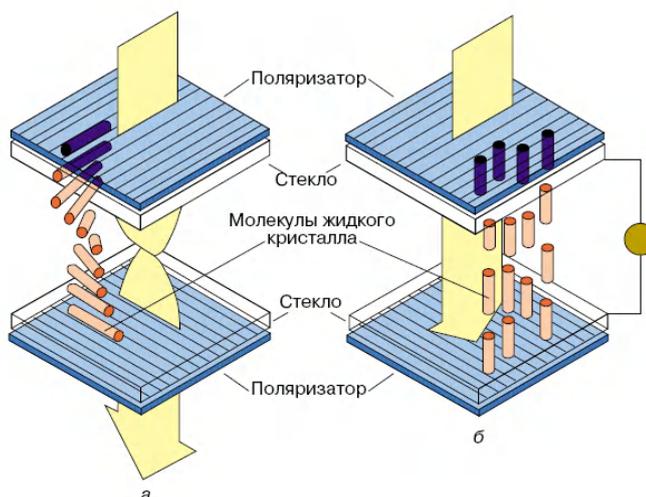
В виду разной степени изученности различных типов жидких кристаллов в современной электронике используются преимущественно нематики и частично холестерики. Сметтики не проявляют выраженного перехода Фредерикса, и потому в настоящее время их использование в электронике ограничено.

Жидкие кристаллы в твердотельной электронике применяют главным образом для создания модуляторов, построения дефлекторов оптического излучения, интегрально-оптических элементов различного назначения, систем обработки оптической информации – оптических транспарантов, управляемых электрическим полем или светом, устройств отображения информации. Использование нематиков базируется преимущественно на использовании эффекта Фредерикса, требующем импульсов напряжений малой амплитуды (до 5 Вольт). Это обстоятельство дополняет хорошую интегрируемость жидкокристаллических элементов с микропроцессорной техникой и позволяет создавать программно-управляемые экономные, малогабаритные и надежные оптоэлектронные элементы, и устройства.

1. Жидкокристаллическая ячейка отображения информации. В основе работы ячеек по отображению информации используют переход Фредерикса в твист-структуре, а также эффект динамического рассеяния света в S -эффекте только в черно-белом цвете. Ячейки следующего поколения уже цветные за счет применения эффекта «гость-хозяин» в нематике либо использования холестерика с электрически регулируемым шагом спирали.

В качестве примера рассмотрим конструкцию планарной черно-белой ячейки, предназначенную для отображения информации на основе твист-эффекта (закрученной на 90 градусов ЖК пленки), как наиболее используемой в оптоэлектронике. Оптическая ячейка образована двумя стеклянными пластинками, на внутренние поверхности которых нанесены металлические электроды с высоким коэффициентом пропускания света. На внешние поверхности нанесены поляроидные пленки, они пропускают свет только одной поляризации. За основу ячейки берут ЖК с положительной оптической анизотропией. Исходная ориентация нематика закрученная, так называемая твист-ориентация молекул (рис. 53, а). Свет падает на верхний поляризатор и становится плоско поляризованным в соответствии с его поляризацией. При отсутствии электрического поля (то есть в выключенном состоянии) свет, “следуя” твист-ориентации нематика, меняет свое направление в соответствии с

оптической осью нематика и на выходе будет иметь то же направление поляризации, что и нижний поляризатор. При включении электрического поля для нематического жидкого кристалла с положительной диэлектрической анизотропией ($\Delta\epsilon > 0$) произойдет переход от закрученной твист-ориентации к гомеотропной ориентации молекул, то есть длинные оси молекул повернутся в направлении, перпендикулярном к электродам, и спиральная структура разрушится (рис. 53, б). Теперь свет, не изменив направления исходной поляризации, совпадающей с поляризацией верхнего поляризатора, будет иметь направление



Риснок 53 – Конструкция и принцип работы оптической ячейки с твист-структурой; исходное состояние – а), состояние после перехода Фредерикса – б)

поляризации, противоположное нижнему поляризатору, а они, как видно на рис. 53, б, находятся в скрещенном положении. В этом случае ячейка — «непрозрачна». Используя это свойство ячейки, можно создать электронный затвор оптического излучения.

2. Жидкокристаллические модуляторы представляют собой оптическую планарную или гомеотропную ячейку, которую размещают между скрещенными или параллельными поляризаторами. Падающее оптическое излучение имеет линейную поляризацию. В зависимости от конструкции модулятора и используемого электрооптического эффекта при наложении электрического поля, создающего один из перечисленных выше структурных эффектов, его пропускная способность изменяется от нуля до 100% в соответствии с законом изменения прикладываемого напряжения. Предельная частота модуляции такого модулятора обычно составляет несколько килогерц. Поэтому такие модуляторы чаще используются как оптические затворы. Однако недалеко то время, когда жидкокристаллические модуляторы будут работоспособны до сотен килогерц и выше.

3. Дефлекторы оптического излучения могут быть созданы на основе многочисленных электрооптических эффектов в жидких кристаллах: переходе Фредерикса, доменах Капустина-Вильямса, использовании нарушения полного внутреннего отражения и т.д. Назначение дефлектора – отклонение оптического луча в нужном направлении. Наиболее просто этого достичь, если в жидкокристаллической ячейке выбором треугольной топологии полевых электродов сформировать клиновидную призму, изменяющую свой относительный показатель преломления в электрическом поле от нуля до величины, определяемой анизотропией диэлектрической проницаемости.

4. Интегрально-оптические элементы на основе жидких кристаллов представляют собой элементы, встроенные в оптическую схему оптоэлектронного устройства. Это активные линзы, волноводные модуляторы, волноводные дефлекторы и т.д. Их применение позволяет программно управлять оптической схемой оптоэлектронного устройства, изменять его функции. Подавая на эти элементы в соответствии с алгоритмом работы устройства

напряжения выше порога Фредерикса, можно жидкокристаллической призмой изменять направление распространения светового луча; активируя жидкокристаллическую линзу, включать-выключать фокусировку; оптической ячейкой осуществлять модуляцию света цифровым сигналом. Однако ввиду значительных оптических потерь в жидких кристаллах интегрально-оптические элементы применяют на малых участках волноводных линий.

5. Оптические транспаранты. Для современных систем обработки информации большое значение имеют *оптоэлектронные транспаранты*, управляемые полем или светом. Транспаранты представляют собой пространственные модуляторы или, точнее, пространственные световые ключи, управляющие по заданному временному закону одним или несколькими параметрами светового потока. Жидкие кристаллы являются удобными веществами для транспарантов ввиду малых управляющих напряжений и потребляемой мощности, хорошего контраста и возможности изготовления транспарантов практически неограниченных размеров.

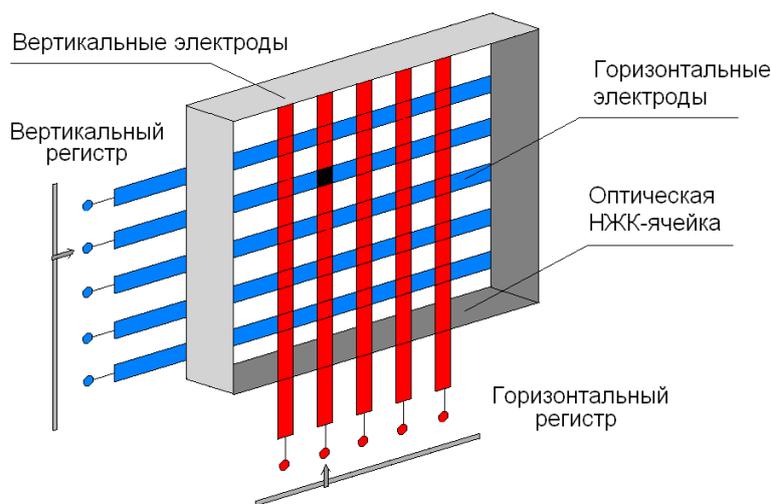


Рисунок 54 – Конструкция и принцип работы жидкокристаллического транспаранта, управляемый электрическим полем

Так, жидкокристаллический транспарант, управляемый электрическим полем, представляет собой матрицу, составленную из большого числа независимых оптических ячеек. На матрицу нанесены прозрачные металлические электроды в виде полосок, направленных взаимно перпендикулярно на противоположных опорных стеклах (рис. 54). При подаче напряжения на определенную комбинацию вертикальных и горизонтальных полосок выбранные таким образом ячейки изменяют свои оптические свойства, в частности, коэффициент пропускания. На представленном рисунке переход Фредерикса происходит в ячейке с координатами (2, 4). Если транспарант работает на просвет, то обе опорные поверхности делают прозрачными. Если же транспарант работает на отражение, то одна опорная поверхность делается зеркальной.

Транспаранты, управляемые светом, отличаются от описанных выше тем, что в них жидкокристаллическая ячейка включается последовательно с высокоомным фотослоем (фоторезистором). Прикладываемое к транспаранту постоянное напряжение берется выше напряжения перехода Фредерикса. В отсутствие освещения из-за большого падения напряжения на фотослое напряжения ячейке недостаточно для структурных изменений в ней. Освещение фотослоя повышает его проводимость в местах освещения, и тогда в освещенных точках транспаранта приложенное напряжение перераспределяется в пользу слоя жидкого кристалла. Это вызывает переход Фредерикса и изменение структуры жидкого кристалла, что легко обнаружить в проходящем или отраженном свете (см. рис.54). Таким образом, изображение на жидкокристаллическом экране в точности повторяет световую

картинку, спроецированную на его заднюю стенку. Если проецируемое на транспарант изображение инфракрасное, то такой транспарант позволяет сделать его видимым.

Жидкокристаллические транспаранты широко используются для формирования двумерных массивов информации, проведения логических операций, преобразования и усиления яркости излучения, перестраиваемой фильтрации изображения и т.д.

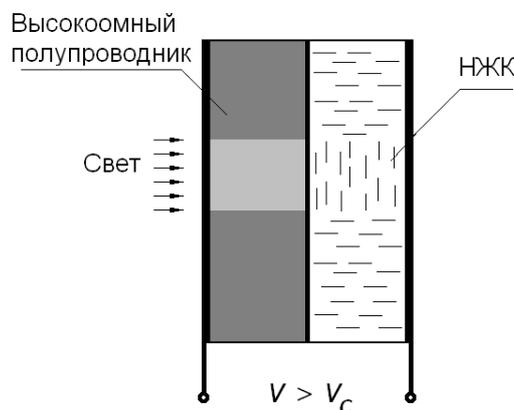


Рисунок 54 – Конструкция и принцип работы жидкокристаллического транспаранта, управляемый светом

6. Устройства отображения информации на жидких кристаллах характеризуются дешевизной, низким энергопотреблением, малыми управляющими напряжениями и хорошей технологической совместимостью с интегральными микросхемами. Они используют переход Фредерикса в твист-структуре, а также эффект динамического рассеяния света в S-эффекте. Однако эти индикаторы отображают информацию в черно-белом цвете. Индикаторы третьего поколения являются полноцветными. Полноцветные устройства отображения информации созданы на основе эффекта «гость-хозяин» в нематике либо на основе холестерика.

7.6. Контрольные вопросы по главе 7

1. В чём различие между эффектами электропоглощения и эффектами электропреломления?
2. Что такое лестницы Штарка и как они образуются в сверхрешётках?
3. В чём суть эффекта Франца – Келдыша и каков механизм его действия на оптическое излучение?
4. Каковы причины появления виртуальных состояний в эффекте Франца – Келдыша?
5. Как использовать эффект Франца – Келдыша для модуляции оптического излучения?
6. Как действует эффект двулучепреломления в 3D полупроводниках?
7. Каков механизм двулучепреломления в жидких кристаллах?
8. Как использовать эффект двулучепреломления для модуляции оптического излучения?
9. В чём суть квантоворазмерного эффекта Штарка?
10. Как использовать квантоворазмерный эффект Штарка для модуляции оптического излучения?

9 ПЕРСПЕКТИВЫ РАЗВИТИЯ ОПТОЭЛЕКТРОНИКИ

В настоящее время достаточно четко просматриваются три направления развития оптоэлектроники на ближайшие годы: использование для создания приборов электроники новых материалов; применение новых физических явлений и принципов; дальнейшая миниатюризация приборов за счет использования квантовых эффектов в твердом теле.

Новые материалы. Применение новых материалов для создания элементов твердотельной электроники способно существенным образом способствовать развитию этой области науки. Так, изготовление генераторов оптического излучения и фотоприемных элементов на основе гетеропереходов (*p-n*-переходы, полученные на основе полупроводниковых материалов с различной шириной запрещенной зоны), создать систему «приемник-передатчик» на оптимальной длине волны для заданного оптического тракта. Это происходит за счет получения узкой спектральной характеристики обоих элементов связи, а также управления спектральной характеристикой в широком диапазоне длин волн.

Широкие возможности в области совершенствования элементной базы электроники и создания новых приборов сулит также использование нового типа полупроводникового материала: сверхрешеток, получаемых из тонких 100-ангстремных слоев полупроводниковых материалов с различной шириной запрещенной зоны. Подбором свойств исходных полупроводников можно получать многослойную структуру, которая будет иметь желаемую ширину запрещенной зоны.

Новые физические явления и принципы – электронная и оптическая синергетика для создания сверхскоростных бистабильных элементов; высокотемпературная сверхпроводимость; голографические оптические методы записи информации в кристаллах; переходы металл-полупроводник; использование эффекта фотонного эха; использование нелинейных свойств жидких кристаллов для преобразования оптического излучения; применение твердотельных автоволн для создания пространственно-временных модуляторов излучения.

Разработка новых технологий. На этапе становления электроники (первые лет 15) основными направлениями ее развития были формирование номенклатуры базовых элементов и их реализация в виде дискретных элементов. Дальнейший прогресс этой области науки связан с переносом центра внимания на создание функционально сложных и конструктивно законченных приборов, устройств, систем. Именно в них наиболее полно проявляются преимущества биоэлектроники перед микроэлектроникой. Особенно глобальные изменения в оптоэлектронике будут связаны с использованием идей и технологии нового направления в микроэлектронике – наноэлектроники (электроники атомного уровня), которое принципиальным образом изменит элементную базу электроники, конструкторские подходы к созданию электронных устройств и систем.

10 ЛИТЕРАТУРА

1. Епифанов, Ю.И. Физические основы микроэлектроники. Ю.И. Епифанов– М.: Высшая школа, 1971. – 388 с.
2. Солимар, Л. Лекции по электрическим свойствам материалов. Л. Солимар, Д. Уолш. Пер. с англ. С.И. Баскакова. – М.: - Мир, 1991. – 501 с.
3. Шарма, Б.Л., Пурохит Р.К. Полупроводниковые гетеропереходы. Б.Л. Шарма, Р.К. Пурохит. Пер. с англ. Под ред. Ю.В. Гуляева. – М.: Сов. Радио, 1979. – 227 с.
4. Жеребцов, И.П. Основы электроники. И.П. Жеребцов. – Л-д: Энергоатомиздат, 1989. – 352 с.
5. Шалимова, К.В. Физика полупроводников. К.В. Шалимова. – М.: Энергия. - 1991, - 416 с.
6. Панков, Ж. Оптические переходы в полупроводниках. Ж. Панков. – М.: Мир, 1973. – 393 с.
7. Пикин, С.А. Жидкие кристаллы. С.А. Пикин, Л.М. Блинов. – М.: Наука, 1982. – 207 с.
8. Зи, С. Физика полупроводниковых приборов. Том 1. С. Зи. Пер. с англ. под ред. Р.А. Суриса. – М.: Мир – 1984. – 453 с.
9. Шуберт, Ф. Светодиоды. Ф. Шуберт. - М.: Мир. – 2008. – 498 с.