МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЕВА
(национальный исследовательский университет)»

В. Д. Паранин, С. А. Матюнин

Элементы и устройства интегральной оптоэлектроники

Курс лекций

Электронное учебное пособие

САМАРА

2012

УДК 621.3
П 180

Авторы : **Паранин Вячеслав Дмитриевич,**
Матюнин Сергей Александрович

Паранин, В. Д. Элементы и устройства интегральной оптоэлектроники [Электронный ресурс] : курс лекций : электрон. учеб. пособие / В. Д. Паранин, С. А. Матюнин; Минобрнауки России, Самар. гос. аэрокосм. ун-т им. С. П. Королева (нац. исслед. ун-т). - Электрон. текстовые и граф. дан. (7 Мбайт). - Самара, 2012. - 1 эл. опт. диск (CD-R).

Режим доступа: http://rtfmoodle.ssau.ru/course/view.php?id=66

Курс лекций предназначен для обучения бакалавров радиотехнического факультета по направлению подготовки 210100.62 «Электроника и наноэлектроника», изучающих дисциплину «Элементная база интегральной оптоэлектроники» в 8 семестре.

Электронное учебное пособие разработано на кафедре электронных систем и устройств.

© Самарский государственный
аэрокосмический университет, 2012

СОДЕРЖАНИЕ

		Стр.
1	ВВЕДЕНИЕ……..	5
2	**Лекция 1.** Концепция и основные понятия интегральной оптоэлектроники. Физические основы интегральной оптоэлектроники. Назначение и состав интегрально-оптических схем. Характеристики интегрально-оптических схем……	6
3	**Лекция 2.** Полупроводниковые источники излучения. Рекомбинационное излучение и поглощение. Энергетические уровни. Квантовый выход. Спонтанная и лазерная люминесценция в полупроводниковых структурах. Когерентность, монохроматичность, поляризация излучения. ……	20
4	**Лекция 3.** Полупроводниковые источники некогерентного излучения. Основные физические параметры. Полупроводниковые лазеры. Лазерные диоды на p-n переходе, с туннельной инжекцией, на гетероструктурах, с распределенной обратной связью. Вопросы надежности……..	44
5	**Лекция 4.** Классификация и основные параметры приемников излучения интегрально-оптических схем. Фотодиод на p-n переходе. Волноводный, лавиный, p-i-n фотодиод. Фотоприемники с гетеропереходом, на основе МДП-структуры. Изменение спектральных характеристик. ……...	90
6	**Лекция 5.** Волноводные усилители. Выходная мощность сигнала, эффективность накачки. Ширина и равномерность полосы усиления. Полупроводниковые усилители. Эрбиевые усилители……	114
7	**Лекция 6.** Принцип работы волноводов. Методы и основные этапы изготовления. Тонкопленочные, диффузионные, имплантированные, эпитаксиальные, электрооптические волноводы. Измерение толщины, затухания, эффективных показателей преломления	123

волноводных мод.....

8 **Лекция 7.** Потери в оптических волноводах. Абсорбционные потери. Потери на излучение. Ввод-вывод излучения волноводов. Основы оптического согласования. Фокусирующие, торцевые, призменные, решеточные, сужающиеся элементы. 163

9 **Лекция 8.** Геометрическая оптика планарных волноводов. Классификация мод планарного волновода. Волноводные моды тонкопленочного волновода. Эффективная толщина волновода. Градиентные планарные волноводы. 196

10 **Лекция 9.** Электромагнитная теория планарных волноводов. Волновые уравнения планарных волноводов. Моды тонкопленочного волновода. Свойства мод тонкопленочного волновода. Волновые уравнения для градиентных планарных волноводов..... 218

11 **Лекция 10.** Элементы интегрально-оптических схем. Связь между волноводами. Разветвители. Рупорные переходы. Фокусирующие элементы. Поляризационные элементы. Спектральные фильтры и резонаторы. 243

12 **Лекция 11.** Интегрально-оптические модуляторы и затворы. Основные технические параметры и характеристики. Электрооптические модуляторы Маха-Цандера, бегущей волны, на связанных волноводах, дифракционные, поглощающие. Акустооптические модуляторы и дефлекторы..... 288

13 **Лекция 12.** Примеры интегрально-оптических схем. Многоканальный переключатель. Когерентный оптический приемник. Анализаторы спектра высокочастотных сигналов. Интегрально-оптические корреляторы, АЦП, ЦАП. Факторы, ограничивающие возможности интегральной оптики..... 322

14 СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ..... 352

ВВЕДЕНИЕ

Одним из наиболее интенсивно развивающихся направлений информационно-измерительной техники и техники связи является интегральная оптоэлектроника. Объединяя достижения физики твердого тела, интегральной и волоконной оптики, а в последнее время – нанотехнологий и нанофотоники, данное направление позволяет обеспечить создание и комплексную миниатюризацию функциональных элементов и устройств. Становится возможным существенное увеличение информационной емкости каналов связи, и, следовательно, повышение качества управления объектами, осуществление мониторинга сложных технологических процессов в реальном времени.

Применение интегральной оптоэлектроники в сочетании с волоконно-оптическими линиями связи позволяет создать мультисенсорные системы для эффективного управления авиационной и космической техникой, промышленными объектами различного назначения. Несомненными достоинствами таких систем являются пожаро- и взрывобезопасность, малые габариты и вес, химическая инертность (стойкость), гальваническая отвязка от токонесущих конструкций, возможность осуществления бесконтактных и дистанционных измерений, возможность осуществления мультиплексирования передаваемой информации.

В учебном пособии изучаются теоретические основы работы полупроводниковых источников и приемников излучения, волоконно-оптичсеких лазеров и усилителей, основные параметры диффузионных и эпитаксиальных волноводов, элементов оптической связи с волноводами. Рассматриваются параметры и характеристики электрооптических модуляторов и коммутаторов дифракционного, интерференционного, отражательного типов, а также акустооптических модуляторов Брэгга.

Лекция 1

Концепция и основные понятия интегральной оптоэлектроники. Физические основы интегральной оптоэлектроники. Назначение и состав интегрально-оптических схем. Характеристики интегрально-оптических схем.

Интерес к использованию для передачи и обработки информации оптических пучков вместо традиционных электрических токов и радиоволн появился в самом начале 60-х гг., когда развитие лазерной техники привело к созданию необходимых для этого надежных источников когерентного излучения. Лазерные пучки могут распространяться в атмосфере, однако атмосферные изменения оказывают нежелательные воздействия на оптические характеристики, причем такие изменения имеют место не только от одного дня к другому, но и, буквально ежесекундно, от одного мгновения к другому. С другой стороны, лазерные пучки можно приспособить для обработки информационных сигналов, однако для этого требуются различные оптические элементы, например призмы, зеркала, модуляторы, детекторы. Подобные элементы обычно устанавливают на оптической скамье длиной до нескольких метров, которую для исключения влияния вибраций подвешивают на специальные опоры (либо ставят на них). Такие схемы приемлемы для проведения лабораторных экспериментов по оптике, однако их использование для прикладных целей существенно затруднено. Поэтому в конце 60-х гг. появилась концепция «интегральной оптики», согласно которой провода и радиолинии заменяют на оптические волоконные световоды, по которым (а не по воздуху) передают оптические сигналы; при этом обычные электрические интегральные схемы заменяют на миниатюризированные оптические интегральные схемы (ОИС). В конце 70-х гг. ситуация сложилась таким образом, что ряд факторов способствовал выходу интегральной оптики за пределы лабораторий в области практического использования; к таким факторам следует отнести разработку оптических волокон с малыми потерями и элементов связи, создание надежных полупроводниковых лазеров

непрерывного действия на основе GaAlAs и развитие методов фотолитографии, с помощью которых можно получать линии шириной, лежащей в субмикронном диапазоне. К настоящему времени разработана богатая элементная база интегральной оптики: разветвители и сумматоры, спектральные, модовые и поляризационные фильтры, модуляторы, коммутаторы, дефлекторы и др. На ее основе реализуются монолитные ОИС, выполняющие передачу и прием сигналов со спектральным уплотнением, анализ спектров сигналов, реализующие вычислительные функции и аппаратную оптическую обработку информации.

Следует отметить тесную взаимосвязь волоконной и интегральной оптики, обусловленную одинаковым характером передачи оптического излучения. В данном курсе буду рассматриваться гибридные элементы, например соединения волокно-волновод и др.

Преимущества интегральной оптики

К основным преимуществам волоконно-оптических и интегрально-оптических элементов относят:

- невосприимчивость к электромагнитным помехам (ЭМП);
- отсутствие риска «короткого» замыкания или замыкания на землю;
- безопасность работы в окружении горючих веществ (топлива);
- безопасность, связанная с радиоперехватом или подслушиванием;
- малые потери при передаче;
- большая пропускная способность (возможность мультиплексирования);
- малые размеры и вес;
- малая стоимость, использование доступных материалов.

В таких областях применения, как бортовая радиолокация, управление ракетами, обнаружение неисправностей в высоковольтных линиях, а также

многоканальная связь, крайне важно, чтобы система продолжала нормально функционировать, в присутствии сильных электромагнитных помех. Конечно, металлические провода можно экранировать, как это имеет место в случае коаксиальных кабелей, однако металлический экран увеличивает вес аппаратуры, ее стоимость и приводит к появлению паразитных емкостных связей, ограничивающих частотные характеристики систем и ширину полосы пропускания. Волоконным световодам свойственна невосприимчивость к большинству видов ЭМП, вследствие отсутствия металлических проводов, в которых за счет электромагнитного взаимодействия с радиопомехой мог бы индуцироваться ток. Кроме того, сравнительно легко исключить нежелательную подсветку, покрывая волокно (или узел из нескольких волокон) непрозрачной оболочкой. Перекрестные помехи, т. е. нежелательное взаимодействие между сигналами, распространяющимися по соседним оптическим волокнам в волоконном кабеле, также минимальны, поскольку каждая светопроводящая жила волоконного световода окружена сравнительно толстой оболочкой, через которую не могут проникать поля распространяющихся оптических волн. Во многих областях применения оптической аппаратуры невосприимчивость распространяющихся оптических волн к электромагнитным помехам сама по себе является весьма существенной причиной предпочтительного использования волоконно-оптических соединений вместо проводов или кабелей. Однако, как будет видно из дальнейшего, имеется еще целый ряд преимуществ, которые делают еще более выгодным использование волоконно-оптических соединений. В отличие от металлических проводов по оптическим волокнам ток не течет; следовательно, «короткие» замыкания не возникают. Это дает возможность связывать волокна в плотные узлы и пропускать их по металлическим трубопроводам, не заботясь об электрической изоляции. Волоконно-оптические соединения особенно полезны в тех случаях, когда имеют дело с высоким напряжением. В подобных случаях такие свойства волоконно-оптических соединений исключают необходимость в использовании дорогостоящих изолирующих

трансформаторов. Поскольку не возникает искр при повреждении световодов, то их можно использовать во взрывоопасных средах.

В военной и других (особо важных) областях применения волоконно-оптические соединения обеспечивают большую невосприимчивость к перехвату или контролю извне по сравнению с электрическими проводами или линиями радиосвязи. Это обусловлено отсутствием электромагнитного излучения, распространяющегося за пределы оптического волокна, сигнал которого можно перехватить за счет наведения электромагнитной индукции в петле или антенне. Чтобы перехватить сигнал из волоконного световода, необходимо каким-то образом повредить светозащитную оболочку волокна. Это очень трудно сделать без нарушения светопередачи по данному волокну до такой степени, чтобы этого нельзя было заметить, что в свою очередь позволяет быстро принять надлежащие контрмеры.

Очевидно, самым важным преимуществом волоконных световодов является то, что их можно использовать для передачи оптических сигналов с малыми потерями на расстояния до 100 км без регенерации сигнала, и до тысяч километров – с регенерацией. Потери в серийно выпускаемых кварцевых волокнах составляют 0,2 – 0,3 дБ/км. Потери в оптическом волокне сравнительно слабо зависят от частоты, в то время как у других конкурирующих средств передачи информационных сигналов они резко возрастают по мере увеличения частоты. Например, для частотной зависимости потерь в скрученном двухпроводном кабеле, который обычно используют для соединения элементов авиационного электронного оборудования, видно, что потери существенно возрастают при частоте модуляции выше 100 кГц. Коаксиальные кабели применяют для передачи сигналов на сравнительно небольшие расстояния на частотах вплоть до 100 МГц, даже если потери велики, однако, выше указанной частоты потери становятся неприемлимо большими. Для сравнения отметим, что затухание в волокне незначительно даже на частотах вплоть до 10 ГГц. Максимальная частота, на которой можно использовать оптическое волокно для передачи сигналов, ограничивается

скорее не самим затуханием, а косвенно явлением дисперсии, т.е. зависимости времени распространения света от длины волны или модового состава.

В большинстве областей использования интегральной оптики такая большая полоса пропускания волоконных световодов будет использоваться не для передачи одного единственного столь высокочастотного сигнала, а скорее для мультиплексирования большого числа сигналов с меньшей шириной на одной и той же несущей оптической волне. Эта возможность мультиплексирования, предоставляемая волоконными световодами, в сочетании с тем, что их диаметр обычно в сотни раз меньше,, чем у коаксиального кабеля, означает, что количество информационных каналов на единицу площади поперечного сечения при использовании волоконных световодов по порядку величины в 10^4 раз больше, чем при использовании токов и радиоволн.

При рассмотрении различных преимуществ волоконных световодов мы уже отмечали, что по ним могут передаваться широкополосные сигналы порядка 10 ГГц. Однако недостаточно иметь возможность передачи таких сигналов, если нет возможности их генерации и обработки. Это привело к созданию оптических интегральных схем (ОИС).

Оптические интегральные схемы (ОИС) имеют целый ряд преимуществ перед любым другим элементом, например, перед электрической интегральной схемой:

- увеличенная ширина полосы;
- большие возможности разделения по частоте (длине волны) при мультиплексировании;
- использование элементов связи с малыми потерями, включая шины для выбора информации;
- большие возможности при переключениях (количество позиций, скорость переключения);
- меньшие размеры, вес, потребляемая мощность
- повышенная надежность.

Оптические интегральные схемы по существу обладают такой же широкой характеристической полосой пропускания, как и волоконные световоды, поскольку в обоих случаях средством; переноса информации является световая волна, а не электрический ток. Поэтому можно пренебречь емкостными и индукционными эффектами, ограничивающими значение частоты. При сравнении с расширенными оптическими системами, состоящими из дискретных элементов, следует ожидать, что ОИС имеют перед ними такие же преимущества, которые имеют электрические интегральные схемы перед дискретными системами, собранными вручную. К ним относятся меньшие размеры, вес, потребляемая мощность, а также повышенная надежность и экономия при серийном производстве. Кроме того, оптическая юстировка и чувствительность к вибрациям, которые представляют собой сложные проблемы для оптических систем, состоящих из дискретных элементов, достаточно контролируются в ОИС автоматизировано.

Материалы, используемые для подложки оптических интегральных схем:

Выбор материала для подложки, на основе которой изготовляется оптическая интегральная схема, очень сильно зависит от функций, выполняемых ОИС. В большинстве случаев ОИС может включать в себя несколько различных оптических элементов, например источник излучения, модуляторы и детекторы, и ни один из материалов для подложки не будет оптимальным для их изготовления. Поэтому необходимо найти некое компромиссное решение. Первый шаг состоит в том, чтобы решить, какой предпочесть подход (гибридный или монолитный), как это рассматривается ниже.

Имеются два основных вида оптических интегральных микросхем. Одним из них являются гибридные микросхемы, в которых два или более материала каким-то способом соединяются вместе для оптимизации характеристик различных устройств. К другому виду относятся монолитные

ОИС, в которых для всех элементов используется один материал. Так как многие ОИС потребуют использования источника света, то монолитные микросхемы можно изготовлять лишь из оптически активных материалов, например из полупроводниковых материалов: арсенид галлия, арсенид-алюминий галлия, фосфид-арсенид галлия, арсенид галлия-индия и др. полупроводники III-V и II – IV групп с прямой запрещенной зоной.

Пассивные материалы, например, кварц, танталат лития, ниобат лития, пятиокись тантала, пятиокись ниобия, также можно использовать в качестве материала для подложки, но с внешним источником света (полупроводниковым лазером), который каким-то образом оптически и механически должен соединяться с подложкой.

Основным преимуществом гибридного подхода является то, что ОИС можно изготовлять с использованием существующей технологии и объединением элементов, которые в основном оптимизированы для данного материала. Например, одной из самых ранних ОИС, выполнявшей функцию сложной системы, был анализатор спектра радиочастоты, который состоял из серийно выпускаемого лазерного диода на основе GaAlAs и кремниевой фотодиодной линейки детекторов с акустооптическим модулятором на подложке из ниобата лития. Гибридное стыковое соединение было использовано для эффективной связи как лазера, так и фотодиодной линейки с подложкой из $LiNbO_3$. В этом случае гибридный подход сделал возможным комбинирование хорошо разработанных GaAlAs-лазеров на гетероструктурах, акустооптических модуляторов на $LiNbO_3$ и кремниевых фотодидных матриц.

Хотя гибридный подход при изготовлении ОИС представляет собой удобный метод получения схем с широким спектром выполняемых операций, он обладает тем недостатком, что места связи различных элементов схемы

могут приводить к рассогласованию и даже ее отказу из-за вибраций и тепловых деформаций. Кроме того, использование подхода, основанного на изготовлении монолитных ОИС, в конце концов обходится дешевле при массовом производстве, так как можно использовать автоматизированное поточное производство.

В настоящее время разработаны следующие элементы ОИС:

Элемент	Возможность согласования
Лазер	+
Детектор	+
Элемент связи	+
Переключатель	+
Модулятор	+
Волноводный фильтр	+
Усилитель	+
Линза	+
Акустический преобразователь	+

Примеры интегрально-оптических схем

Многие из монолитных интегрально-оптических устройств имеют эксплуатационные характеристики, сравнимые или даже лучшие по сравнению с аналогами на основе дискретных элементов. Достигнутый относительно невысокий уровень интеграции, по-видимому, связан скорее с инерцией разработчиков, чем с принципиальными техническими ограничениями. Некоторые специалисты предпочитали разрабатывать лазеры на гетеропереходах, в то время как другие свое внимание сосредоточивали на волноводах или детекторах. Подобная специализация была выгодна на ранней стадии развития этой области, поскольку приводила к изобретению и разработке сравнительно сложных элементов, входящих в состав ОИС различного назначения. Такие приборы, как лазеры на гетеропереходах с распределенной обратной связью, акустооптические волноводные модуляторы и переключаемые электронооптические направленные ответвители, являются примерами надлежащим образом проработанной технологии. Это должно привести к разработке нескольких базовых монолитных ОИС, объединяющих лазеры, модуляторы или переключатели, а также детекторы. Такая тенденция, без сомнения, будет продолжаться и в дальнейшем должна привести к значительно большей интеграции. Примеры существующих ОИС приведены на рисунках 2 - 7.

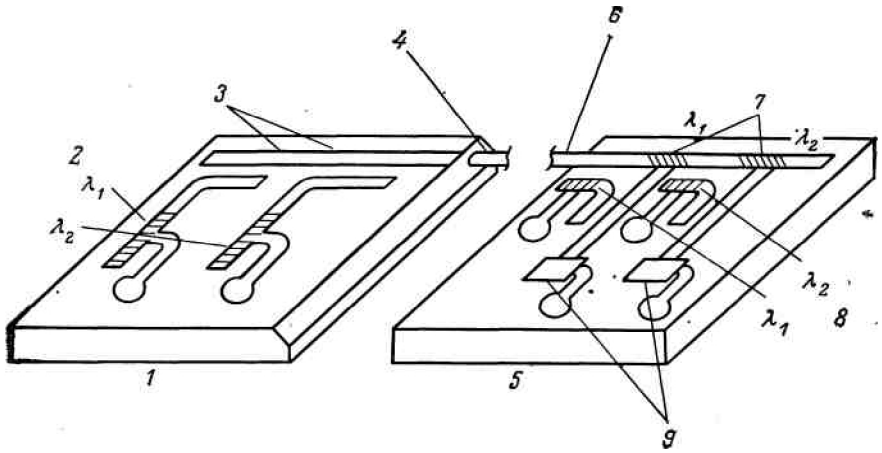


Рисунок 2 – Монолитные приемная и передающая ОИС для оптической связи: 1 - подложка передатчика; 2 - излучатели с распределенной обратной связью; 3 -

пассивные направленные ответвители; 4 - элемент связи пленка - оптическое волокно; 5 подложка чип приемника; 6 - элемент связи оптическое волокно - пленка; 7 - селективные частотные фильтры; 8 — местные излучатели с распределенной обратной связью; 9 - интегральные детекторы.

В качестве источника излучения используются интегральные лазеры с распределенной обратной связью, работающие на различных длинах волн. Для простоты показано 2 таких лазера, однаков в реальных схемах их может быть несколько десятков. Излучение различных лазеров с помощью направленных ответвителей поступает в общий волновод, а затем в волоконную линию связи, где распространяется без смешивания информации. В приемном модуля излучение разделяется по длинам волн и поступает к индивидуальному детектору. Для гетеродинного детектирования иногда используют дополнительные «местные» излучатели.

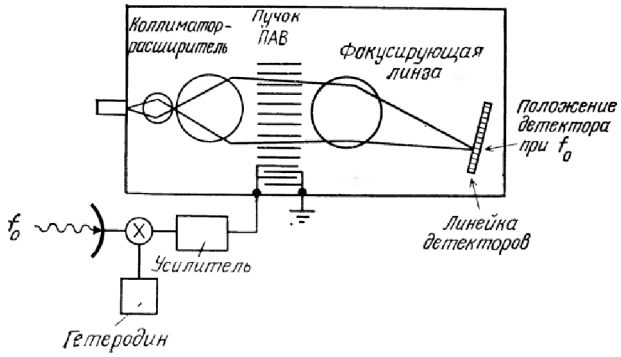


Рисунок 3 – Схема ИОС ВЧ спектроанализатора

Свет от лазерного источника вводится в планарный волновод, в котором он проходит сначала через кол-лимирующую линзу, а затем через акустооптический модулятор брэгговского типа. ВЧ-сигнал, спектр которого необходимо про-анализировать, подается на акустический преобразователь, который генерирует звуковые волны с изменяемым периодом. Таким образом, угол отклонения

оптического луча на выходе модулятора зависит от ВЧ-сигнала. Для фокусировки оптического луча на линейку фотодетекторов используется вторая линза. Если в ВЧ-сигнале присутствует более одной частотной компоненты, луч света разделяется на соответствующие компоненты, которые фокусируются на различные детекторные элементы. Каждый детекторный элемент представляет собой определенный частотный канал, и, поскольку фотодиоды обычно имеют квадратичные характеристики, выходной сигнал с каждого канала пропорционален ВЧ-мощности на данной частоте. Преимущество интегрально-оптического спектроанализатора по сравнению с электронным состоит в том, что требуется лишь несколько оптических элементов для выполнения тех функций, которые в противном случае требовали бы тысячи электронных элементов.

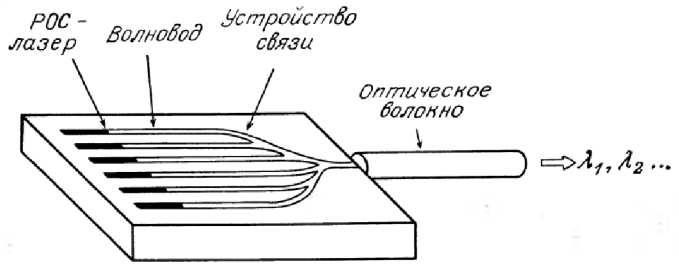


Рисунок 4 – Передающий оптический модуль со спектральным уплотнением

Передающий модуль объединяет несколько интегральных лазеров, снабженных спектрально-избирательными элементами. Излучение с различными длинами волн вводится в интегральный, а затем в волоконный световод. На каждой длине волны передается своя информация.

Интегрально-оптические корреляторы (ИО-корре-ляторы) электрических и оптических сигналов могут использоваться для реализации операций типа свертки. В настоящее время существует несколько разновидностей ИО-корреляторов цифровых и аналоговых сигналов. В основе их работы используется дифракция когерентных пучков оптического излучения в планарных ОВ на дифракционных решетках (индуцированных с помощью ПАВ и/или электрооптического эффекта),

последующее фурье-преобразование пространственно модулированного оптического сигнала посредством планарных волноводных или внешних линз, а также выделение пространственным фильтром и фотоприемником (или матрицы фотоприемников) корреляционного сигнала или сигнала свертки двух входных сигналов *S1(t)* и *S2(t)*.

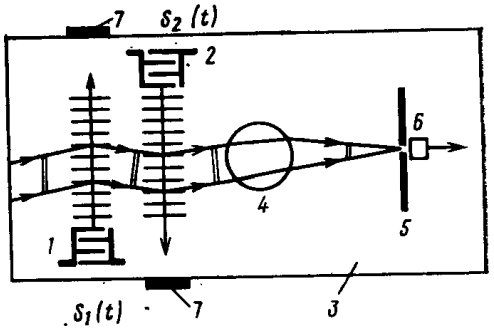


Рисунок 5 – Акустооптический пространственно-интегрирующий коррелятор аналоговых сигналов: 1, 2 – встречно-штыревые преобразователи; 3 – планарный волновод; 4 – линза; 5 – фильтр пространственных частот; 5 – фотодетектор; 6 – поглотитель поверхностной акустической волны.

Из множества разработанных ИО-корреляторов аналоговых сигналов можно выделить два основных типа: пространственно-интегрирующие и время-интегрирующие корреляторы. Схемы быстродействующих ИО-корреляторов цифровых сигналов представлены на рисунках 6 и 7. Для выполнения интегральной операции корреляции или свертки необходимо обеспечить запоминание фильтрующей функции (эталонного кода) — сигнала *S2(t)*, прием изменяющегося во времени информационного сигнала S1*(t)*, который складывается из полезного сигнала и шумов, а также умножение и интегрирование сигналов.

В быстродействующем ИО-корреляторе цифровых сигналов (рисунок 6) эталонный код формируется с помощью соответствующих электронных устройств электрооптического пространственного модулятора света. Входной информационный сигнал *S1(t)* вызывает в волноводе генерацию радиоимпульсов

ПАВ, распространяющихся на участке, длина которого (произведение длительности импульса на скорость ПАВ) согласована с длиной элемента электрооптического пространственного модулятора света. Логическая единица формируется высокочастотной ПАВ (преобразователь *2* с центральной частотой f_02 = 875 МГц), а нуль (или минус единица, если сигнал биполярный) - низкочастотной ПАВ (преобразователь / с центральной частотой f_0i =459 МГц). Умножение сигналов осуществляется в результате двукратной дифракции параллельного светового пучка в планарном ОВ, а интегрирование - с помощью оптической линзы.

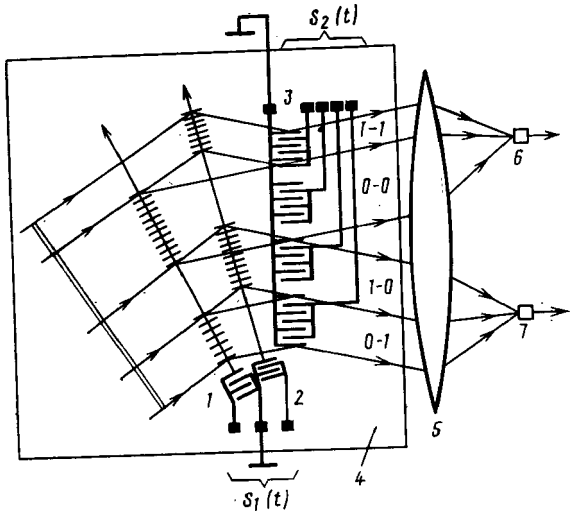


Рисунок 6 - Схема быстродействующего ИО-коррелятора цифровых сигналов: 1, 2 - встречио-штыревые преобразователи; 3 - электрооптический пространственный модулятор света; 4 - планарный ОВ; 5 - линза; 6, 7 – фотодетекторы

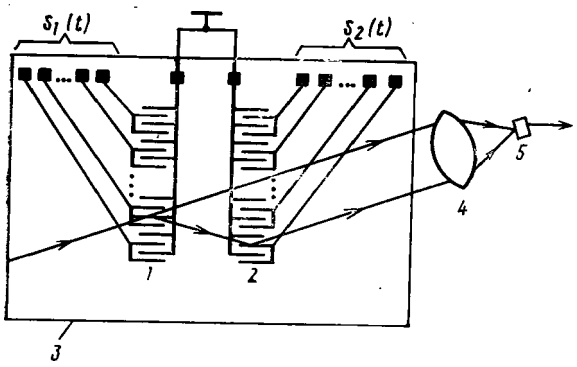


Рисунок 7 - Схема электрооптического брэгговского ИО-коррелятора цифровых сигналов:

1, 2 — система встречно» штыревых сигнальных и кодирующих электродов соответственно: 3 — планарный ОВ; 4 — линза; 5 — фотодетектор

Лекция 2

Полупроводниковые источники излучения. Рекомбинационное излучение и поглощение. Энергетические уровни. Квантовый выход. Спонтанная и лазерная люминесценция в полупроводниковых структурах. Когерентность, монохроматичность, поляризация излучения.

Основные принципы излучения света в полупроводниках

При рассмотрении взаимодействия света с электронами полупроводников удобной для использования моделью является модель микроскопического момента (в противоположность макроскопической) в кристалле, в которой электроны рассматриваются как имеющие в кристалле моменты, устанавливаемые соотношением

$$p = \frac{hk}{2\pi}, \qquad (1)$$

где k - волновой вектор электронного состояния (постоянная Планка h = 6,624*10^{-27} эрг*с). Вектор р не является классическим механическим моментом свободного электрона mv, где т — масса, а v - скорость. Кристаллический момент р включает в себя воздействие на электроны атомов кристаллической решетки и связан с энергией и скоростью электрона довольно сложным образом. В уравнениях движения электрона в кристалле величина р играет роль момента, а эффективная масса m* — роль массы по аналогии с классическими уравнениями движения свободного электрона в вакууме. Кристаллический момент и эффективную массу можно определить для дырок в валентной зоне так же, как это делается для электронов в зоне проводимости. Однако в общем случае величины р и m* для электронов будут отличными от аналогичных величин для дырок в данном материале.

Взаимодействие электронов с дырками в полупроводнике может приводить к фотонам света. Эти дискретные кванты или относительно локализованные единицы оптической энергии сохраняют свою тождественность в качестве таковых во всех процессах эмиссии, передачи, отражения, дифракции, поглощения и т. д. Фотон описывается в микроскопической модели путем установления его массы и момента следующим образом. Первоначально рассмотрим фотон света в вакууме. Энергия E фотона дается уравнением:

$$E = h\nu, \qquad (2)$$

где ν - частота излучения ($с^{-1}$). Если выражение (2) скомбинировать с известным соотношением:

$$c = \nu\lambda_0, \qquad (3)$$

где с - скорость света в вакууме, λ_0 - длина волны света в вакууме, то энергию фотона можно выразить как:

$$E = \frac{hc}{\lambda_0}, \qquad (4)$$

Часто удобно помнить, что выражение (4) для энергии принимает вид:

$$E = \frac{1{,}24}{\lambda_0}, \qquad (5)$$

где в обычно используемой системе единиц E выражается в электронвольтах (эВ), а длина волны - в микрометрах (мкм).

Таким образом, с помощью выражений (1) - (5) можно описать эквивалентно с помощью энергии, длины волны, частоты, массы или момента. Заметим, что в случае первых четырех параметров полное описание фотона требует указания направления движения. Следует помнить, что также важна фаза фотона в случае рассмотрения когерентного излучения. Уравнения (1) - (5) были выведены для фотона в вакууме. Однако, чтобы воспользоваться ими для описания фотона в твердом теле, достаточно заменить величины c и λ_0 на скорость v и длину волны λ света в материале соответственно. Взаимодействие фотонов излучения с электронами и дырками в полупроводнике можно рассматривать как взаимодействие частиц, используя определения предыдущих разделов. В этом случае оптические процессы регулируются обычными законами сохранения энергии и импульса.

Законы сохранения энергии и момента (импульса)

Законы сохранения энергии и момента (импульса) удобно иллюстрировать путем рассмотрения явления поглощения света. Наибольшее поглощение в полупроводниках происходит за счет возбуждения электронов, когда электрон, находящийся на данном энергетическом уровне, поглощает фотон и переходит в более высокое энергетическое состояние. Электронные переходы подчиняются некоторым правилам отбора, большинство из которых базируется на сохранении энергии и импульса электрона и фотона. (С теоретической точки зрения - это сохранение волнового вектора k, являющееся следствием периодичности кристаллической структуры, которая приводит к сохранению момента, определяемого как $hk/2\pi$.) Таким образом, если принимаются во внимание только взаимодействие фотонов и электронов, то:

$$E_i + h\nu_{фот} = E_f, \quad (6)$$

$$p_i + p_{фот} = p_f$$

где индексы i и f относятся к начальному и конечному состояниям электрона. Момент фотона в видимой или инфракрасной области спектра во много раз меньше момента электрона при тепловом возбуждении. Следовательно, селективным правилом

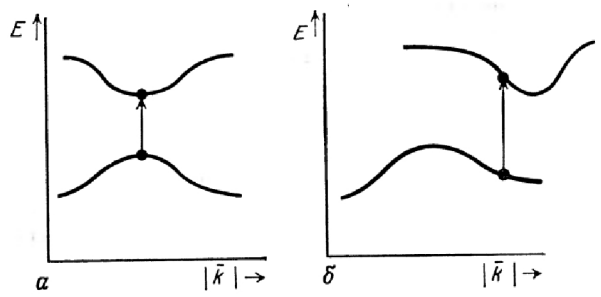


Рисунок 1 - Прямые переходы электронов с поглощением в материалах с прямой (а) и непрямой (б) запрещенными зонами.

отбора является приближенное выполнение равенства $k_i=k_f$, и электрон практически не испытывает никакого изменения момента. Такой тип перехода, включающий лишь электрон и фотон, называется прямым переходом. Этот переход показан на графике E-k (рисунок 1) как вертикальный переход. Заметим, что в общем случае минимумы зон проводимости не совпадают с максимумами валентных зон в одной точке k-пространства. Если совпадение имеет место, то говорят, что материал имеет прямую запрещенную зону. Если совпадения нет, то материал имеет непрямую запрещенную зону.

Хотя фотон не может передать значительную часть момента электрону, можно иметь электронные переходы, в которых изменяется k. В этих процессах участвует также фонон или, говоря другими словами, момент передается или поглощается атомами решетки. Колебательная энергия атомов решетки квантована (как световая энергия) на фононы с энергией hν, где ν - частота (1/с). Фононы характеризуются волновым вектором q. Условия, ограничивающие поглощающий переход, в котором участвуют фотон, электрон и фонон, определяются выражениями:

$$k_i \pm q = k_f ,$$

$$E_i + h\nu_{фот} \pm h\nu_{фон} = E_f , \quad (7)$$

где в выражении (7) волновым вектором $\mathbf{k} = (2\pi/\lambda_0)\mathbf{u}$ и фотона можно пренебречь. Схемы процесса на графике E - k приведены на рисунке 2.

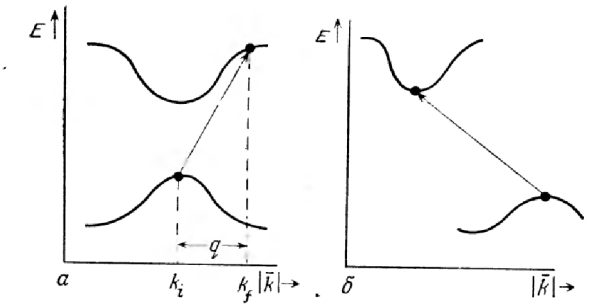


Рисунок 2 - Непрямые переходы электронов с поглощением в материале с прямой (а) и непрямой (б) запрещенными зонами

Электрон поглощает фотон и в то же самое время поглощает или испускает фонон. Такие переходы, называемые непрямыми, выглядят как диагональные переходы на графике E - k, представленном на рисунке 2. Так как для непрямых переходов необходимы как фотон, так и фонон, а для того, чтобы вызвать прямой переход, необходим только фотон, прямой переход является более вероятным. По этой причине полупроводники с прямой запрещенной зоной оптически более активны по сравнению с полупроводниками с непрямой запрещенной зоной. Эта особенность очень важна в случае источников света.

На рисунках 1, 2 представлены лишь межзонные переходы (между валентной зоной и зоной проводимости). Однако как прямые, так и непрямые переходы также имеют место внутри зоны (внутризонные) или между энергетическими состояниями, обусловленными атомами примесей и (или) дефектами. Во всех

случаях применимы принципы сохранения энергии и импульса (волнового вектора). Поглощающие переходы такого типа представлены на рисунке 3. Внутризонное поглощение может иметь место как для электронов в зоне проводимости, так и для дырок в валентной зоне, и поэтому оно называется поглощением на свободных носителях. Такое поглощение обычно включает в себя, переходы электронов из донорных состояний в зону проводимости и переходы дырок из акцепторных состояний в валентную зону. Так как графики соответствуют энергии для электронов, возрастание энергии для дырок соответствует движению по графикам вниз.

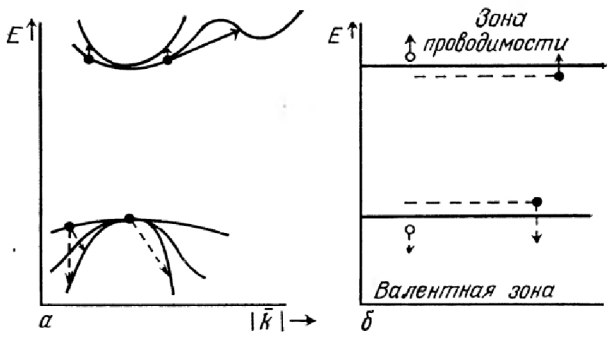


Рисунок 3 - Поглощение на свободных носителях

Действие законов сохранения энергии и импульса, которое проиллюстрировано на примере процессов оптического поглощения, также одинаково применимо к процессам излучения фотонов в полупроводниках. В действительности такое рассмотрение является даже более важным в случае излучения света, как это будет показано в следующем разделе.

Излучение света в полупроводниках

Излучение света в полупроводниках обычно происходит благодаря рекомбинации носителей. Процесс излучения можно разделить на два типа:

спонтанное излучение, в котором дырки и электроны рекомбинируют случайно, и стимулированное излучение, в котором дырки и электроны стимулируются к рекомбинации уже существующими фотонами.

Спонтанное излучение

При поглощении света электрон поглощает фотон и переходит в более высокое энергетическое состояние. При излучении происходит противоположный эффект. Электрон переходит из более высокого энергетического состояния в более низкое и при этом теряет энергию, которая излучается в виде фотона. Переходы с излучением света в полупроводниках являются межзонными переходами; они происходят между зоной проводимости и валентной зоной и (или) определенными состояниями внутри запрещенной зоны, обусловленными легирующими примесями или дефектами. Так как запрещенная зона полупроводников обычно составляет от нескольких десятых до нескольких единиц электронвольта, длина волны излучения обычно лежит в инфракрасной области спектра и приблизительно соответствует длине волны края полосы поглощения в полупроводнике. Как и в случае поглощения, существуют определенные правила отбора, ограничивающие возможные излу-чательные переходы. В частности, требуется, чтобы выполнялись законы сохранения энергии импульса. Таким образом, для прямых переходов имеют место соотношения

$$E_i - E_f = h\nu_{фот},$$
$$k_i - k_f = (2\pi / \lambda_{фот})u, \qquad (8)$$

где u - единичный вектор, определяющий лишь направление. Отметим, что для прямых переходов $2\pi/\lambda_{фот} << |k_i|$ и $|k_f|$, так что $|k_i| \approx |k_f|$, как и для прямого перехода с поглощением. Излучение фотона может также происходить при непрямом переходе,

причем в этом случае также требуется выполнение законов сохранения энергии и импульса:

$$k_i - k_f \pm q = (2\pi / \lambda_{фот})u, \quad (9)$$

где q - волновой вектор поглощаемого или испускаемого фотона. Как и в случае поглощения, необходимость участия фонона приводит к значительному уменьшению вероятности перехода и, следовательно, к генерации фотона. Третьим требованием для излучательного перехода является то, чтобы были заполнены верхние (исходные) энергетические уровни и свободны соответствующие нижние (конечные) энергетические уровни, причем разность в энергии между ними должна равняться энергии излучаемого фотона. Это правило выглядит очевидным и тривиальным, но играет малую роль в случае поглощения ввиду того, что поглощающий материал обычно находится в тепловом равновесии и, следовательно, всегда имеет заполненные нижние энергетические уровни и свободные верхние. (За исключением случая, когда энергия фотонов меньше межзонной. Тогда не существует возможного верхнего состояния для электронов и, следовательно, поглощение не происходит). В случае излучения упомянутое правило является более важным. Например, для характерных полупроводников при тепловом равновесии и комнатной температуре и ниже существует относительно малое число дырок в валентной зоне и электронов в зоне проводимости. В случае легированного полупроводника в зоне проводимости может быть много электронов, как, например, в материале я-типа, но при этом очень мало дырок в валентной зоне (вследствие постоянства произведения яр). Противоположная картина имеет место для полупроводников р-типа. Следовательно, излучательная рекомбинация, происходящая при тепловом равновесии, дает очень мало фотонов, которые перепоглощаются до того, как покинут кристалл. Для получения значительного светового излучения от полупроводника необходимо каким-либо образом вывести полупроводник из состояния теплового равновесия, чтобы создать больше электронов в зоне

проводимости и дырок в валентной зоне. После накачки до более высоких энергетических уровней электроны и дырки «усаживаются» на состояния, соответствующие краям их зон.

Увеличения концентрации дырок и электронов можно достичь разными путями, и этот процесс обычно называется накачкой материала. Например, если свет большой интенсивности с энергией фотонов, большей энергетической ширины запрещенной зоны, направлен на полупроводник, то происходит межзонное поглощение, в результате чего будет возникать большое количество электронно-дырочных пар. Эти дырки и электроны первоначально горячие, но очень быстро термолизуются, т. ё, электроны садятся на дно зоны проводимости, а дырки поднимаются к верхней границе валентной зоны через взаимодействие с решеткой. Переход в тепловое равновесие (термолизация) обычно происходит очень быстро - за время порядка 10^{-14} с. После его установления электроны и дырки рекомбинируют; эта рекомбинация может быть излучательной, приводящей к излучению фотона с энергией, приблизительно равной ширине запрещенной зоны. Время жизни, или время, в среднем необходимое для рекомбинации электронно-дырочных пар после их возникновения, обычно составляет около 10^{-11} с в накачиваемом полупроводнике (с прямой запрещенной зоной). Заметим, что в полупроводниках с непрямой запрещенной зоной переход в тепловое равновесие вынуждает электроны и дырки занимать состояния со значительно различающимися волновыми векторами. Это означает, что, для того чтобы произошла рекомбинация, необходим фонон, так же как и дырка и электрон, а следовательно, это менее вероятный процесс. Типичное время жизни для непрямой рекомбинация достигает 0,25 с. Прежде чем он произойдет, дырка и электрон в среднем уже рекомбинируют в результате безызлучательного процесса с передачей энергии решетке или дефектам и т. д. Следовательно, эффективность непрямой излучательной рекомбинации очень мала. Квантовая эффективность для материала с прямой запрещенной зоной может достигать 1, в то время как для материала с непрямой зоной он обычно составляет 0,001 или менее. Однако непрямые излучательные переходы имеют место, и в некоторых материалах с непрямой за-

прещенной зоной (например, GaP) может происходить вполне измеримое и действительно полезное излучение света. Тем не менее чрезвычайно трудно изготовить лазеры на материалах «с непрямым переходом.

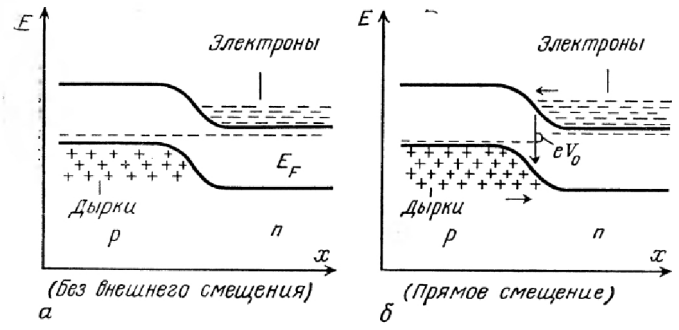


Рисунок 4 - Расположение энергетических зон p-n перехода светодиода. а) в отсутствие внешнего напряжения смещения; б) внешнее напряжение смещения V_0 приложено в прямом направлении

Метод оптической накачки, который был описан как способ получения необходимого увеличения концентрации электронов ш дырок, является сравнительно громоздким, так как требует •очень интенсивных источников света; этот процесс малоэффективен, и большая часть энергии превращается в тепло. Кроме того, выходящее излучение необходимо фильтровать, чтобы выделить излучение полупроводника из отраженного света накачки

Другой способ накачки, который одновременно эффективен и очень прост, использует свойства p-n перехода. Расположение энергетических зон светоизлучающего p-n перехода представлено на рисунке 4. В отсутствие напряжения смещения существует много электронов в зоне проводимости на n-стороне и много дырок в валентной зоне на p-стороне, но лишь некоторые из них могут преодолеть барьер и попасть в область перехода. После приложения прямого напряжения смещения V_0 барьер понижается и многие дырки и электроны

проникают (инжектируют) в область перехода, где они рекомбинируют, генерируя фотоны. Свет, излучаемый под воздействием электрического напряжения, называется электролюминесценцией, в та время как свет, получаемый путем накачки оптическим источником с более короткой длиной волны излучения, называется фотолюминесценцией. Типичная схема электролюминесцентного диода, или светоизлучающего диода (СИД), представлена на рисунке 5.

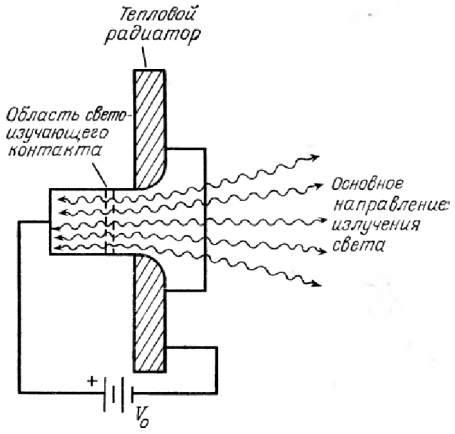


Рисунок 5 - Светоизлучающий диод (СИД) на p-n переходе

Свет генерирует в p-n переходе и покидает диод, проходя через объемный материал вне области перехода. Большая часть света перепоглощается в полупроводнике до того, как он покинет диод. Следовательно, можно говорить как о внешнем, так и о внутреннем квантовых выходах, определяемых следующим образом:

$$\eta_{внешн} = \frac{число\,фотонов,\ \ излученных\ \ в\ \ желаемом\ \ направлении}{число\ \ \ электронно-дырочных\ \ \ пар}, \quad (10)$$

$$\eta_{внутр} = \frac{число\ \ \ генерируемых\ \ \ фотонов}{число\ \ \ электронно-дырочных\ \ \ пар},$$

При разработке светоизлучающих диодов обычно стремятся минимизировать перепоглощение света путем изготовления очень тонкого слоя материала между переходом и поверхностью, а также (если возможно) путем подбора материала, имеющего очень малое поглощение. Такие светодиоды получили широкое распространение в панельных индикаторах и дисплеях вследствие того, что они очень малы, дешевы, более надежны, чем устройства на лампах накаливания, и легко объединяются в комплекты интегральных схем. СИД также очень часто используются в качестве источников света в волоконно-оптических линиях связи, так как их излучение можно модулировать на больших частотах (хотя и не таких высоких, как в лазерах), а сами СИД можно удовлетворительно соединять с сердцевинами многомодовых волокон.

Спектр длин волн полупроводниковых излучателей света независимо от используемого типа накачки является довольно простым по сравнению, например, со спектром газоразрядной дуговой лампы, которая имеет много линий или эмиссионных максимумов. Хороший полупроводниковый светоизлучающий диод с малым количеством дефектов или низким содержанием загрязняющих примесей имеет один эмиссионный максимум вблизи длины волны, соответствующей краю полосы поглощения, т. е. энергии запрещенной зоны. Этот максимум имеет полуширину обычно 20 - 30 нм, которая значительно больше ширины линий газового источника. Соответствующие загрязняющим примесям и дефектам состояния внутри запрещенной зоны порождают эмиссионные пики на других длинах волн. Эти пики показывают, что некоторые из инжектируемых электронно-дырочных пар рекомбинируют с излучением фотонов, не принадлежащих основному пику, соответствующему краю зоны. Следовательно, квантовая эффективность на желаемой длине волны уменьшается.

Даже если полупроводниковый излучатель сформирован из сравнительно чистых и бездефектных материалов, так что излучение почти точно приходится на длину волны, соответствующую краю полосы, то длина волны пика излучения все

же зависит от температуры и содержания легирующих примесей, поскольку ширина запрещенной зоны изменяется с изменением температуры, а легирование изменяет эффективную ширину запрещенной зоны, приводя к излучательным переходам между донорными и акцепторными состояниями. Таким образом, эмиссионную длину волны светоизлучающего полупроводника можно изменять в пределах некоторого ограниченного диапазона за счет соответствующего введения примесей. Для примера рассмотрим случай GaAs, легированного примесями гс-типа (Te, Se или Sn). Длина волны центра эмиссионного пика смещается, как показано на рисунке 6. При сравнительно низком уровне легирования имеет место рекомбинация между электронами, заполняющими донорные состояния на 6 мэВ ниже дна зоны проводимости, и незаполненными состояниями у верхней границы валентной зоны.

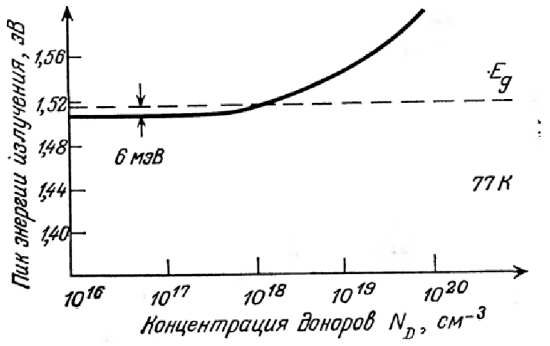


Рисунок 6 - Зависимость пика эмиссионной длины волны для GaAs n-типа от концентрации доноров N_D [10.6].

Когда уровень легирования поднимается выше $5*10^{17}$ см-3, донорная полоса сливается с зоной проводимости и рекомбинация происходит между заполненными состояниями зоны проводимости вблизи квазиуровня Ферми и незаполненными состояниями валентной зоны. Примеси р-типа (Cd и Zn) также вызывают сдвиг длины волны эмиссионного пика, как показано на рисунке 7. При низкой концентрации примесей р-типа рекомбинация происходит из заполненных состояний на дне зоны проводимости на незаполненные акцепторные состояния, расположенные на 30 мэВ выше границы валентной зоны. Если примеси

превышают уровень 10^{18} $см^{-3}$, то акцепторные состояния расширяются в полосу, как показано на рисунке 8.

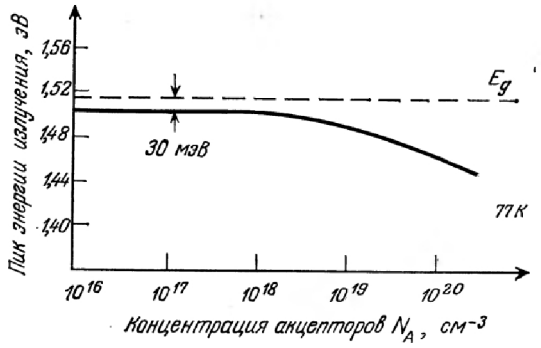


Рисунок 7 - Зависимость пиковой длины волны излучения для GaAs p-типа от концентрации акцепторов N_A

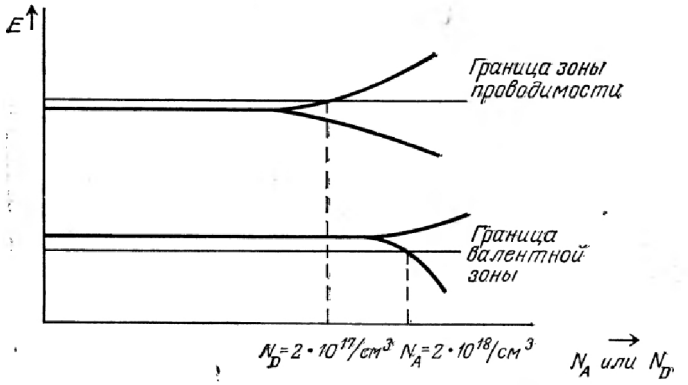


Рисунок 8 - Зависимость энергетических уровней (зон) доноров и акцепторов в GaAs от концентрации примесей

Рекомбинация происходит между заполненными состояниями на дне зоны проводимости и незаполненными состояниями в верхней границе акцепторной полосы. Таким образом, эмиссионный пик по энергии снижается при легировании полупроводника примесями p-типа в противоположность тому, что он увеличивается при легировании примесями n-типа. Если примеси p-типа возрастают до уровня выше $2*10^{18}$ $см^{-3}$, то акцепторная полоса сливается с валентной зоной и,

следовательно, излучательную рекомбинацию можно рассматривать как межзонную, хотя она на самом деле происходит между состояниями в зоне проводимости и акцепторными состояниями. Рекомбинация не происходит только между состояниями, имеющими точно одинаковые энергии в зоне проводимости и валентной зоне. На самом деле она происходит между состояниями, простирающимися по некоторому диапазону энергий. Следовательно, излучение линии порядка сотен ангстрем. По мере возрастания примесей n- или p-типа в GaAs ширина линии излучения обычно возрастает от 10 нм для легколегированного материала, до нескольких десятков нм для сильнолегированного материала. Это возрастание ширины линии излучения происходит за счет примесной «размытости» зон, которая делает возможным существование пар из заполненных и незаполненных состояний, способных рекомбинировать в широком диапазоне энергий.

Суммируя сказанное, отметим, что основными свойствами светового излучения, спонтанно генерируемого рекомбинирующими электронами и дырками в полупроводнике, являются следующие. Длины волн фотонов лежат в диапазоне, простирающемся на нескольких сотен ангстрем, с пиком на длине волны, соответствующей приблизительно ширине запрещенной зоны, но зависящей -каюке от концентрации и типа примесей. Фотоны излучаются более или менее изотропно, за исключением влияния геометрии излучателя, и не существует фиксированного фазового соотношения между фотонами. Излучение каждого фотона является спонтанным событием, совершенно не связанным с каким-либо другим фотоном. Характеристики излучаемого света, генерируемые стимулированной эмиссией, сильно отличаются от рассмотренного спонтанного излучения.

Стимулированное излучение

Спонтанно излученные фотоны могут, конечно, перепоглощаться за счет каких-либо процессов поглощения, которые были описаны ранее. В

противоположность процессу поглощения если фотон сталкивается с электронно-дырочной парой, имеющей подходящую энергию связи, то он может стимулировать ее рекомбинацию. В этом случае появляется новый фотон, который точно воспроизводит первоначальный, стимулировавший процесс рекомбинации, т. е. он имеет точно такую же энергию частоту, направление, фазу и поляризацию. Это стимулированное, или вынужденное, излучение, которое происходит в дополнение к спонтанному излучению. Таким образом, в общем случае интенсивность излучения, возникающего в присутствии внешнего излучения, состоит из двух частей. Одна часть независима от внешнего излучения (спонтанная эмиссия), другая имеет интенсивность, пропорциональную внешнему излучению, е то время как его частота, фаза, направление распространения и поляризация являются такими же, как и у внешнего излучения. Заметим, что внешнее излучение может быть излучением, порождаемым либо спонтанным, либо стимулированным излучением внутри самого образца, но не должно быть внешним нолем.

Исключая специфические условия, стимулированное излучение является менее вероятным, чем спонтанное, и, следовательно, обычный светоизлучатель дает главным образом спонтанное излучение. Вместе с тем необходимы специальные усилия, чтобы достичь значительной величины стимулированного излучения. Во-первых, концентрации дырок и электронов в области испускания света должны быть столь значительными, чтобы большинство состояний вблизи дна зоны проводимости было заполнено электронами, а большинство состояний около вершины валентной зоны было пусто. Это приводит к более высокой вероятности, что фотон вызовет переход электрона из зоны проводимости ш валентную зону и рождение фотона-близнеца, по

сравнению с вероятностью поглощения фотона при переходе электрона из валентной зоны в незаполненные состояния в зоне проводимости. Во-вторых, чтобы стимулированное излучение доминировало над спонтанным, должно существовать интенсивное поле фотонов соответствующей энергии.

При тепловом равновесии (т. е. в отсутствие любого источника накачки энергии) большинство электронов в материале будет находиться на нижнем энергетическом уровне. Концентрации электронов N_1 и N_2 на двух уровнях связаны следующим выражением в предположении распределения Больцмана:

$$N_2 / N_1 = e^{-(E_2 - E_1)/kT},$$

Следовательно, чтобы стимулированное излучение преобладало над поглощением, отношение N2/N1 в выражении должно быть сильно увеличенным за счет введения источника накачки энергии. Условие $N_2 >> N_1$ называется условием инвертированной населенности. Очевидно, что для создания стимулированного излучения, превышающего спонтанное, необходим интенсивный поток фотонов подходящей энергии. Необходимая концентрация фотонов обычно создается путем введения положительной обратной связи от зеркал или других отражателей.

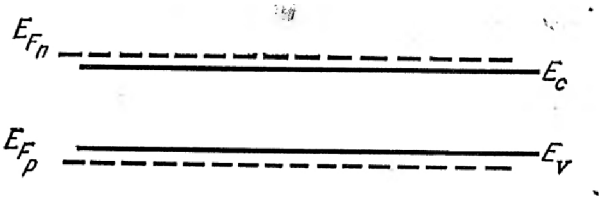


Рисунок 9 - Диаграмма энергетических зон полупроводника с внешней энергией на входе (E_{Fn} и E_{Fp}, представляют собой квазиуровни Ферми для электронов и дырок соответственно)

Два требования, заключающиеся в наличии инверсной населенности и интенсивного потока фотонов с подходящей энергией, являются общими для всех типов лазеров.

Упрощенная двухуровневая модель, которая использовалась для демонстрации этих принципов, гораздо лучше подходит для газовых лазеров, чем для полупроводниковых, поскольку электроны и дырки в полупроводниках распределены по энергетическим зонам, которые значительно шире дискретных энергетических уровней. Тем не менее действуют те же принципы. В полупроводниках, которые подвержены накачке, электроны заполняют состояния от дна зоны проводимости вверх приблизительно до энергетического уровня E_{Fn}, в то время как дырки заполняют состояния от вершины валентной зоны вниз приближенно до энергетического уровня E_{Fp}, как показано на рисунке 9. Энергетические уровни E_{Fn} и E_{Fp} известны как квазиуровни Ферми для электронов и дырок, так как, строго говоря, уровень Ферми определен только для случая теплового равновесия. Для неравновесного случая при подаче внешней энергии квазиуровни Ферми должны определяться отдельно для электронов и дырок. Из рисунка 9 видно, что состояние инверсии населенностей реализуется для любого фотона с энергией $h\nu_{фот}$, удовлетворяющей условию

$$E_{Fn} - E_{Fp} > h\nu_{фот}, \qquad (11)$$

Минимальная энергия, при которой существует инверсная населенность, определяется из выражения

$$E_{Fn} - E_{Fp} = E_c - E_v = E_g, \qquad (12)$$

где E_g - ширина запрещенной зоны. Таким образом, в накачиваемом полупроводнике инверсная населенность существует, и, следовательно, стимулированное излучение преобладает над поглощением фотонов с энергией, лежащей в диапазоне:

$$E_g \le h\nu_{фот} \le (E_{Fn} - E_{Fp}), \qquad (13)$$

Лазерный эффект

Если выполнены условия инверсии населенностей и существования высокой плотности фотонов, то возникает процесс селекции, который изменяет характер излучения и в результате которого появляется значительное количество стимулированного излучения. При наличии соответствующей оптически резонирующей структуры может иметь место лазерный эффект.

Полупроводниковые лазерные структуры

Обычно используемая структура для дискретных полупроводниковых лазеров представляет собой прямоугольный параллелепипед, аналогичный представленному на рисунке 10, который изготавливается путем скалывания кристалла вдоль плоскостей кристаллической решетки, которые перпендикулярны светоизлучающему слою. Например, для изготовления такой структуры пластина GaAs <100> может легко раскалываться по четырем перпендикулярным плоскостям <110>. Типичная длина L и ширина W составляют 300 и 50 мкм соответственно, однако возможны значительные отклонения от этих величиню. Две торцевые поверхности шириной W образуют оптически резонансную отражательную структуру, известную как эталон Фабри-Перо. При использовании такой структуры увеличение накачиваемой мощности будет приводить к лазерному эффекту, если превышен определенный критический пороговый уровень.

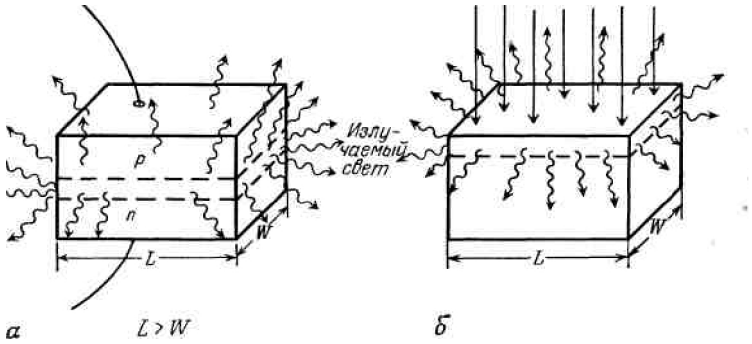


Рисунок 10 - Полупроводниковые лазерные структуры: а - накачиваемый через p-n переход лазерный диод; б - накачка электронами или фотонами. Штриховые линии ограничивают зону максимальной накачки.

Пороговый уровень для лазерного эффекта

По мере возрастания от нуля накачиваемой мощности увеличивается число фотонов, испускаемых за счет спонтанной эмиссии. Даже если мощность накачки превышает требуемую для создания инверсной населенности, излучение вначале преимущественно состоит из спонтанно образующихся фотонов. Спонтанные фотоны излучаются более или менее одинаково во всех направлениях, как показано на рисунке 10. Тем не менее, большинство этих фотонов начинает свой путь в направлениях, которые очень быстро выводят их из области с инверсной населенностью, где может происходить преимущественная стимулированная эмиссия, и, таким образом, они оказываются неспособными воспроизводить самих себя. Некоторые фотоны, случайно бегущие вдоль области с инверсией населенности, оказываются способными создавать себе подобные фотоны много раз за счет стимулированной эмиссии, прежде чем они покинут лазер. Кроме того, для любой данной энергии межзонной щели и распределения дырок и электронов будет существовать одна особая энергия (длина волны), которая является преимущественной (переходы более вероятны) по сравнению с другими. Эта длина волны или энергия обычно соответствует в первом приближении длине волны пика спонтанного излучения в материале. В результате этих преимуществ по энергии и направлению, когда возникает стимулированное излучение, испускаемое излучение сужается как по пространственной расходимости, так и по спектральной ширине полосы (уменьшающейся от нескольких сотен ангстрем приблизительно до 2,5 нм). Заметим, что после возникновения стимулированного излучения расходуется больше электронно-дырочных пар в пересчете на 1 с. Следовательно, для любой данной скорости генерации электронно-дырочных пар спонтанное излучение подавляется, поскольку стимулированное излучение расходует генерируемые пары прежде, чем они спонтанно рекомбинируют. Такое преобладающее стимулированное излучение, как оно было описано, известно как суперлюминесцентное излучение. Это еще не когерентное излучение света, производимое за счет лазерного эффекта. Такое излучение обусловлено усилением спонтанного излуче-

ния. Фотоны не движутся точно в одном направлении; они могут слегка отличаться по энергиям, но не синхронизованы по фазе. Иными словами, излучение еще некогерентно.

В том случае, когда светоизлучающая область ограничена параллельными плоскостями, как показано на рисунке 10, будет иметь место некоторое отражение от поверхностей и, кроме того, часть излучения будет проходить назад и снова вперед через область с инверсной населенностью и еще более усиливаться. Чтобы отразиться назад и затем вперед в узком слое инверсной населенности, фотоны должны быть строго в плоскости этого слоя и двигаться в направлении, строго перпендикулярном поверхностям Фабри-Перо. Излучение также должно быть однородным по частоте и фазе, чтобы избежать деструктивной интерференции, так как отражатели Фабри-Перо формируют резонаторную полость, в которой устанавливаются оптические моды. Таким образом, только та малая часть суперлюминесцентного излучения, которая удовлетворяет всем этим требованиям, продолжает существовать и становится доминирующей разновидностью излучения, выходящего из лазера. Такое излучение, однородное по длине волны, фазе, направлению и поляризации, является когерентным. Под поляризацией понимается направление колебаний электрической составляющей световой волны, детектируемое с помощью двулучепреломляющих материалов и фотоприемников. В этом случае говорят, что устройство обладает лазерным эффектом.

Переход от суперлюминесценции к лазерному эффекту (по мере нарастания входной мощности) принимается как точка, при которой усиление преимущественной моды за счет стимулированной эмиссии становится больше поглощения и всех других потерь фотонов, так что излучение при отражении проходит туда и обратно между поверхностями Фабри-Перо, возрастая по интенсивности при каждом проходе (т. е. устанавливаются осцилляции). Конечно, излучение не будет нарастать по интенсивности до бесконечности. По мере роста оно вызывает больше рекомбинаций до тех пор, пока концентрации дырок и электронов не будут использоваться с той же скоростью, с какой они образуются при данном источнике накачки, т. е. пока не установится динамическое равновесие. В этом стационарном состоянии лазер генерирует на одной пиковой длине волны (или, возможно, нескольких дискретных длинах волн, соответствующих различным продольным модам), излучение его имеет высокую монохроматичность (приблизительно ширину полосы 0,1

нм); внутри кристалла возникает стоячая волна с целым числом полуволн, укладывающихся между поверхностями Фабри-Перо для каждой моды. Вследствие описанной ранее селективности по направлениям отражающих торцевых поверхностей фотоны лазерной моды излучаются преимущественно в направлении, перпендикулярном отражающим поверхностям, как показано на рисунке 11.

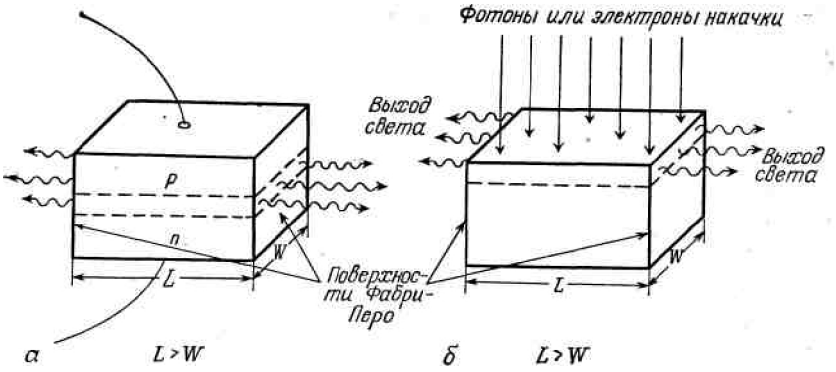


Рисунок 11 - Полупроводниковые лазеры, работающие при накачке, превышаю щей порог генерации: а - накачка током в p-n переходе; б - накачка электронами или фотонами.

Световые волны (фотоны), излучаемые из такой резонансной структуры, находятся в фазе, и, кроме того, волны, которые могут наиболее легко распространяться, являются волнами с электрическим вектором, перпендикулярным плоскости слоя инверсной населенности. Следовательно, излучаемый свет в этом случае является нормально поляризованным.

Переходы между областями спонтанной, суперлюминесцентной и лазерной эмиссии по мере возрастания накачиваемой мощности происходят очень резко. Действительно, хотя переход между суперлюминесценцией и лазерным излучением был описан в деталях с педагогической целью, реальный переход в обычном лазере настолько быстрый, что суперлюминесцентное излучение не наблюдается. Когда накачиваемая мощность превышает пороговую точку, характер светового излучения резко изменяется и приобретает лазерный тип генерации с убыванием полосы излучения от нескольких десятков нм до 0,1 нм и с наблюдаемой когерентностью фазы. Промежуточная стадия «суженной» полосы излучения, возникающая за счет преимущественно стимулированного излучения без когерентности по фазе (т. е. суперлюминесцентная), проходит настолько быстро, что ее невозможно наблюдать. Только в лазерных диодных структурах с

поврежденными отражателями Фабри-Перо, которые сдерживают лазерный эффект, можно наблюдать суперлюминесцентное излучение.

Эффективность излучения света

Одной из важнейших характеристик светового излучения полупроводниковых лазеров, которая еще не обсуждалась, является эффективность преобразования электрической энергии в оптическую. Лазерные диоды, работающие выше порогового уровня, имеют приблизительно в 100 раз большую эффективность, чем светоизлучающий диод, изготовленный из того же материала. Такое улучшение эффективности происходит за счет нескольких факторов.

Как ранее упоминалось, генерация фотонов с помощью стимулированного излучения имеет сравнительно низкую вероятность возникновения по сравнению с вероятностью спонтанного излучения в случае неинвертированнои населенности, как это происходит в светоизлучающем диоде (СИД). Однако, когда есть инверсия населенностей, как в лазере, вероятность стимулированного излучения становится очень большой, даже больше вероятностей спонтанного излучения и безызлучательных рекомбинаций электронно-дырочных пар. Например, внутренняя квантовая эффективность в лазерном диоде примерна в 10 раз выше, чем в СИД. Фактически, если лазерный диод охлаждается до 77К, для снижения температурного разброса электронов и дырок по энергиям его внутренний квантовый выход может приблизиться к 100%.

Если фотоны образовались, то они также имеют меньшие потери в лазерах, чем в СИД. Фотоны в лазерных модах проходят главным образом в области инверсионной населенности, в которой внутризонное поглощение подавляется за счет недостатка электронов вблизи края валентной зоны и незаполненных состояний вблизи края зоны проводимости. Фотоны в СИД двигаются главным образом в безынверсном объеме материала. Таким образом, внутреннее перепоглощение испускаемых фотонов в лазере намного меньше, чем в СИД. Кроме того, фотоны, излучаемые из лазеров, образуют пучок света с много меньшей расходимостью, чем фотоны, излучаемые из СИД, что позволяет собирать (или вводить) больше света в заданном месте. Вследствие совместного эффекта

снижения внутреннего перепоглощения высокой коллимации пучка и возрастания внутреннего квантового выхода можно ожидать, что полупроводниковые лазеры будут иметь внешний квантовый выход примерно в 100 раз выше, чем у сопоставимых СИД. Так как лазеры и СИД работают при одинаковых напряжениях, возрастание внешнего квантового выхода переводится в пропорциональное возрастание всей энергетической эффективности.

Лекция 3

Полупроводниковые источники некогерентного излучения. Основные физические параметры. Полупроводниковые лазеры. Лазерные диоды на p-n переходе, с туннельной инжекцией, на гетероструктурах, с распределенной обратной связью. Вопросы надежности.

Излучение света в полупроводниках

Излучение света в полупроводниках обычно происходит благодаря рекомбинации носителей. Процесс излучения можно разделить на два типа: спонтанное излучение, в котором дырки и электроны рекомбинируют случайно, и стимулированное излучение, в котором дырки и электроны стимулируются к рекомбинации уже существующими фотонами.

Спонтанное излучение

При поглощении света электрон поглощает фотон и переходит в более высокое энергетическое состояние. При излучении происходит противоположный эффект. Электрон переходит из более высокого энергетического состояния в более низкое и при этом теряет энергию, которая излучается в виде фотона. Переходы с излучением света в полупроводниках являются межзонными переходами; они происходят между зоной проводимости и валентной зоной и (или) определенными состояниями внутри запрещенной зоны, обусловленными легирующими примесями или дефектами. Так как запрещенная зона полупроводников обычно составляет от нескольких десятых до нескольких единиц электронвольта, длина волны излучения обычно лежит в инфракрасной области спектра и приблизительно соответствует длине волны края полосы поглощения в полупроводнике. Как и в случае поглощения, существуют определенные правила отбора, ограничивающие возможные излучательные переходы. В частности,

требуется, чтобы выполнялись законы сохранения энергии импульса. Таким образом, для прямых переходов имеют место соотношения

$$E_i - E_f = h\nu_{фот},$$

$$k_i - k_f = (2\pi / \lambda_{фот})u, \qquad (1)$$

где u - единичный вектор, определяющий лишь направление. Отметим, что для прямых переходов $2\pi/\lambda_{фот} << |k_i|$ и $|k_f|$, так что $|k_i| \approx |k_f|$, как и для прямого перехода с поглощением. Излучение фотона может также происходить при непрямом переходе, причем в этом случае также требуется выполнение законов сохранения энергии и импульса:

$$k_i - k_f \pm q = (2\pi / \lambda_{фот})u, \quad (2)$$

где q - волновой вектор поглощаемого или испускаемого фотона. Как и в случае поглощения, необходимость участия фонона приводит к значительному уменьшению вероятности перехода и, следовательно, к генерации фотона.

Третьим требованием для излучательного перехода является то, чтобы были заполнены верхние (исходные) энергетические уровни и свободны соответствующие нижние (конечные) энергетические уровни, причем разность в энергии между ними должна равняться энергии излучаемого фотона. Это правило выглядит очевидным и тривиальным, но играет малую роль в случае поглощения ввиду того, что поглощающий материал обычно находится в тепловом равновесии и, следовательно, всегда имеет заполненные нижние энергетические уровни и свободные верхние. (За исключением случая, когда энергия фотонов меньше межзонной. Тогда не существует возможного верхнего состояния для электронов и, следовательно, поглощение не происходит). В случае излучения упомянутое правило является более важным. Например, для характерных полупроводников при тепловом равновесии и комнатной температуре и ниже существует относительно малое число дырок в валентной зоне и электронов в зоне проводимости. В случае легированного полупроводника в зоне проводимости может быть много электронов, как, например, в материале я-типа, но при этом очень мало дырок в валентной зоне

(вследствие постоянства произведения яр). Противоположная картина имеет место для полупроводников р-типа. Следовательно, излучательная рекомбинация, происходящая при тепловом равновесии, дает очень мало фотонов, которые перепоглощаются до того, как покинут кристалл. Для получения значительного светового излучения от полупроводника необходимо каким-либо образом вывести полупроводник из состояния теплового равновесия, чтобы создать больше электронов в зоне проводимости и дырок в валентной зоне. После накачки до более высоких энергетических уровней электроны и дырки «усаживаются» на состояния, соответствующие краям их зон.

Увеличения концентрации дырок и электронов можно достичь разными путями, и этот процесс обычно называется накачкой материала. Например, если свет большой интенсивности с энергией фотонов, большей энергетической ширины запрещенной зоны, направлен на полупроводник, то происходит межзонное поглощение, в результате чего будет возникать большое количество электронно-дырочных пар. Эти дырки и электроны первоначально горячие, но очень быстро термолизуются, т. ё, электроны садятся на дно зоны проводимости, а дырки поднимаются к верхней границе валентной зоны через взаимодействие с решеткой. Переход в тепловое равновесие (термолизация) обычно происходит очень быстро - за время порядка 10^{-14} с. После его установления электроны и дырки рекомбинируют; эта рекомбинация может быть излучательной, приводящей к излучению фотона с энергией, приблизительно равной ширине запрещенной зоны. Время жизни, или время, в среднем необходимое для рекомбинации электронно-дырочных пар после их возникновения, обычно составляет около 10^{-11} с в накачиваемом полупроводнике (с прямой запрещенной зоной). Заметим, что в полупроводниках с непрямой запрещенной зоной переход в тепловое равновесие вынуждает электроны и дырки занимать состояния со значительно различающимися волновыми векторами. Это означает, что, для того чтобы произошла рекомбинация, необходим фонон, так же как и дырка и электрон, а следовательно, это менее вероятный процесс. Типичное время жизни для непрямой рекомбинация достигает 0,25 с. Прежде чем он произойдет, дырка и электрон в среднем уже рекомбинируют в результате безызлучательного процесса с передачей энергии решетке или дефектам и т. д. Следовательно, эффективность непрямой излучательной рекомбинации очень мала. Квантовая эффективность для материала с прямой

запрещенной зоной может достигать 1, в то время как для материала с непрямой зоной он обычно составляет 0,001 или менее. Однако непрямые излучательные переходы имеют место, и в некоторых материалах с непрямой запрещенной зоной (например, GaP) может происходить вполне измеримое и действительно полезное излучение света. Тем не менее чрезвычайно трудно изготовить лазеры на материалах «с непрямым переходом.

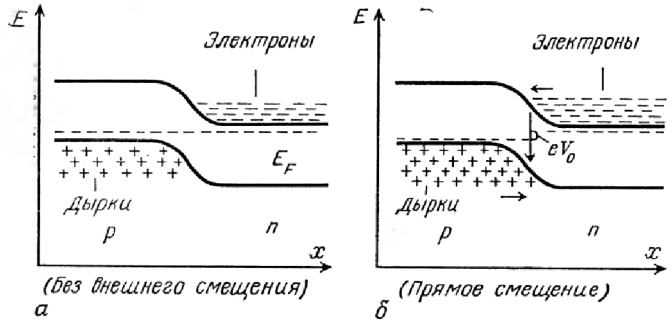


Рисунок 1 - Расположение энергетических зон p-n перехода светодиода.
а) в отсутствие внешнего напряжения смещения; б) внешнее напряжение смещения V_0 приложено в прямом направлении

Метод оптической накачки, который был описан как способ получения необходимого увеличения концентрации электронов ш дырок, является сравнительно громоздким, так как требует •очень интенсивных источников света; этот процесс малоэффективен, и большая часть энергии превращается в тепло. Кроме того, выходящее излучение необходимо фильтровать, чтобы выделить излучение полупроводника из отраженного света накачки

Другой способ накачки, который одновременно эффективен и очень прост, использует свойства p-n перехода. Расположение энергетических зон светоизлучающего p-n перехода представлено на рисунке 1. В отсутствие напряжения смещения существует много электронов в зоне проводимости на n-стороне и много дырок в валентной зоне на p-стороне, но лишь некоторые из них могут преодолеть барьер и попасть в область перехода. После приложения прямого напряжения смещения V_0 барьер понижается и многие

дырки и электроны проникают (инжектируют) в область перехода, где они рекомбинируют, генерируя фотоны. Свет, излучаемый под воздействием электрического напряжения, называется электролюминесценцией, в то время как свет, получаемый путем накачки оптическим источником с более короткой длиной волны излучения, называется фотолюминесценцией. Типичная схема электролюминесцентного диода, или светоизлучающего диода (СИД), представлена на рисунке 2.

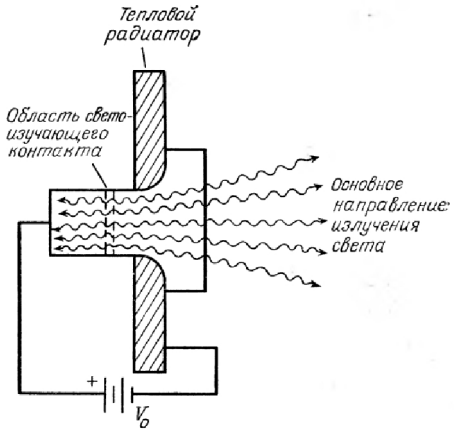


Рисунок 2 - Светоизлучающий диод (СИД) на p-n переходе

Свет генерирует в p-n переходе и покидает диод, проходя через объемный материал вне области перехода. Большая часть света перепоглощается в полупроводнике до того, как он покинет диод. Следовательно, можно говорить как о внешнем, так и о внутреннем квантовых выходах, определяемых следующим образом:

$$\eta_{внешн} = \frac{число\,фотонов,\ \ излученных\ \ в\ \ желаемом\ \ направлении}{число\ \ \ электронно-дырочных\ \ \ пар}, \quad (3)$$

$$\eta_{внутр} = \frac{число\ \ \ генерируемых\ \ \ фотонов}{число\ \ \ электронно-дырочных\ \ \ пар},$$

При разработке светоизлучающих диодов обычно стремятся минимизировать перепоглощение света путем изготовления очень тонкого слоя материала между переходом и поверхностью, а также (если возможно) путем подбора материала, имеющего очень малое поглощение. Такие светодиоды получили широкое распространение в панельных индикаторах и дисплеях вследствие того, что они очень малы, дешевы, более надежны, чем устройства на лампах накаливания, и легко объединяются в комплекты интегральных схем. СИД также очень часто используются в качестве источников света в волоконно-оптических линиях связи, так как их излучение можно модулировать на больших частотах (хотя и не таких высоких, как в лазерах), а сами СИД можно удовлетворительно соединять с сердцевинами многомодовых волокон.
Спектр длин волн полупроводниковых излучателей света независимо от используемого типа накачки является довольно простым по сравнению, например, со спектром газоразрядной дуговой лампы, которая имеет много линий или эмиссионных максимумов. Хороший полупроводниковый светоизлучающий диод с малым количеством дефектов или низким содержанием загрязняющих примесей имеет один эмиссионный максимум вблизи длины волны, соответствующей краю полосы поглощения, т. е. энергии запрещенной зоны. Этот максимум имеет полуширину обычно 20 - 30 нм, которая значительно больше ширины линий газового источника. Соответствующие загрязняющим примесям и дефектам состояния внутри запрещенной зоны порождают эмиссионные пики на других длинах волн. Эти пики показывают, что некоторые из инжектируемых электронно-дырочных пар рекомбинируют с излучением фотонов, не принадлежащих основному пику, соответствующему краю зоны. Следовательно, квантовая эффективность на желаемой длине волны уменьшается.

Лазерный эффект

Если выполнены условия инверсии населенностей и существования высокой плотности фотонов, то возникает процесс селекции, который изменяет характер излучения и в результате которого появляется значительное количество стимулированного

излучения. При наличии соответствующей оптически резонирующей структуры может иметь место лазерный эффект.

Полупроводниковые лазерные структуры

Обычно используемая структура для дискретных полупроводниковых лазеров представляет собой прямоугольный параллелепипед, аналогичный представленному на рисунке 3, который изготавливается путем скалывания кристалла вдоль плоскостей кристаллической решетки, которые перпендикулярны светоизлучающему слою. Например, для изготовления такой структуры пластина GaAs <100> может легко раскалываться по четырем перпендикулярным плоскостям <110>. Типичная длина L и ширина W составляют 300 и 50 мкм соответственно, однако возможны значительные отклонения от этих величиню. Две торцевые поверхности шириной W образуют оптически резонансную отражательную структуру, известную как эталон Фабри-Перо. При использовании такой структуры увеличение накачиваемой мощности будет приводить к лазерному эффекту, если превышен определенный критический пороговый уровень.

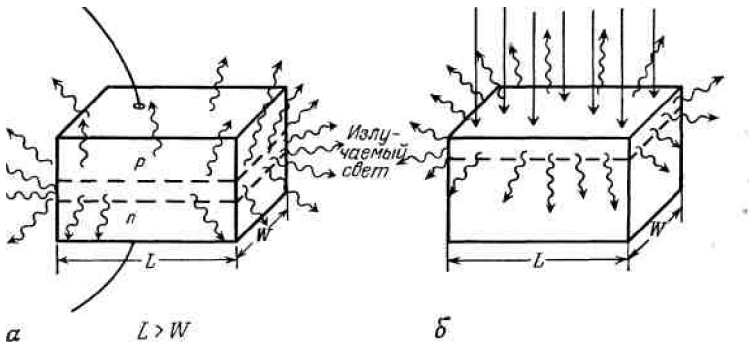


Рисунок 3 - Полупроводниковые лазерные структуры: а - накачиваемый через p-n переход лазерный диод; б - накачка электронами или фотонами. Штриховые линии ограничивают зону максимальной накачки.

Пороговый уровень для лазерного эффекта

По мере возрастания от нуля накачиваемой мощности увеличивается число фотонов, испускаемых за счет спонтанной эмиссии. Даже если мощность накачки превышает требуемую для создания инверсной населенности, излучение вначале преимущественно состоит из спонтанно образующихся фотонов. Спонтанные фотоны излучаются более или менее одинаково во всех направлениях, как показано на рисунке 3. Тем не менее, большинство этих фотонов начинает свой путь в направлениях, которые очень быстро выводят их из области с инверсной населенностью, где может происходить преимущественная стимулированная эмиссия, и, таким образом, они оказываются неспособными воспроизводить самих себя. Некоторые фотоны, случайно бегущие вдоль области с инверсией населенности, оказываются способными создавать себе подобные фотоны много раз за счет стимулированной эмиссии, прежде чем они покинут лазер. Кроме того, для любой данной энергии межзонной щели и распределения дырок и электронов будет существовать одна особая энергия (длина волны), которая является преимущественной (переходы более вероятны) по сравнению с другими. Эта длина волны или энергия обычно соответствует в первом приближении длине волны пика спонтанного излучения в материале. В результате этих преимуществ по энергии и направлению, когда возникает стимулированное излучение, испускаемое излучение сужается как по пространственной расходимости, так и по спектральной ширине полосы (уменьшающейся от нескольких сотен ангстрем приблизительно до 2,5 нм). Заметим, что после возникновения стимулированного излучения расходуется больше электронно-дырочных пар в пересчете на 1 с. Следовательно, для любой данной скорости генерации электронно-дырочных пар спонтанное излучение подавляется, поскольку стимулированное излучение расходует генерируемые пары прежде, чем они спонтанно рекомбинируют. Такое преобладающее стимулированное излучение, как оно было описано, известно как суперлюминесцентное излучение. Это еще не когерентное излучение света, производимое за счет лазерного эффекта. Такое излучение обусловлено усилением спонтанного излучения. Фотоны не движутся точно в одном направлении; они могут слегка отличаться по энергиям, но не синхронизованы по фазе. Иными словами, излучение еще некогерентно.

В том случае, когда светоизлучающая область ограничена параллельными плоскостями, как показано на рисунке 3, будет иметь место некоторое отражение от

поверхностей и, кроме того, часть излучения будет проходить назад и снова вперед через область с инверсной населенностью и еще более усиливаться. Чтобы отразиться назад и затем вперед в узком слое инверсной населенности, фотоны должны быть строго в плоскости этого слоя и двигаться в направлении, строго перпендикулярном поверхностям Фабри-Перо. Излучение также должно быть однородным по частоте и фазе, чтобы избежать деструктивной интерференции, так как отражатели Фабри-Перо формируют резонаторную полость, в которой устанавливаются оптические моды. Таким образом, только та малая часть суперлюминесцентного излучения, которая удовлетворяет всем этим требованиям, продолжает существовать и становится доминирующей разновидностью излучения, выходящего из лазера. Такое излучение, однородное по длине волны, фазе, направлению и поляризации, является когерентным. В этом случае говорят, что устройство обладает лазерным эффектом.

Переход от суперлюминесценции к лазерному эффекту (по мере нарастания входной мощности) принимается как точка, при которой усиление преимущественной моды за счет стимулированной эмиссии становится больше поглощения и всех других потерь фотонов, так что излучение при отражении проходит туда и обратно между поверхностями Фабри-Перо, возрастая по интенсивности при каждом проходе (т. е. устанавливаются осцилляции). Конечно, излучение не будет нарастать по интенсивности до бесконечности. По мере роста оно вызывает больше рекомбинаций до тех пор, пока концентрации дырок и электронов не будут использоваться с той же скоростью, с какой они образуются при данном источнике накачки, т. е. пока не установится динамическое равновесие. В этом стационарном состоянии лазер генерирует на одной пиковой длине волны (или, возможно, нескольких дискретных длинах волн, соответствующих различным продольным модам), излучение его имеет высокую монохроматичность (приблизительно ширину полосы 0,1 нм); внутри кристалла возникает стоячая волна с целым числом полуволн, укладывающихся между поверхностями Фабри-Перо для каждой моды. Вследствие описанной ранее селективности по направлениям отражающих торцевых поверхностей фотоны лазерной моды излучаются преимущественно в направлении, перпендикулярном отражающим поверхностям, как показано на рисунке 4.

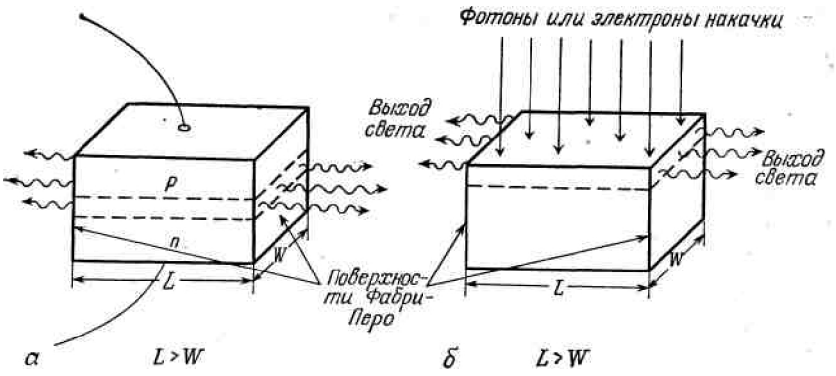


Рисунок 4 - Полупроводниковые лазеры, работающие при накачке, превышаю щей порог генерации: а - накачка током в p-n переходе; б - накачка электронами или фотонами.

Световые волны (фотоны), излучаемые из такой резонансной структуры, находятся в фазе, и, кроме того, волны, которые могут наиболее легко распространяться, являются волнами с электрическим вектором, перпендикулярным плоскости слоя инверсной населенности. Следовательно, излучаемый свет в этом случае является нормально поляризованным.

Переходы между областями спонтанной, суперлюминесцентной и лазерной эмиссии по мере возрастания накачиваемой мощности происходят очень резко. Действительно, хотя переход между суперлюминесценцией и лазерным излучением был описан в деталях с педагогической целью, реальный переход в обычном лазере настолько быстрый, что суперлюминесцентное излучение не наблюдается. Когда накачиваемая мощность превышает пороговую точку, характер светового излучения резко изменяется и приобретает лазерный тип генерации с убыванием полосы излучения от нескольких десятков нм до 0,1 нм и с наблюдаемой когерентностью фазы. Промежуточная стадия «суженной» полосы излучения, возникающая за счет преимущественно стимулированного излучения без когерентности по фазе (т. е. суперлюминесцентная), проходит настолько быстро, что ее невозможно наблюдать. Только в лазерных диодных структурах с поврежденными отражателями Фабри-Перо, которые сдерживают лазерный эффект, можно наблюдать суперлюминесцентное излучение.

Эффективность излучения света

Одной из важнейших характеристик светового излучения полупроводниковых лазеров, которая еще не обсуждалась, является эффективность преобразования электрической энергии в оптическую. Лазерные диоды, работающие выше порогового уровня, имеют приблизительно в 100 раз большую эффективность, чем светоизлучающий диод, изготовленный из того же материала. Такое улучшение эффективности происходит за счет нескольких факторов.

Как ранее упоминалось, генерация фотонов с помощью стимулированного излучения имеет сравнительно низкую вероятность возникновения по сравнению с вероятностью спонтанного излучения в случае неинвертированнои населенности, как это происходит в светоизлучающем диоде (СИД). Однако, когда есть инверсия населенностей, как в лазере, вероятность стимулированного излучения становится очень большой, даже больше вероятностей спонтанного излучения и безызлучательных рекомбинаций электронно-дырочных пар. Например, внутренняя квантовая эффективность в лазерном диоде примерна в 10 раз выше, чем в СИД. Фактически, если лазерный диод охлаждается до 77К, для снижения температурного разброса электронов и дырок по энергиям его внутренний квантовый выход может приблизиться к 100%.

Если фотоны образовались, то они также имеют меньшие потери в лазерах, чем в СИД. Фотоны в лазерных модах проходят главным образом в области инверсионной населенности, в которой внутризонное поглощение подавляется за счет недостатка электронов вблизи края валентной зоны и незаполненных состояний вблизи края зоны проводимости. Фотоны в СИД двигаются главным образом в безынверсном объеме материала. Таким образом, внутреннее перепоглощение испускаемых фотонов в лазере намного меньше, чем в СИД. Кроме того, фотоны, излучаемые из лазеров, образуют пучок света с много меньшей расходимостью, чем фотоны, излучаемые из СИД, что позволяет собирать (или вводить) больше света в заданном месте. Вследствие совместного эффекта снижения внутреннего перепоглощения высокой коллимации пучка и возрастания внутреннего квантового выхода можно ожидать, что полупроводниковые лазеры будут иметь внешний квантовый выход примерно в 100 раз выше, чем у сопоставимых СИД. Так как лазеры и СИД работают при одинаковых напряжениях, возрастание внешнего квантового

выхода переводится в пропорциональное возрастание всей энергетической эффективности.

Полупроводниковые лазеры

Вероятно, наиболее значительной особенностью этого излучения является возможность разработки такого источника, в котором стимулированное излучение фотонов будет преобладать над спонтанной эмиссией и поглощением. Если в конструкции предусмотреть резонансные отражающие структуры, например две параллельные плоскости, то в результате могут возникнуть моды генерации и будет иметь место когерентное оптическое излучение. В этом разделе мы рассмотрим несколько основных полупроводниковых лазерных структур и разработаем количественную теорию, необходимую для расчета их ожидаемых рабочих характеристик.

Лазерные диоды

Со времени своего появления в 1962 г. полупроводниковые лазерные диоды превратились из лабораторных образцов в надежные и годные для продажи устройства, которые получили широкое распространение в качестве источника света во многих случаях применения. В этом разделе мы обсудим основные структуры и работу дискретных лазерных диодов в противоположность монолитно интегрированным лазерным диодам оптических интегральных схем.

Основные структуры

Лазерные диоды с p-n переходом являются превосходными источниками света для использования в оптических интегральных схемах и передачи сигнала в волоконно-оптических системах благодаря своим малым размерам, относительно простой конструкции и высокой надежности. К настоящему времени большинство лазерных диодов изготовлено из GaAs или $Ga_{(1-x)}AL_xAs$, однако другие материалы, несомненно,

также будут использоваться для получения излучения на различных длинах волн, когда будет разработана технология их изготовления. Основные структуры лазерных диодов на p-n переходе представлены на рисунке 5. Обычно p-n переход формируется «путем эпитаксиального выращивания слоя p-типа на подложке n-типа. В каждой области создаются омические контакты, чтобы пропускать электрический ток, который является источником энергии накачки, необходимой для создания инверсной населенности в активной области, примыкающей к переходу. Две параллельные торцевые поверхности изготовляются для работы в качестве зеркала и создания оптической обратной связи, необходимой для генерации оптических мод. Устройство, представленное на рисунке 5, является дискретным диодом, который можно использовать совместно с волоконно-оптической передающей линией. В последующих главах мы рассмотрим более сложные многослойные структуры с более современными оптическими схемами обратной связи, которые лучше подходят для монолитных конструкций в оптических интегральных схемах. Однако основная структура является удобной базой для теоретического описания работы лазера, которую можно легко приспособить к более сложным устройствам.

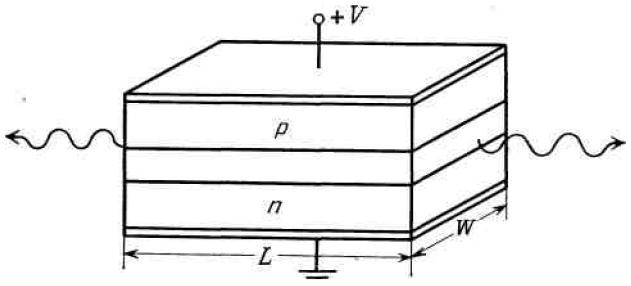


Рисунок 5 - Основная структура лазера на p-n переходе.

Оптические моды

Частично отражающие торцевые поверхности лазерного диода обеспечивают оптическую обратную связь, приводящую к установлению одной или нескольких продольных оптических мод. Благодаря подобию этих плоскостей параллельным зеркалам интерферометра Фабри-Перо торцевые поверхности лазера часто называются

поверхностями Фабри-Перо. Когда ток проходит через лазерный диод, то свет может генерироваться за счет инверсного населенного слоя одновременно путем спонтанного и стимулированного излучения фотонов. Вследствие отражений, которые происходят на поверхностях Фабри-Перо, некоторые фотоны будут многократно проходить туда и обратно через область инверсной населенности и будут преимущественно усиливаться стимулированным излучением. Эти фотоны, которые двигаются строго перпендикулярно поверхностям Фабри-Перо, имеют наибольшую вероятность выживания в слое инверсной населенности, в которой они воспроизводятся стимулированным излучением. Следовательно, они становятся фотонами оптических мод или мод, которые устанавливаются, когда достигается стационарное состояние при данном уровне тока. Имеется также возможность появления других мод, соответствующих фотонам, двигающимся зигзагообразно с отражением от боковых стенок лазера, однако в реальных устройствах эти моды обычно подавляются путем создания неровностей на боковых поверхностях или за счет использования специальных методов ослабления нежелательных мод.

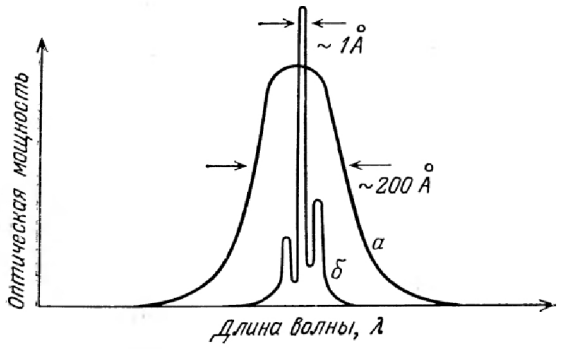


Рисунок 6 - Эмиссионный спектр лазерного диода:
а - спонтанная эмиссия ниже порога; б - лазерная модовая структура выше порога.

Излучение лазерных мод должно быть также однородным по частоте и фазе, чтобы избежать ослабления при интерференции. В результате внутри лазерного диода

устанавливается стоячая волна с целым числом полупериодов между параллельными торцевыми поверхностями. Модовое число m задается числом полуволн:

$$m = 2Ln / \lambda_0 , \qquad (4)$$

где L - расстояние между торцами, n - коэффициент преломления лазерного материала, λ_0 - длина волны излучаемого света в вакууме. Модовое разделение можно установить, взяв производную $dm/d\lambda_0$ и принимая во внимание, что полупроводниковые лазеры всегда работают вблизи длины волны, соответствующей ширине запрещенной зоны, где n - функция длины волны. Таким образом:

$$dm / d\lambda_0 = -2Ln / \lambda_0^2 + (2L / \lambda_0)(dn / d\lambda_0), \qquad (5)$$

При dm = -1 модовое разделение $d\lambda_0$ задается выражением

$$d\lambda_0 = \lambda_0^2 / (L(n - \lambda_0 (dn / d\lambda_0))). \qquad (6)$$

Мы выбрали dm= -1 потому, что уменьшение на единицу величины m соответствует потере одной полуволны между торцевыми поверхностями Фабри-Перо, т. е. увеличению длины волны λ_0.

Типичный модовый спектр лазерного диода представлен на рисунке 6. Обычно могут сосуществовать несколько продольных мод, имеющих длины волн вблизи пика спонтанной эмиссии. Обычно модовое разделение для лазера на основе GaAs составляет $d\lambda_0 \approx 0{,}3$ нм. Чтобы добиться работы в одномодовом режиме, необходимо модифицировать лазерную структуру, чтобы подавить все моды, кроме основной.

Пороговые условия для генерации света

Когда на лазерный диод подается прямое смещение и начинает протекать ток, устройство не сразу начинает работать в лазерном режиме. При низком уровне тока испускаемый свет возникает главным образом за счет спонтанной эмиссии и имеет спектральную характеристику шириной порядка сотен ангстрем — некогерентный свет. По мере нарастания тока накачки в области перехода создается высокая степень инверсной населенности и излучается больше фотонов. Спонтанное излучение происходит более или менее равномерно во всех направлениях. Большинство фотонов двигается в направлениях, приводящих их к быстрому выходу из области с инверсной населенностью, в которой может происходить стимулированная эмиссия, и, таким образом, они оказываются неспособными воспроизводить самих себя. Однако те несколько фотонов, которые проходят строго в плоскости перехода и перпендикулярно отражающим торцевым поверхностям, в состоянии многократно повторить себя до того, как они выйдут из лазера. Кроме того, для любой данной энергии между зонами и для распределения дырок и электронов существует одно особое значение энергии (длины волны), которое является преимущественным по сравнению с другими. В первом приближении эта длина волны обычно соответствует длине волны эмиссионного пика в данном материале. В результате такого преимущества по энергии и направлению, когда стимулированное излучение возникает с возрастанием тока, испускаемое излучение существенно сужается одновре: менно по ширине спектра и по пространственной расходимости. Когда возникает стимулированное излучение, возрастает и плотность фотонов (интенсивность) оптической моды, что приводит к дальнейшему возрастанию стимулированного излучения и большему расходу электронно-дырочных пар в единицу времени. Следовательно, спонтанное излучение подавляется при данной скорости генерации электронно-дырочных пар из-за того, что стимулированное излучение использует генерируемые пары до того, как они спонтанно рекомбинируют. Вследствие фазового условия (4) свет, создаваемый стимулированным излучением в оптической резонансной структуре, подобной структуре лазерного диода, является когерентным, а устройство, как говорят, работает в режиме генерации.

Переход от спонтанной эмиссии к лазерному излучению происходит резко после того, как ток превысит пороговое значение. Как только будет превышен порог, можно

экспериментально наблюдать начало лазерной генерации, отметив резкий излом на кривой зависимости мощности оптического излучения от тока накачки (рисунок 7), который является результатом резкого увеличения квантового выхода в лазерном процессе.

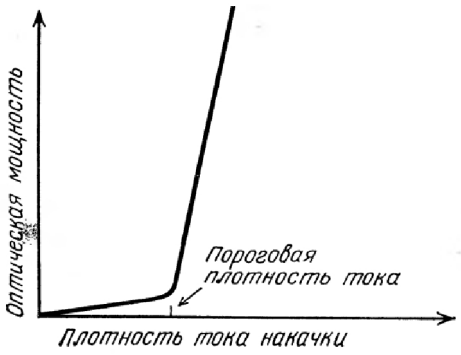


Рисунок 7 - Зависимость светового излучения на выходе лазерного диода от тока накачки на входе.

Вместе с тем форма кривой спектрального распределения резко изменяется от широкой кривой распределения спонтанной эмиссии к кривой с несколькими узкими модами, как показано на рисунке 6. С количественной точки зрения порог лазерной генерации соответствует точке, при которой возрастание числа фотонов (в секунду) лазерных мод за счет стимулированного излучения точно равно числу фотонов (в секунду), теряемых за счет рассеяния, поглощения или излучения из лазера. В общепринятых терминах, используемых для описания осциллятора, можно сказать, что устройство имеет усиление на циклический проход, равный единице. Пользуясь этим, можно получить функциональное выражение порогового тока в зависимости от различных материалов и геометрических параметров.

Начнем с рассмотрения лазерной структуры на основе p-n перехода, показанной на рисунке 5. Свет из лазера излучается преимущественно в направлении, перпендикулярном поверхностям Фабри-Перо. Поперечное пространственное распределение энергии световых волн (плотность фотонов) представлено на рисунке 8. Распределение фотонов простирается в неактивную (неинвертируемую) область с каждой стороны перехода в

основном за счет дифракции. Таким образом, существует светоизлучающий слой толщиной D, который больше толщины d слоя с активной или инверсной населенностью. Например, в диоде из GaAs d≈1 мкм, D≈10 мкм. На основе идеализированной кривой пространственного распределения энергии можно заключить, что из полного числа фотонов, существующих в лазерных модах в любое данное время, только часть d/D остается в активной области и может генерировать дополнительные фотоны за счет стимулированного излучения. Этот эффект снижает усиление, возможное в этом устройстве.

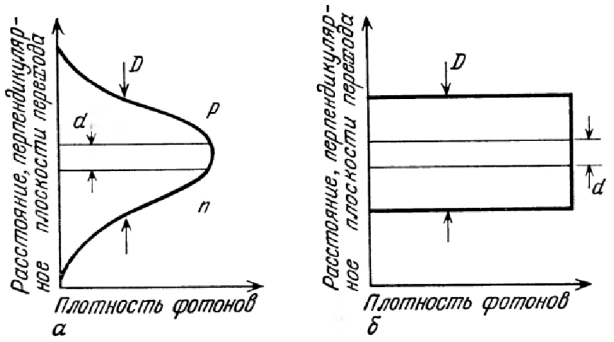


Рисунок 8 - Поперечное пространственное распределение энергии в лазерном диоде: а - фактическое; б - идеализированное (область под идеализированной кривой соотнесена к реальной так, чтобы отношение d/D оставалось постоянным).

Чтобы вывести количественное выражение для пороговой плотности тока, необходимого для появления лазерного эффекта, рассмотрим один проход волны лазерного света (поток фотонов) от одной поверхности Фабри-Перо до другой. Диаграмма потока мощности выглядит так, как это показано на рисунке 9.

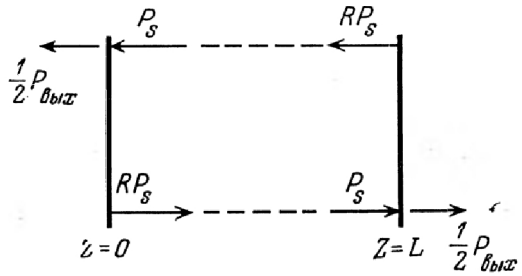


Рисунок 9 – Диаграмма потоков мощностей для лазерного диода.

Величина Ps представляет собой оптическую мощность, падающую изнутри на каждую торцевую поверхность, а R - коэффициент отражения мощности. При выполнении условий генерации (лазерной генерации) произведение RP_S экспоненциально увеличивается в зависимости от расстояния, достигая за один проход величины Ps на противоположной поверхности Фабри-Перо (Ф-П). Существующий механизм потерь преодолевается механизмом лазерного усиления за счет стимулированного излучения. На каждой Ф-П поверхности излучается мощность, составляющая $1/2P_{вых}=(1-R)P_S$. Если α [$см^{-1}$] — коэффициент потерь для волны по мере ее распространения (включая все типы потерь), а g [$см^{-1}$] —коэффициент усиления, то зависимость мощности от расстояния представляется в виде:

$$P = RP_S \exp(g(d/D) - \alpha)Z, \quad (7)$$

Заметим, что усиление имеет место только в области с инверсной населенностью и, следовательно, величину g необходимо умножить на d/D, в то время как потери имеются повсюду, где существует поле, и, следовательно, коэффициент а не имеет такого множителя. При генерации (усиление за один цикл равно 1) должно выполняться условие:

$$P_S = RP_S \exp(g(d/D) - \alpha)L, \quad (8)$$

или

$$\ln(1/R) = (g(d/D) - \alpha)L\,. \qquad (9)$$

Таким образом, условие лазерной генерации будет следующим:

$$g(d/D) = \alpha + (1/L)\ln(1/R), \qquad (10)$$

Коэффициент усиления g связан с плотностью инжектируемого тока электронов и дырок. Можно показать, что функциональная зависимость для величины g имеет вид:

$$g = \eta_1 \lambda_0^2 J / 9\pi e n^2 d\Delta\nu\,, \qquad (11)$$

Для GaAs при температуре 300 К:

- внутренняя квантовая эффективность η_q=0,7;
- длина волны излучения в вакууме λ_0=0,9 мкм;
- коэффициент преломления n=3,34 на λ_0;
- ширина полосы спонтанного излучения $\Delta\nu=1{,}5*10^{13}$ c^{-1};
- заряд электрона $e=1{,}6*10^{-19}$ Кл;
- толщина активной области d=100 мкм;
- плотность инжектируемого тока J.

Таким образом, подставляя выражение (11), которое справедливо как выше, так и ниже порога, в соотношение, определяющее условия порога (10), получим:

$$\frac{\eta_q \lambda_0^2 J_n}{8\pi e n^2 \Delta\nu D} = \alpha + \frac{1}{L}\ln\frac{1}{R}$$

или

$$J_n = \frac{8\pi e n^2 \Delta\nu D}{\eta_q \lambda_0^2}\left(\alpha + \frac{1}{L}\ln\frac{1}{R}\right), \qquad (12)$$

Таким образом, пороговый ток (плотность) является таким, который необходим, чтобы создать усиление, достаточное для покрытия потерь. Заметим, что с точки зрения пороговых условий выход света из торцевой поверхности лазера должен учитываться как потери. Этот факт учитывается членом (1/L)ln(1/R). Может быть, непосредственно не очевидно, что этот член описывает потери, обусловленные излучением с торцевой поверхности, однако если мы подставим коэффициент пропускания T=1-R и разложим ln[l/(l-T)] в ряд, то получим

$$(1/L)\ln(1/R) = (1/L)\ln(1/(1-T)) = (1/L)\left(T - \left(T^2/2\right) + \left(T^3/3\right) - \left(T^4/4\right) + \ldots\right), \qquad (13)$$

Пренебрегая членами высокого порядка по T:

$$(1/L)\ln(1/R) \approx T/L, \qquad (14)$$

Величина T/L представляет собой коэффициент потерь [см$^{-1}$], полученный путем усреднения коэффициента T потерь на торцах по всей длине. Так как обычно T≈0,6 (для GaAs), то отбрасывание T^2 и членов более высокого порядка приводит к неточному результату, и в более строгих количественных расчетах их нельзя отбрасывать. Однако цель данного примера - качественно показать, что природа члена (l/L)ln(1/R) связана с усреднением коэффициента потерь по единице длины за счет выхода фотонов из лазера. Из теоретических выкладок, приводящих к выражению (12), можно видеть, что распределение фотонов (оптических полей) от области инверсной населенности в окружающую пассивную область приводит к существенному возрастанию порогового тока. Этот факт предполагает, что лазерные диоды должны конструироваться так, чтобы для достижения оптимальных характеристик отношение D/d=1.

Выходная мощность и эффективность

Выражения для общей эффективности по мощности и выходной мощности лазерного диода можно установить из следующего. Во-первых, рассмотрим потери на небольшом (по приращению) расстоянии AZ. В первом приближении потери мощности на этом расстоянии будут следующими:

$$P_{пот} = P - Pe^{-\alpha\Delta Z} = P\left(1 - e^{-\alpha Z}\right) \approx P\left(1 - \left(1 - \alpha\Delta Z\right)\right) = \alpha P\Delta Z\,, \quad (15)$$

или

$$dP_{пот} = \alpha P dZ\,, \quad (16)$$

Следовательно, мощность поглощаемая на длине L, составляет

$$P_{пот} = \int_{Z=0}^{Z=L} dP_{пот} = \alpha \int_{0}^{L} P dZ\,, \quad (17)$$

Аналогично мощность, генерируемая на одном проходе, описывается выражением:

$$P_{ген} = \int_{Z=0}^{Z=L} dP_{ген} = g\frac{d}{D}\int_{0}^{L} P dZ\,, \quad (18)$$

Таким образом, внутренняя эффективность η (по мощности) определяется следующим образом:

$$\eta = \left(P_{ген} - P_{пот}\right) / P_{ген} = \left(g\left(d / D\right) - \alpha\right) / \left(gd / D\right), \quad (19)$$

Заменяя переменные, согласно выражению (20), находим:

$$\eta = (1/L)\ln(1/R)/[\alpha + (1/L)\ln(1/R)], \quad (20)$$

Заметим снова, что α — коэффициент полных потерь, включающий внутризонное поглощение и поглощение на свободных носителях, а также любые потери на рассеяние. В этом случае мощность на выходе определяется формулой

$$P_{вых} = \eta P_{вн} = \eta[(J/e)\eta_q (LxW)h\nu], \quad (21)$$

где $P_{вн}$ — оптическая мощность, генерируемая внутри лазера; L - длина; W - ширина; ν - частота генерации. Подставляя η из выражения (20), получим

$$P_{вых} = (1/L)\ln(1/R)/[\alpha + (1/L)\ln(1/R)][(J\eta_q/e)(LxW)h\nu], \quad (22)$$

где $P_{вых}$ - энергия выхода с обеих торцевых поверхностей. В случае когда последовательным сопротивлением диода можно пренебречь:

$$\eta_{полн} = (1/L)\ln(1/R)/[\alpha + (1/L)\ln(1/R)]\eta_{й}, \quad (23)$$

Заметим, что эти формулы для мощности и эффективности справедливы при пороговом уровне и выше порогового уровня. Для плотностей тока ниже порога (т. е. до начала генерации) равенство (10) не выполняется, и в этих выражениях должны быть оставлены величины g, D, d и α, а не L, R и α.

Эффективность полупроводниковых лазеров уменьшается с ростом температуры за счет появления двух явлений. Во-первых, увеличения коэффициента α и, во-вторых, понижения квантовой эффективности η_q. Последний эффект имеет место потому, что электроны и дырки размазываются по энергиям в широкую зону за счет температурного возбуждения, и, следовательно, меньше электронно-дырочных пар с подходящим энер-

гетическим разделением доступно для стимулированного излучения в генерируемую моду при любом заданном уровне инжектируемого входного тока. Эти два эффекта ответственны за возрастание плотности порогового тока, необходимого для появления лазерного излучения . Очевидно, что для данной плотности тока на входе мощность на выходе уменьшается при возрастании температуры под влиянием тех же двух факторов. Вследствие ранее упомянутых отрицательных явлений нагрева лазеры на p-n переходе обычно работают в импульсном режиме. Времена включения и спада генерации диодов очень коротки (10^{-10} с или короче); следовательно, диоды могут легко работать в режиме прямоугольного импульса длительностью 100 нс. Общая эффективность по мощности обычно низка (несколько процентов), а пиковая выходная мощность обычно составляет 10 Вт. Полупроводниковые лазеры на p-n переходах могут работать при комнатной температуре (или выше) в непрерывном режиме, за исключением того, что ток накачки, необходимый для достижения порога, не должен превышать тока, приводящего к разрушению устройства за счет его перегрева. Типичные плотности порогового тока для основных полупроводниковых лазеров на p-n переходах, подобных представленному на рисунке 6, составляют порядка 10^4 А/см^2, но, так как площадь диода составляет около 10^{-3} см^2, максимальный ток обычно достигает порядка 10 А. Однако эффективность и мощность на выходе сильно зависят от свойств материала и геометрических характеристик устройства. В последних главах мы рассмотрим усовершенствованные геометрические структуры, которые позволяют получить лазерные диоды непрерывного действия с пороговым током менее 100 мА.

Гетероструктурные лазеры с ограниченным полем

Ранее было показано, что ограничение оптического поля в области лазера, в которой существует инверсная населенность, приводит к существенному снижению пороговой плотности тока и соответствующему возрастанию эффективности. Еще в 1963 г. было предсказано, что гетеропереходы могут использоваться для изготовления волноводных структур с желаемыми свойствами оптического ограничения. Примерно в то же время предсказывалась возможность использования гетеропереходных лазерных структур не

для ограничения оптического поля, а для достижения более высокой эффективности инжекции носителей в p-n переход, а также для ограничения носителей в области перехода. Фактически все эти три механизма представлены в гетероструктурном лазере, и их совместный эффект приводит к созданию устройства, которое значительно превосходит лазер на обычном p-n переходе.

Вследствие технологических трудностей, связанных с выращиванием многослойных гетероструктур, понадобилось несколько лет, прежде чем в 1969 г. были изготовлены действующие гетероструктурные лазеры. Все эти первые устройства были изготовлены на основе $Ga_{1-x}Al_xAs$, так как постоянные решетки GaAs и AlAs близки между собой, что приводит к минимальным поверхностным напряжениям между гетероструктурными слоями и делает указанный материал наиболее подходящим по сравнению с другими, например GaAsP.

Даже в первых таких гетероструктурных лазерах плотность порогового тока составляла порядка 10^3 А/см^2, а не 10^4 - 10^5 А/см^2, как у сравниваемых гомопереходных полупроводниковых лазеров. Начиная с 1969 г. был введен ряд усовершенствований в основной гетероструктур-ный лазер, которые привели к созданию устройства с плотностью порогового тока порядка 10^2 А/см2.

Лазеры с одинарным гетеропереходом (ОГ)

Самым простым для изготовления гетеропереходным лазером является ОГ-структура, представленная на рисунке 10. При изготовлении этого устройства аномально быстрая диффузия Zn в GaAs используется для формирования диффузного p-n перехода, лежащего на 1-2 мкм ниже гетероперехода $Ga_{1-x}Al_xAs$ — GaAs. Если концентрации примесей n- и p-типов примерно одинаковы на обеих сторонах p-n перехода, то инжекционный ток будет существовать главным образом за счет электронов, инжектируемых в слой p-типа, так как эффективная масса электронов почти в 7 раз меньше массы дырок в $Ga_{1-x}Al_xAs$. Таким образом, слой с инверсной населенностью (активный слой) в ОГ-лазерах такого типа находится в материале GaAs p-типа, как показано на рисунке 10.

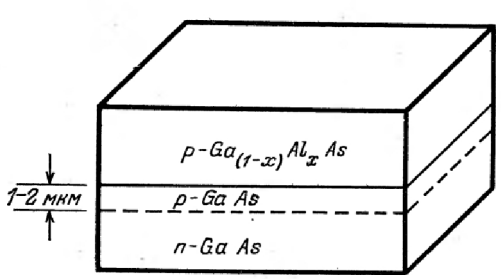


Рисунок 10 - Лазерный диод с одинарной гетероструктурой.

Этот ОГ-лазер можно изготовить, используя метод эпитаксиального выращивания из жидкой фазы, за исключением того, что эпитаксия проводится при относительно высокой температуре (900-1000 °С) для обеспечения требуемой диффузии Zn в подложку. Толщину активного слоя можно выбрать, контролируя время и температуру эпитаксиального роста (следовательно, диффузии Zn) для достижения значений 1-5 мкм. Однако, поскольку длина диффузии инжектируемых электронов составляет лишь около 1 мкм, увеличение толщины слоя p = GaAs свыше этого значения приводит к уменьшению эффективности и более высокой пороговой плотности тока. Это обусловлено тем, что область инверсной населенности все же ограничена толщиной около 1 мкм, где в основном происходит рекомбинация электронов. Таким образом, даже если оптическая мода простирается по всему слою p-GaAs, она может накачиваться стимулированным излучением лишь в слое толщиной 1 мкм, прилегающем к p-n переходу, что снижает эффективность устройства. В некоторых случаях возникает необходимость в увеличении толщины слоя p-GaAs более 1 мкм даже за счет возрастания плотности порогового тока, так как коррекшш дифракции оптического пучка в более толстом слое приводит к малому углу расхождения излучаемого света в плоскости, перпендикулярной p-n переходу.

В ОГ-лазере оптическое ограничение происходит лишь на одной стороне светоизлучающего перехода, а именно на поверхности раздела между слоями p-GaAs и p-$Ga_{1-X}Al_XAs$. Хотя в обедненном слое самого p-n перехода существует волноводный эффект вследствие снижения концентрации носителей, его влияние пренебрежимо мало по сравнению со значительным ограничением, которое имеет место благодаря изменению

коэффициента преломления в гетеропереходе. Таким образом, структура ОГ-лазера лишь частично эффективна в достижении желаемого оптического ограничения. В результате ОГ-лазер обладает более высокой плотностью порогового тока по сравнению с ДГ-лазером. В действительности ОГ-лазеры при комнатной температуре должны работать в импульсном, а не в непрерывном режиме. Во многих случаях применения импульсный режим не вредит работе системы в целом и может быть даже полезным, например для того, чтобы в процессе обработки информации улучшить отношение сигнал - шум. Следовательно, ОГ-лазеры могут широко использоваться как источники световых импульсов в интегрально-оптических устройствах. Однако, когда при комнатной температуре лазерный диод работает в непрерывном режиме, необходимо использовать лазер с двойной гетероструктурой.

Лазеры с двойной гетероструктурой (ДГ-лазеры)

Реальная структура обычного ДГ-лазера представлена на рисунке 11 вместе с профилем показателя преломления в направлении, перпендикулярном плоскости p-n перехода. Основная трехслойная волноводная структура GaAlAs обычно выращивается на сильнолегированной подложке n+-типа и покрывается -сильнолегированным слоем p+-GaAs, чтобы облегчить формирование электрических контактов. Активная область создается на стороне p-типа p-n перехода. Очень часто активный слой должен содержать определенную концентрацию атомов Al, чтобы сдвинуть оптическое излучение в сторону более коротких длин волн.

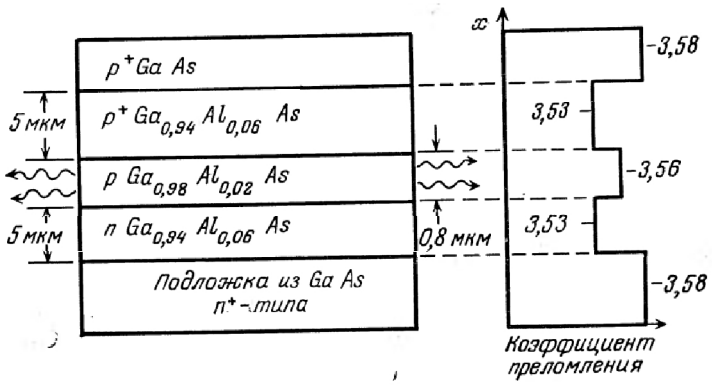


Рисунок 11 - Лазерный диод с двойной гетероструктурой.

Толщина активного слоя должна быть менее 1 мкм, чтобы гарантировать существование инверсной населенности повсеместно внутри этого слоя, а не в слое, ограниченном длиной диффузии инжектируемых электронов. На самом деле толщина активного слоя часто уменьшается до 0,2—0,3 мкм, чтобы достичь больших значений инверсной населенности и плотности генерируемых фотонов. Типичные значения концентрации примесей составляют $N_A \approx 2*10^{19}$ см$^{-3}$ для p+-слоя, $N_A \approx 10^{16}$ см$^{-3}$ для активного слоя, $N_D \approx 10^{17}$ см$^{-3}$ в n-GaAlAs и $N_D \approx 2*10^{18}$ см$^{-3}$ для подложки. Выбор этих концентраций диктуется желанием уменьшить последовательное сопротивление в объеме материала и в то же время ограничить поглощение на свободных носителях в светоизлучающей области. Многослойные $Ga_{1-x}Al_xAs$ структуры, представленные на рисунке 11, обычно получают эпитаксиальным выращиванием из жидкой фазы методом скользящих расплавов.

Лазерные структуры с двойным гетропереходом обеспечивают ограничение с двух сторон активной зоны. Вследствие такого ограничения как фотонов генерируемых мод, так и инжектируемых носителей в области инверсной населенности, в которой возможно усиление за счет стимулированного излучения, ДГ-лазеры являются высокоэффективными устройствами и характеризуются минимальным пороговым током по сравнению с другими п/п лазерами.

Сравнение характеристик излучения лазеров

Пороговая плотность тока для типичного гомопереходного лазера имеет порядок 10^4 А/см^2, а дифференциальная квантовая эффективность (выше пороговой плотности) составляет ~10%. Хотя ОГ-лазер обеспечивает ограничение фотонов и носителей только на одной границе активного слоя, этого вполне достаточно для уменьшения пороговой плотности тока приблизительно до $5*10^3$ А/см^2 и увеличения квантовой эффективности до ~40%. Необходимо напомнить, конечно, что $J_п$ и η_D являются функциями температуры,

толщины активного слоя и концентрации легирующих примесей в любом конкретном случае; поэтому упомянутые здесь числовые значения являются типичными, приведенными лишь для сравнения. Так как величина $J_П$ в ОГ-лазерах все же относительно велика, а величина η_D мала, то они не могут работать в непрерывном режиме, а должны быть пульсирующими, обычно с длительностью импульса 100 нс и периодичностью повторения 100—1000 Гц, чтобы обеспечить время, необходимое для охлаждения. Тем не менее, достижима импульсная выходнвямощность 10 - 30 Вт.

ДГ-лазеры, которые обеспечивают ограничение с обеих границ активной области, имеют типичные пороговые плотности тока 400-800 А/см^2 и дифференциальную квантовую эффективность вплоть до 55%. ДГ-лазеры могут работать в непрерывном режиме с выходной мощностью до 390 мВт, хотя более обычен режим с непрерывной мощностью менее 100 мВт. У лазеров с двойной гетероструктурой имеется лишь один недостаток, заключающийся в большой угловой расходимости луча в плоскости, перпендикулярной плоскости перехода (20-40°), вместо 15-20°, обычно наблюдаемых у ОГ-лазеров. Эту расходимость, возникающую из-за дифракции в тонком активном слое, можно уменьшить использованием специальных ДГ-структур с большим оптическим резонатором, которые обеспечивают раздельное ограничение фотонов и носителей. Для гомопереходных лазеров, а также для ОГ- и ДГ-лазеров с широкой рабочей зоной, о которых говорилось ранее, нет никакого бокового ограничения в плоскости перехода, и свет свободно распространяется по всей ширине чипа (20-80 мкм). В результате во всех случаях расходимость луча в плоскости перехода составляет всего около 10°.

Лазеры с распределенной обратной связью

Работа всех описанных до сих пор лазеров зависит от оптической обратной связи, обеспечиваемой парой отражающих поверхностей, которые формируют эталон Фабри-Перо. В оптической интегральной схеме, в которой лазерные диоды монолитно объединены в составе полупроводниковой пластины, обычно трудно сформировать такие отражающие поверхности. Их можно получить путем травления или скалывания. Однако при этом плоская поверхность подложки прерывается, что затрудняет электрический

монтаж и отвод тепла. Другой подход, использующий распределенную обратную связь (РОС) на дифракционной решетке брэгговского типа, обеспечивает достижение целого ряда преимуществ и использование при этом планарной геометрии.

Теоретическое рассмотрение

В лазерах с распределенной обратной связью (РОС-лазерах) решетка обычно создается путем гофрирования поверхности раздела между полупроводниковыми слоями, образующими лазер. Такой гофр обеспечивает отражение на угол 180° лишь на строго определенных длинах волн в зависимости от периода решетки.

Спектральные характеристики брэгговского отражения

Принципы селективного отражения определенных длин волн можно понять из рассмотрения рисунка 12, который иллюстрирует отражение падающей плоской волны совокупностью отражателей, расположенных друг от друга на расстоянии d. В первоначальном случае дифракции рентгеновских лучей, рассмотренной Брэггом, такими отражателями служили атомные плоскости кристаллической решетки. Однако тот же эффект наблюдается и в случае отражения от гофрированной решетки, сформированной в плоскости перехода полупроводникового лазера, как показано на рисунке 13. Из рисунка 12 очевидно, что для поддержания фазовой когерентности плоского волнового фронта (перпендикулярного показанным лучам) и, следовательно, для предотвращения деструктивной интерференции пути лучей, отраженных от соседних отражателей, должны отличаться на целое число полных длин волн.

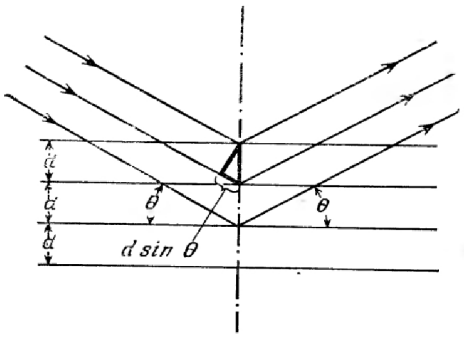


Рисунок 12 – Брэгговское отражение от периодической структуры

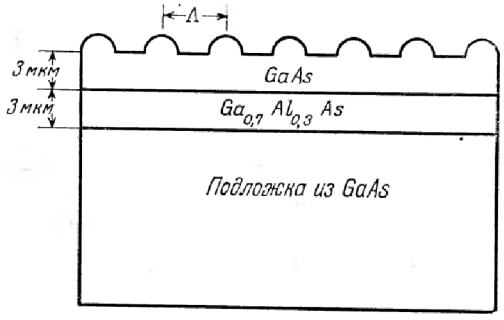


Рисунок 13 – Поперечное сечение оптически накачиваемого РОС-лазера

Таким образом, из геометрических соображений мы получаем соотношение Брэгга:

$$2d \sin \Theta = l\lambda, \quad l = 1,2,3,..., \qquad (24)$$

где Θ - угол между падающим лучом и плоскостью отражателя, λ - оптическая длина волны в среде. Чтобы применить соотношение (24) к случаю отражения решеткой РОС-лазера на угол 180°, необходимо лишь положить d равным периоду решетки Λ, $\lambda=\lambda_0/n_g$ где n_g - эффективный показатель преломления волновода для рассматриваемой моды, а Θ = 90°. Для света, который будет отражаться такой решеткой на угол 180°, длина волны в вакууме будет составлять:

$$\lambda_0 = 2\Lambda n_g / l, \quad l = 1,2,3,..., \qquad (25)$$

Хотя решетка может отражать много различных продольных мод, соответствующих различным значениям *l*, обычно лишь одна мода будет лежать внутри полосы усиления лазера. В действительности из-за трудности изготовления решетки первого порядка (*l*=1) используется решетка третьего порядка.

Доля оптической мощности, отраженной решеткой, зависит от многих факторов, включая толщину волноводного слоя, глубину штриха решетки и длину решеточной области. Определение части отраженной оптической мощности с математической точки зрения является довольно сложной задачей.

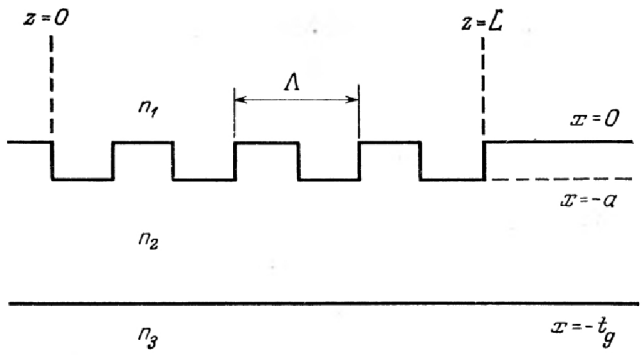


Рисунок 14 – Оптический волновод с прямоугольной решеткой распределенной обратной связи

С использованием теории связанных мод, было показано, что для прямоугольной решетки с глубиной a штриха, связь можно охарактеризовать коэффициентом χ, который определяется формуллой:

$$\chi = \frac{2\pi^2}{3l\lambda_0}\frac{n_2^2 - n_1^2}{n_2}\left(\frac{a}{t_g}\right)^3\left[1 + \frac{3(\lambda_0 / a)}{2\pi\left(n_2^2 - n_1^2\right)^{1/2}} + \frac{3(\lambda_0 / a)^2}{4\pi^2\left(n_2^2 - n_1^2\right)}\right], \quad (26)$$

где t_g – толщина волноводного слоя с показателем преломления n_2; *l* – порядок гармоники, ответственной за связь, определяемый соотношением:

$$l \approx \beta_m \Lambda / \pi , \qquad (27)$$

где β_m – константа распространения рассматриваемой моды.

Для случая первого порядка (l=1) процесс отражения в основном ограничивается связью между прямой и обратной волнами, распространяющимися в направлениях ±z. Тогда амплитуды падающей (A_m+) и отраженной (A_m-) волн являются функциями расстояния z, которое падающая волна проходит под решеткой, и описываются выражениями:

$$A_m^+(z) = A_m^+(0)\frac{ch[\chi(z-L)]}{ch(\chi L)},$$
$$A_m^-(z) = A_m^+(0)\frac{\chi}{|\chi|}\frac{sh[\chi(z-L)]}{ch(\chi L)}, \qquad (28)$$

где L — длина решетки. Падающая и отраженная оптические мощности выражаются через $|A_m+|^2$ и $|A_m-|^2$ соответственно. При достаточно больших аргументах гиперболических функций в выражении (28) падающая оптическая мощность уменьшается экспоненциально с увеличением z, поскольку мощность отражается в волну, распространяющуюся в обратном направлении.

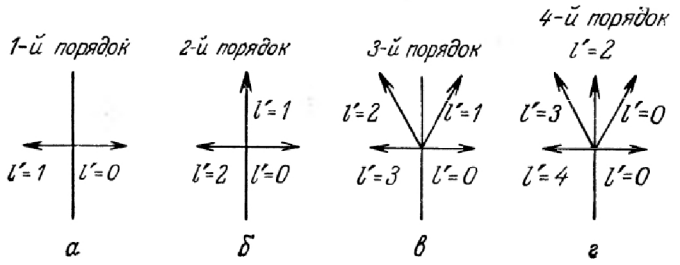


Рисунок 15 - Направления рассеянных волн для брэгговских решеток высокого порядка.

Эффективный коэффициент отражения $R_{эфф}$ и соответствующий коэффициент пропускания $T_{эфф}$ на брэгговской длине волны определяются выражениями:

$$R_{эфф} = \left| A_m^-(0) / A_m^+(0) \right|^2, \qquad (29)$$
$$T_{эфф} = \left| A_m^+(L) / A_m^+(0) \right|^2.$$

Соотношения (28) - (29), строго говоря, применимы лишь для брэгговского отражения первого порядка. Для решеток более высокого порядка, период Λ которых составляет:

$$\Lambda = l\lambda_0 / 2n_g, \quad l > 1, \qquad (30)$$

часть оптической энергии перекачивается в волны, распространяющиеся в направлении от плоскости волновода, как показано на рисунке 15. Эти волны определяют потери энергии для генерируемой моды и, таким образом, уменьшают значения как $R_{эфф}$ так и $T_{эфф}$.

Рассмотрим волну, распространяющуюся в гофрированном волноводе вправо для решетки более высокого порядка (рисунок 16).

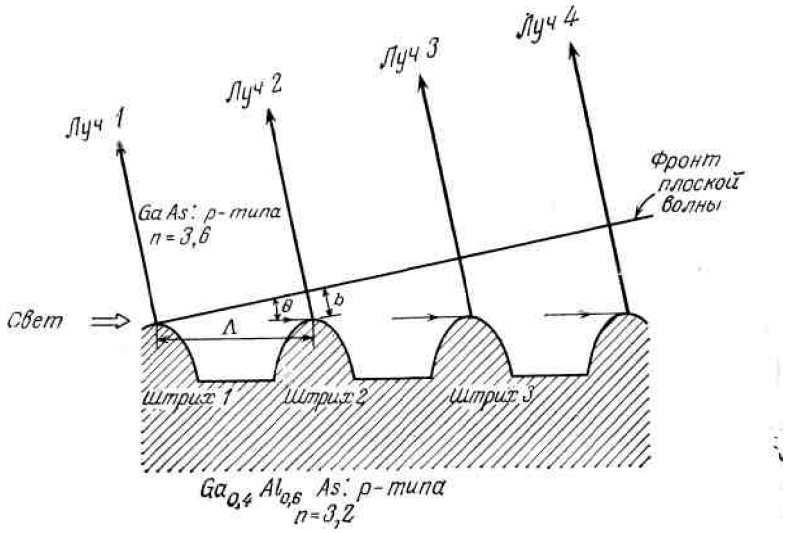


Рисунок 16 - Лучи, рассеянные из волновода брэгговской решеткой высокого порядка.

Чтобы все лучи, рассеянные от каждого последующего штриха, были в фазе, длина пути каждого луча должна составлять целое число длин волн. Таким образом, налагается условие, которое требует, чтобы

$$b + \Lambda = l' \lambda_0 / n_g, \quad l' = 0,1,2,..., \qquad (31)$$

где b - расстояние, показанное на рисунке 16. Из геометрических соображений можно видеть, что:

$$b = \Lambda \sin \Theta, \qquad (32)$$

где Θ — угол между волновым фронтом рассеянной волны и плоскостью волновода. Таким образом, на основе выражений (31) и (32) угол Θ определяется следующим образом:

$$\sin \Theta = \frac{l' \lambda_0}{n_g \Lambda} - 1, \quad l' = 0,1,2,.... \qquad (33)$$

Рассмотрим случай брэгговской дифракции второго порядка, для которого $\Lambda = \lambda_0 / n_g$. В этом случае выражение (33) принимает форму

$$\sin \Theta = l' - 1, \quad l' = 0,1,2,.... \qquad (34)$$

Решение при l'=0 соответствует свету, рассеянному в прямом направлении, в то время как решение при l'=2 – обратному рассеянию, необходимому для РОС. Однако из решения при l'=1 следует Θ=0, что соответствует свету, излучаемому перпендикулярно плоскости волновода.

В основном, сильная связь в поперечном направлении нежелательна для РОС-лазеров. Чтобы обеспечить оптимальную работу, необходима решетка первого порядка.

Если способ изготовления не допускает создания решетки первого порядка, лучше использовать решетку третьего порядка, поскольку связь с волнами, выходящими из плоскости волновода, значительно слабее, чем в решетке второго порядка.

Интегрально-оптический РОС-лазер

В интегрально-оптическом РОС-лазере необходимо вырастить один или несколько эпитаксиальных слоев но поверхности волновода, после того как на ней сформирована решетка. Дефекты кристаллической решетки, образующиеся в процессе создания гофра, будут определять качество эпитаксиальных слоев, поскольку точечные дефекты и дислокации имеют тенденцию к распространению из области решетки в эпитаксиальные слои. Эти дефекты действуют как центры оптического поглощения и безызлучательной рекомбинации. Таким образом, они являются источником уменьшения квантовой эффективности и увеличения пороговой плотности тока.

Очень часто используемый подход для устранения влияния искажений кристаллической решетки заключается в физическом разделении гофра и активного слоя лазера. Например, изготовлен инжекционный РОС-лазер на p-n переходе с использованием процесса диффузии, которой разделяет p-n переход и решетку, как показано на рисунке 17.

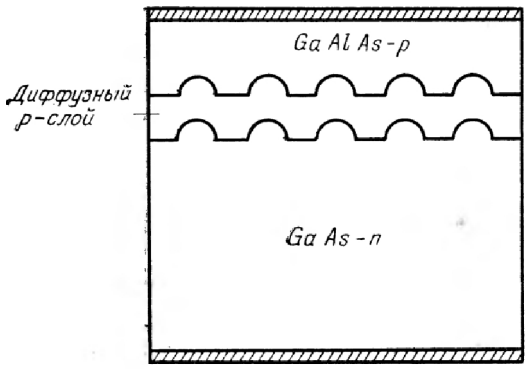


Рисунок 17 - РОС-лазер с диффундированным переходом

Сначала подложка из n-GaAs была гофрирована с помощью ионно-лучевого травления для создания решетки третьего порядка с периодом 350 нм. Затем на подложке был выращен слой p-GaAlAs. В качестве легирующей добавки использовался яинк, благодаря его высокой диффузии в GaAs. Таким образом, в

процессе роста слоя GaAlAs цинк диффундировал в подложку и на глубине около 1 мкм под гофрированной поверхностью раздела GaAs – GaAlAs формировался p-n переход. Получившийся в результате одногетероструктурный РОС-лазер имел пороговые плотности, в основном равные плотностям подобных ОГ-лазеров, со сколотыми торцами. Таким образом, можно сделать вывод, что влияние дефектов, образующихся при изготовлении решетки, было не сильным.

Другой подход к разделению решетки и p-n перехода показан на рисунке 18. Здесь использован гетеропереход с раздельным ограничением в совокупности с решеткой Брэгга третьего порядка в полосковом лазере шириной 50 мкм. Длина активной области составляла 700 мкм. Ненакачиваемый волновод длиной 2-3 мм являлся непрерывным продолжением активной зоны и служил для изоляции поверхности заднего скола, гарантируя тем самым, что генерация происходила лишь за счет распределенной обратной связи, а не отражения от зеркала.

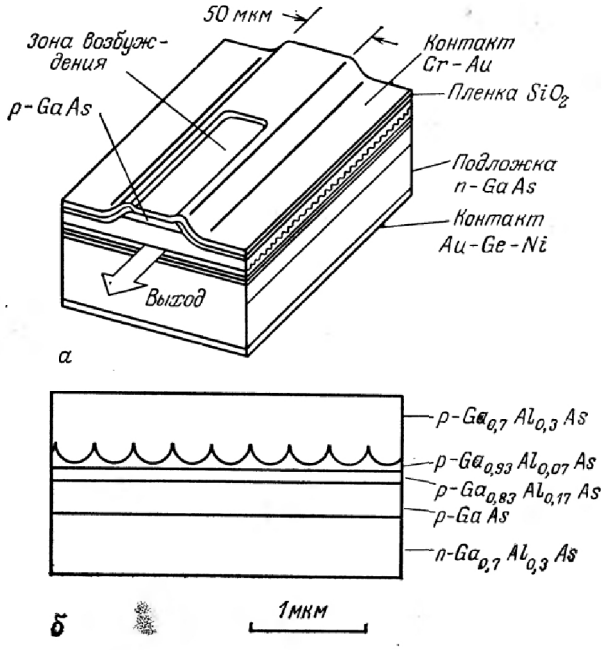


Рисунок 18 - Полосковый РОС-лазер с раздельным ограничением оптического поля и носителей

В структуре, представленной на рисунке 18, инжектируемые электроны ограничены в активном слое толщиной 0,2 мкм слоем p-$Ga_{0,83}Al_{0,17}As$ толщиной 0,1 мкм, а фотоны оптической моды достигают поверхности раздела со слоем p-$Ga_{0,93}Al_{0,07}As$. Таким образом, оптическая мода взаимодействует с периодической структурой, а активная область, в которой происходит рекомбинация, находится достаточно далеко от решетки, поэтому безызлучательная рекомбинация из-за дефектов кристаллической решетки минимизирована. Устройства подобного типа работали в непрерывном режиме при комнатной температуре при комнатной температуре при плотностях тока около $3,5*10^3$ А/см^2. При периоде 368,4 нм наблюдалась одна продольная мода 846,4 нм.

РБЗ-лазер

Активную зону можно также изолировать от области решетки использованием структуры с распределенными брэгговскимн зеркалами (РБЗ), показанной на рисунке 19. В таком устройстве используются две брэгговские решетки, которые расположены по обе стороны лазера вне электрически накачиваемой активной зоны. Кроме исключения безызлучательной рекомбинации, обусловленной кристаллическими дефектами, размещение двух решетчатых зеркал вне активной зоны позволяет делать их неодинаковыми и создавать лазер с выходом через один торец. Для достижения эффективной работы в одной продольной моде один из распределенных отражателей должен иметь узкую полосу и высокий коэффициент отражения на генерируемой длине волны, а другой – сравнительно низкую отражательную способность для обеспечения оптимальной связи на выходе.

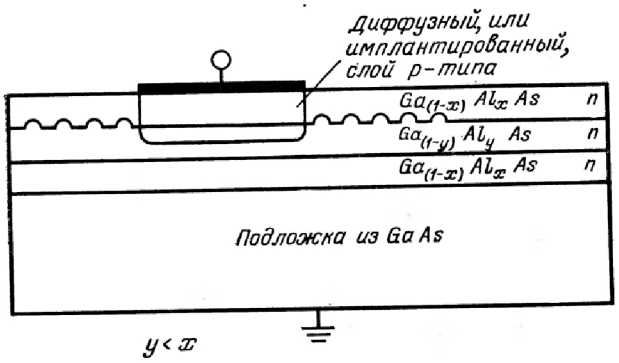


Рисунок 19 - Интегрально-оптический РБЗ-лазер.

Благодаря удобству обеспечения однонаправленного выхода на заранее определенной длине волны, зависящей от периода решетки, РБЗ-лазеры хорошо подходят для решения задач частотного мультиплексимрования. РБЗ-лазеры имеют плотности порогового тока (3-6)*10^3 А/$см^2$ и спектральную ширину 0,03 нм.

Лазеры с туннельной инжекцией

Как следует из названия, эти устройства накачиваются током электронов или дырок, которые достигают активной области путем туннелирования через энергетический барьер. Лазеры с туннельной инжекцией не приобрели большой известности, так как до сих пор не смогли изготовить ни одного устройства с достижимой плотностью тока. Тем не менее весьма полезно рассмотреть лазер с туннельной инжекцией, так как он является первым предложенным лазером с ограничением поля и служит основой для простой модели, которая объясняет, как такое ограничение поля, используемое в современных лазерах на гетеропереходах, приводит к существенному снижению пороговой плотности тока и более высокой эффективности.

Основная структура

Полупроводниковым лазером, который в принципе объединяет лучшие характеристики лазеров на p-n переходах (малые размеры, простота, низкое рабочее напряжение) с лучшими

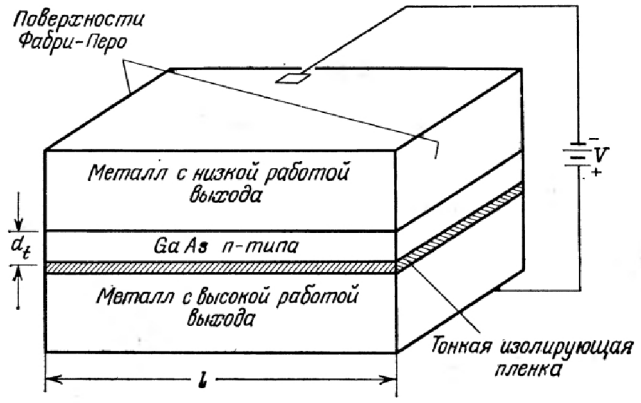


Рисунок 20 – Основная структура лазера с туннельной инжекцией

характеристиками лазеров, накачиваемых пучком электронов (единый тип полупроводникового материала, отсутствие необходимости формировать переход), является лазер с туннельной инжекцией, показанный на рисунке 20. В этом лазере используется единый кристалл из однородно-легированного примесями полупроводникового материала без перехода. Электронно-дырочные пары инжектируются в полупроводник туннелированием и диффузией. Если используется полупроводник p-типа, то электроны инжектируются через изолятор путем туннелирования, в то время как дырки диффундируют от источника в металле 1 к полупроводнику. Если используется полупроводник n-типа, то дырки туннелируют через изолятор, в то время как электроны инжектируются в металл 1.

Процесс туннелирования можно лучше понять, если рассмотреть рисунок 21.

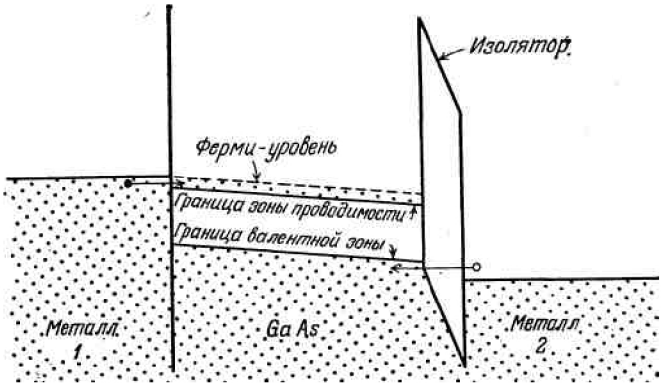


Рисунок 21 - Энергетические зоны лазера с туннельной инжекцией с прямым напряжением смещения.

Рисунок показывает картину энергетических зон после приложения требуемого напряжения смещения, в результате чего ферми-уровни сдвинуты к новому положению. Заметим, что большая часть напряжения падает на изоляторе; следовательно, в этой области существует очень высокое электрическое поле.

Из схемы можно видеть, что дырки в металле 2 не имеют достаточной энергии (говоря языком классической физики), чтобы преодолеть энергетический барьер. Однако квантовоме-ханическая теория предсказывает существование малой вероятности того, что дырки могут проникать через изолятор в полу-проводник, даже если, согласно классической теории, они не имеют достаточно энергии для преодоления барьера. Чтобы через барьер туннелировало значительное число дырок, изолятор должен быть очень тонким или напряжение очень большим, поскольку плотность туннельного тока описывается выражением:

$$J_t = J_0 (E / E_0)^2 e^{-E_0 / E}, \qquad (35)$$

где E — напряженность электрического поля в изоляторе, E_0 и J_0 - постоянные, задаваемые выражениями

$$E_0 = 4\varphi^{3/2} (2m^*)^{1/2} / 3hq, \qquad (36)$$

$$J_0 = 2q\varphi^2 m^* / 9h\pi^2.$$

где φ - высота барьера, а другие величины были определены ранее.

Чтобы достичь необходимой величины тока для лазерного порога (порядка 100 А/см^2), как показывают расчеты, необходимо достичь напряженности 10^7 В/см. Таким образом, если использовать 10 В источник напряжения, толщина изолятора должна составлять около 0,01 мкм. До сих пор все попытки реализовать лазер с туннельной инжекцией потерпели неудачу, так как очень трудно сделать сверхтонкую пленку без микропор. Тем не менее на основе принципа туннельной инжекции были сделаны светоизлучатели в материалах GaAs и CdS. В случае CdS туннельная инжекция особенно важна, так как еще никто не мог добиться p-n перехода в CdS. Поскольку для светоизлучателей требуемый ток меньше необходимого для лазерного порога, можно использовать более толстые и более практичные пленки (~ 100 нм), и это вполне реализуемо.

Вопросы надежности

Важной характеристикой лазерных диодов, которая сильно влияет на конструкцию системы, является надежность устройства. Лазерные диоды являются высоконапряженными с точки зрения существующих в них плотностей тока и напряженностей оптического поля. Обычно им свойственны постепенное ухудшение рабочих характеристик по мере старения и даже (в некоторых условиях) катастрофическое разрушение.

Мгновенное разрушение

Лазерные диоды подвержены катастрофическому разрушению, когда оптическая плотность мощности на торцевых зеркалах превышает некоторый критический уровень. Механическое повреждение в виде углублений и выступов обычно появляется прежде всего вблизи центра грани, где напряженность оптического поля максимальна. Порог плотности мощности, выше которого происходит

катастрофическое повреждение, у различных лазеров изменяется благодаря тому факту, что присутствие первоначальных дефектов роста на торцевой поверхности увеличивает вероятность последующего повреждения. Типичные пороговые плотности мощности для лазеров с гомопереходом при температуре 300К составляют 6 - 8 мВт/см^2 . Гетероструктурные лазеры также подвержены катастрофическому разрушению при оптических мощностях, превышающих некоторые пороговые значения. Сообщается о пороговом разрушении при плотностях в диапазоне 2-4 мВт на 1 мкм ширины полоски лазеров на двойном гетеропереходе на основе GaAlAs с d = 0,2-0,3 мкм.

При работе лазерного диода в импульсном режиме по сравнению с непрерывным пороговый уровень в основном повышается. В случае импульсов длиной 1 нс установлено , что для типичных лазеров с одной гетероструктурой плотность мощности P_c составляет 80 мВт/мкм, а для более длинных импульсов пороговый уровень убывает, изменяясь обратно пропорционально квадратному корню длины импульса. Порог повреждения можно также повысить путем нанесения на грани просветляющих покрытий с целью снижения различия в напряженности оптического поля внутри и вне лазерной полости, однако это также ведет к повышению плотности порогового тока, необходимого для появления лазерной генерации /$_{п}$.

Катастрофического разрушения можно, конечно, избежать, работая с лазером просто ниже порога разрушения. Однако необходимо отметить, что переходные процессы длительностью несколько микросекунд достаточны для разрушения лазеров. Следовательно, переходные процессы при включении должны отфильтровываться в непрерывном режиме работы, а в импульсном режиме необходимо исключить переходные процессы типа «звон». Кроме того, фильтры источника питания должны быть способны поглотить случайные переходные процессы, которые могут возникнуть в линиях.

Постепенное ухудшение характеристик

Проблема постепенного ухудшения характеристик, при которой по мере работы устройства плотность порогового тока лазера возрастает, а дифференциальная квантовая эффективность убывает, была так сложна в первых лазерных диодах, что срок службы устройств составлял лишь 1-100 ч. В то время полагали, что практически никогда не удастся создать надежного устройства. Тем не менее, методические работы в различных лабораториях в течение нескольких лет привели к выявлению и исключению большинства механизмов, ухудшающих рабочие характеристики, поэтому в настоящее время изготовляются лазеры с проектируемым сроком службы 10^6 ч.

Постепенное ухудшение характеристик лазерных диодов происходит вследствие образования дефектов внутри активной зоны, которые действуют как безызлучательные рекомбинационные центры. Точное происхождение этих дефектов не установлено. Тем не менее были предложены некоторые возможные механизмы, включая осаждение примесных атомов в дислокациях, релаксацию напряжений возникавших при монтаже, за счет генерации дислокаций, оптическую генерацию дефектов в напряженном материале и диффузию дефектов при световом воздействии. По мере ухудшения характеристик дефекты становятся видимыми как дефекты темных линий или трехмерные дислокационные сети, которые растут по величине и количеству во время лазерной генерации. Дефекты и загрязняющие атомы, исходно присутствующие в лазере, содействуют ускорению роста дефектов темных линий. Следовательно, поддержание большинства из разработанных процедур роста является существенным при изготовлении лазеров с большим сроком службы. Диоды на $Ga_{1-x}Al_xAs$ с величиной x в пределах 0,05-0,1 в активной области характеризуются наименьшим ухудшением. Так как дислокации имеют тенденцию к распространению из подложки в эпитаксиальные слои, важен выбор подложек с малой плотностью дислокаций.

Особое внимание к деталям процесса изготовления, включая выбор соответствующей подложки, сверхвысокую чистоту выращивания эпитаксиальных слоев, тщательный контроль за контактной металлизацией и

крепежом, может привести к созданию диодов с довольно большим сроком службы. Например, сообщается лазере со сроком службы более двух лет и о снижении его выходной мощности менее чем на 15% при постоянном токе после одного года работы. Лазер на GaAlAs с двойной гетероструктурой имел полосковую геометрию, сформированную бомбардировкой протонами. Из статистического анализа тестов ускоренного старения, проведенных на этих устройствах при 70 °С, авторы оценили среднюю наработку на отказ при 22°С, равную $1,3*10^6$ ч (т. е. более 100 лет). Аналогичные результаты получены для лазера на $Al_{0,3}Ga_{0,7}As/Al_{0,08}Ga_{0,92}As$ с двойной гетероструктурой и $A1_20_3$-покрытием торцов. Ускоренные тесты на старение их устройств привели к выводу, что экстраполированный срок службы составляет около 10^6 ч.

Лекция 4

Классификация и основные параметры приемников излучения интегрально-оптических схем. Фотодиод на p-n переходе. Волноводный, лавиный, p-i-n фотодиод. Фотоприемники с гетеропереходом, на основе МДП-структуры. Изменение спектральных характеристик.

Фотодиоды с обедненным слоем

Наиболее распространенным типом полупроводникового оптического детектора, который используется в интегральной оптике, а также в качестве отдельного прибора, является фотодиод с обедненным слоем. Он является по существу полупроводниковым диодом с обратным смещением, в котором обратный ток модулируется электронно-дырочными парами, образующимися в обедненном слое или около него в результате поглощения фотонов света. Диод в основном работает в фотодиодном режиме при сравнительно большом напряжении смещения, а не в фотогальваническом, при котором сам диод является электрическим генератором, и на него напряжение смещения не подается.

Простейшим типом фотодиода с обедненным слоем является диод на p-n переходе. Схема энергетических зон такого устройства с обратным напряжением смещения V_a представлена на рис унке 1. Полный ток фотодиода с обедненным слоем состоит из двух компонент: дрейфовой, образующейся носителями, генерированными в области б, и диффузионной, образующейся в областях а и в. Дырки и электроны, образовавшиеся в области б, разделяются полем обратного смещения, причем дырки уносятся в p-область в, а электроны — в n-область а. Дырки, образующиеся в n-области, или электроны, образующиеся в p-области, имеют определенную вероятность диффузии к краю обедненной области б, где они переносятся через нее полем. Основные носители (электроны в а-области или дырки в в-области) удерживаются в соответствующих областях напряжением обратного смещения и не переносятся через обедненный слой.

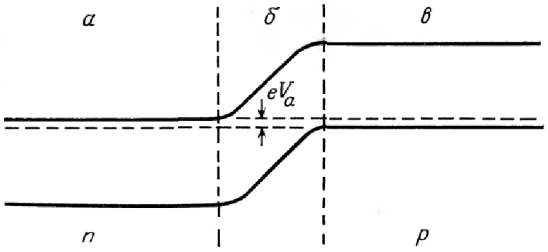


Рисунок 1 - Схема энергетических зон диода на p-n переходе при подаче обратного смещения V_a.

В практических фотодиодах, чтобы минимизировать последовательное сопротивление, сохранив при этом максимальную ширину обедненной области, обычно одна область легируется значительно сильнее другой. В этом случае обедненный слой формируется почти полностью на менее легированной стороне перехода, как показано на рисунке 2.Такое устройство называется асимметричным резким переходом. В GaAs и его трехкомпонентных сплавах подвижность электронов в основном значительно выше подвижности дырок. Поэтому p-область обычно делается тоньше и легируется значительно сильнее, чем n-область, так что устройство формируется в основном в материале n-типа, а p-область тогда служит по существу лишь контактным слоем.

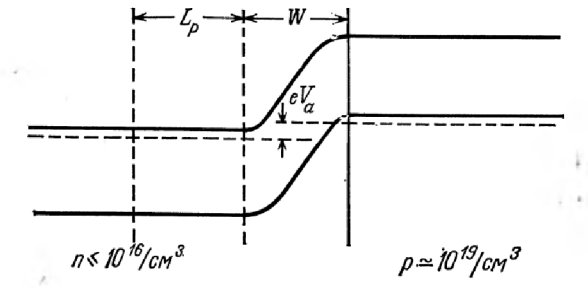


Рисунок 2 - Схема энергетических зон диода на p+-n (асимметричном)-переходе при подаче обратного смещения V_a.

Можно показать, что для устройства с асимметричной геометрией перехода, показанного на рисунке 2, полная плотность тока $J_{полн}$ выражается следующим образом:

$$J_{полн} = q\varphi_0\left(1 - e^{-\alpha W}/\left(1 + \alpha L_p\right)\right) + qp_{n0}\left(D_p / L_p\right), \quad (1)$$

где φ_0 - полный поток фотонов [в фотон/(см^2*с)], W - ширина обедненного слоя, q - величина заряда электрона, α - коэффициент оптического межзонного поглощения, L_p - длина диффузии дырок, D_p - коэффициент диффузии для дырок, p_{n0} - равновесная плотность дырок. Последний член выражения (1) представляет собой обратный ток утечки (или темновой ток), являющийся результатом термической генерации дырок в n-материале. Это объясняет, почему этот член не пропорционален потоку фотонов φ_0. Первый член выражения (1) соответствует фототоку, который пропорционален φ_0 и включает в себя как ток дрейфа носителей внутри обедненного слоя, так и диффузию и дрейф дырок в пределах длины диффузии L_p от края обедненного слоя. Квантовая эффективность детектора η_q, или число носителей, генерируемых одним падающим фотоном, задается выражением:

$$\eta_q = 1 - e^{-\alpha W}/\left(1 + \alpha L_p\right), \quad (2)$$

которое может принимать любое значение от нуля до единицы. Необходимо заметить, что выражения (1) и (2) основываются на молчаливом предположении, что потери на рассеяние и поглощение свободными носителями пренебрежимо мало. Из выражения (2) можно видеть, что для максимизации величины η_q желательно сделать произведения αW и αL_p возможно большими. Когда произведения αW и αL_p настолько велики, что $\eta_q \approx 1$, то диодный ток по существу пропорционален φ_0, поскольку темновой ток обычно пренебрежимо мал. Если коэффициент межзонного поглощения α.слишком мал по сравнению с W и L_p,

многие из падающих фотонов пройдут полностью через активные слои диода в подложку, как показано на рисунке 3.

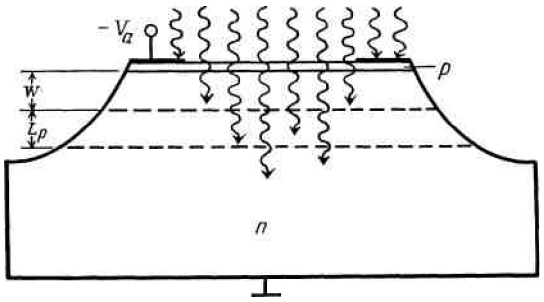


Рисунок 3 - Схема обычного фотодиода мезогеометрии на p+-n переходе, показывающая проникновение фотонов.

Только фотоны, поглощенные в обедненном слое толщиной W, имеют максимальную эффективность генерации носителей. Фотоны, поглощенные на глубинах вплоть до длины диффузии L_p от края обедненного слоя, в некоторой степени эффективны для генерации фотоносителей; в этой области дырки могут диффундировать в обедненный слой. Фотоны, которые проникают на глубину более $W+L_p$, не будучи поглощенными, по существу потеряны для процесса фотогенерации, поскольку имеют очень низкую статистическую вероятность создания дырки, которая могла бы достигнуть обедненного слоя и быть перенесенной через него. В полупроводнике поток фотонов φ(x) уменьшается экспоненциально с увеличением расстояния x от поверхности, как показано на рисунке 4, а. Следовательно, если значение α не достаточно велико, то многие фотоны проникнут слишком глубоко, прежде чем будут поглощены, образуя таким образом носители, которые (в среднем) рекомбинируют прежде, чем диффундируют на расстоянии, достаточном для достижения обедненного слоя.

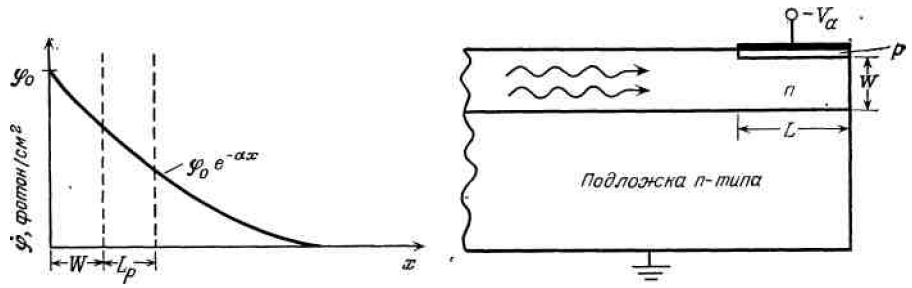


Рисунок 4 – а) зависимость оптического поглощения от расстояния до поверхности в обычном мезофотодиоде; б) схема волноводного фотодетектора

В полупроводниках межзонное поглощение сильно зависит от длины волны. Коэффициент поглощения а обычно имеет спектральную характеристику, которая резко возрастает на длине волны края поглощения (зоны), затем насыщается на длине волны немного короче длины волны, соответствующей ширине запрещенной зоны, медленно увеличиваясь к более коротким длинам волн.

Следовательно, невозможно создать диод с идеальной шириной W для всех длин волн. Для длин волн вблизи края поглощения длинноволновая часть характеристики диода ограничивается повышенным проникновением фотонов в подложку, как показано на рисунках 3 и 4, а; характеристику в коротковолновой части может ограничить слишком сильное поглощение фотонов в p+-слое вблизи поверхности, где вероятность рекомбинации велика.

Помимо уменьшения квантовой эффективности в результате плохого подбора величин α, W и L_p существуют некоторые другие также важные ограничения на характеристики фотодиода с обедненным слоем. Поскольку величина W обычно относительно мала (в пределах 0,1—1,0 мкм), емкость перехода может ограничить высокочастотную характеристику через обычную постоянную времени RC. Время, требуемое для диффузии носителей с глубин между W и $W+L_p$, может также ограничить высокочастотную характеристику обычного фотодиода. Волноводный фотодиод с обедненным слоем, который обсуждается в следующем разделе, существенно смягчает многие из этих проблем обычных фотодиодов.

Волноводные фотодиоды

Если исходный фотодиод с обедненным слоем объединить с волноводной структурой, как показано на рисунке 4, б, то можно добиться улучшений целого ряда характеристик. В этом случае свет падает на активный объем детектора в поперечном направлении, а не по нормали к плоскости перехода. Тогда плотность фототока диода описывается выражением

$$J = q\varphi_0\left(1 - e^{-\alpha L}\right), \quad (3)$$

где L - длина детектора в направлении распространения света. Поскольку W и L - два независимых параметра, концентрацию носителей в объеме детектора и напряжение смещения V_a можно выбрать таким образом, чтобы толщина обедненного слоя была равна толщине волновода, а L может быть сделана столь

длинной, сколько необходимо для того, чтобы $\alpha L >> 1$. Таким образом, простой подгонкой длины L можно достичь 100%-ной квантовой эффективности для любой величины α. Например, для материала с относительно небольшой величиной $\alpha = 30$ см$^{-1}$ длина L = 3 мм обеспечивает $\eta_q = 0{,}99988$. (Вновь в выражении (3) мы предполагаем, что потери на рассеяние и поглощение свободными носителями пренебрежимо малы).

Поскольку волноводный детектор можно сформировать в узком канальном волноводе, емкость может быть очень малой, даже если величина L относительно велика. Например, для материала с относительной диэлектрической постоянной $\varepsilon=12$, например GaAs, детектор длиной 3 мм, сформированный в канальном волноводе шириной 3 мкм, имеет емкость лишь 0,32 пФ. Эта емкость приблизительно в 10 раз меньше, чем у типичных используемых мезофотодиодов. Следовательно, можно ожидать, что высокочастотные характеристики соответственно улучшатся. Экспериментально продемонстрированная ширина полосы 2 ГГц была получена с волноводными детекторами, имеющими площадь $1{,}23*10^4$ мкм^2 на материале подложки GaAs. Благодаря тому что все падающие фотоны поглощаются непосредственно внутри обедненного слоя волноводного фотодетектора, не только улучшается значение η_q, но и устраняется временная задержка, связанная с диффузией носителей. Это приводит к дальнейшему улучшению высокочастотных характеристик.

Благодаря многим улучшениям характеристик, заложенным в поперечной структуре волноводного детектора по сравнению с аксиальной геометрией обычного мезофотодиода, волноводные детекторы следует рассматривать для случаев применения, в которых требуется использование отдельных приборов, а также в составе оптических интегральных схем. В настоящее время волноводные детекторы отдельно промышленностью не выпускаются. Однако их можно сравнительно легко изготовить во многих лабораториях. Следовательно, их наличие не будет составлять проблему.

Влияние рассеяния и поглощения свободными носителями

В выражениях (1) - (3) мы пренебрегли влиянием поглощения свободными носителями и рассеянием фотонов на квантовую эффективность детектора. Поскольку оба этих механизма дают потери фотонов без генерации новых носителей, они имеют тенденцию уменьшать квантовую эффективность. Во многих случаях ими можно пренебречь, и выражения (1) - (3) дадут точные результаты. Однако, когда коэффициентом поглощения свободными носителями α_{FC} и коэффициентом потерь на рассеяние α_S нельзя пренебречь по сравнению с коэффициентом межзонного поглощения α_{IB}, требуется более сложное выражение для η_q. Это выражение можно вывести следующим образом. Предположим, что поток фотонов в произвольной точке на расстоянии x от поверхности детектора, на которую первоначально падают фотоны, задается выражением:

$$\varphi(x) = \varphi_0 e^{-\alpha x}, \quad (4)$$

где коэффициент полных потерь a определяется как сумма:

$$\alpha = \alpha_{IB} + \alpha_{FC} + \alpha_S, \quad (5)$$

Скорость генерации электронно-дырочных пар G(x) составляет:

$$G(x) = \alpha_{IB} \varphi_0 e^{-\alpha x}, \quad (6)$$

поскольку лишь α_{IB} дает вклад в генерацию носителей. Таким образом, плотность фототока определяется выражением:

$$J = q\int_0^L G(x)dx, \quad (7)$$

или

$$J = q\varphi_0 [\alpha_{IB} / (\alpha_{IB} + \alpha_{FC} + \alpha_S)] \left(1 - e^{-(\alpha_{IB}+\alpha_{FC}+\alpha_S)L}\right), \quad (8)$$

Сравнивая выражения (8) и (3), с очевидностью получаем, что эффект дополнительных потерь из-за рассеяния и поглощения свободными носителями заключается в уменьшении квантовой эффективности в α_{IB}/α раз, даже когда длина L выбрана достаточно большой для максимизации величины η_q.
Если значения α_S и α_{FC} малы по сравнению с α_{IB}, что в основном верно, выражение (8) переходит в выражение (3). Однако если волновод неоднороден или имеет необычно грубые поверхности или если объем детектора сильно легирован, так что значениями α_S и α_{FC} пренебречь нельзя, то необходимо использовать выражение (8).

Специальные фотодиодные структуры

Существуют две весьма пригодные фотодиодные структуры, которые можно изготовить либо в волноводной, либо в обычной, неволноводной форме. Ими являются фотодиоды на основе барьера Шоттки и лавинные фотодиоды.

Фотодиоды на основе барьера Шоттки

Фотодиод на основе барьера Шоттки является простым фотодиодом с обедненным слоем, в котором p-n переход заменен запирающим (блокирующим) контактом металл - полупроводник. Например, если бы в устройствах, представленных на рисунках 3 и 4, слои p-типа были заменены металлом, который формирует запирающий контакт для полупроводника, получился бы фотодиод на основе барьера Шоттки. Фототок также задавался бы выражениями (1) и (3), и приборы имели бы по существу те же самые рабочие характеристики, что и их аналоги на p+-n переходах. Схемы энергетических зон диода на основе барьера Шоттки при нулевом и обратном смещении приведены на рисунке 5.

Видно, что обедненная область простирается в материал n-типа, так же как и в случае p+-n перехода. Высота барьера φ_B зависит от использования комбинации металл - полупроводник. Типичная величина φ_B составляет около 1 В.
В обычных мезоприборах часто используется тонкий оптически прозрачный контакт Шоттки (а не p+-n переход) для увеличения отклика в коротковолновой области благодаря устранению сильного поглощения этих фотонов более высоких энергий, которое происходит в p+-слое. В волноводных фотодиодах нет необходимости в контакте Шоттки для улучшения характеистик в области коротких длин волн, поскольку фотоны проникают в активную область в поперечном направлении. Однако легкость в изготовлении часто делает фотодиод на основе барьера Шоттки наилучшим вариантом в интегрально-оптических приложениях. Например, почти любой металл (кроме серебра), будучи напыленным на GaAs, создает запирающий барьер Шоттки при комнатной температуре. Часто используются золото, алюминий или платина. Фоторезистивная маска вполне при годна для ограничения поверхностных размеров в процессе напыления, и не требуется строгого контроля времени и температуры, как в случае диффузии мелких p+-слоев.

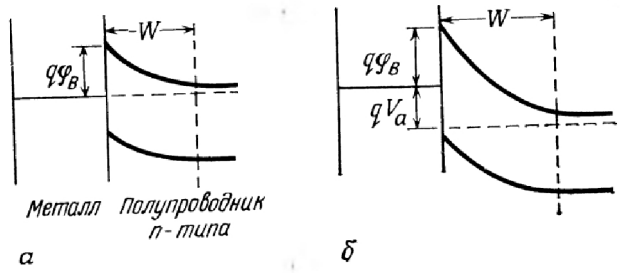


Рисунок 5 - Схемы энергетических зон диода на основе барьера Шоттки.
а — нулевое смещение; б — обратное смещение напряжением V_a.

Лавинные фотодиоды

Коэффициент усиления (т. е. квантовая эффективность) фотодиода на основе либо p-n перехода, либо барьера Шоттки может по крайней мере равняться единице в обычных условиях обратного смещения. Однако, если прибор смещен точно в точку лавинного пробоя, умножение носителей, обусловленное ударной

ионизацией, может обеспечить существенный выигрыш в смысле увеличения отношения числа носителей к числу фотонов. В самом деле, лавинные коэффициенты усиления вплоть до 10^4 не являются чем-то необычным. Типичные вольт-амперные характеристики лавинного фотодиода показаны на рисунке 6. Верхняя кривая вычерчена для условий затемнения, а нижняя отражает влияние освещения. При сравнительно небольших напряжениях обратного смещения диод имеет темновой ток насыщения I_{d0} и фототок насыщения I_{ph0}. Однако, если его сместить в точку лавинного пробоя, умножение носителей в результате приведет к увеличению темнового тока I_d, так же как и увеличение фототока I_{ph0}. Можно определить коэффициент фотоумножения M_{ph} в виде:

$$M_{ph} = I_{ph} / I_{ph0}, \qquad (9)$$

а коэффициент умножения M в виде:

$$M_{ph} = (I_{ph} + I_d)/(I_{ph0} + I_{d0}), \qquad (10)$$

В области смещений, в которой происходит лавинный пробой, трудно вывести точное уравнение для вольтамперной характеристики (рисунок 6).

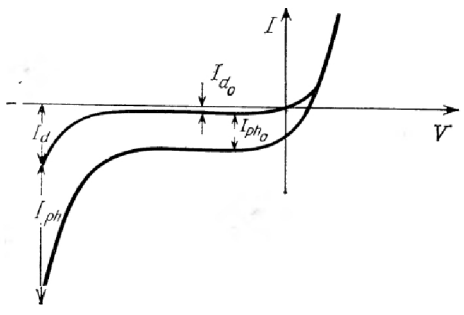


Рисунок 6 - Вольтамперные характеристики лавинного фотодиода.

Однако функциональная зависимость коэффициента умножения может даваться выражением:

$$M_{ph} = 1/\left[1-(V_a/V_b)^n\right], \quad (11)$$

где V_b — напряжение пробоя, n - эмпирически определенный показатель степени, зависящий от длины волны света, концентрации легирующей добавки и, разумеется, от типа материала полупроводника, из которого сделан диод. Для случая большого фототока $I_{ph0} >> I_{d0}$ коэффициент умножения описывается формулой:

$$M = 1/\left[1-((V_a - IR)/V_b)^n\right], \quad (11)$$

где I - полный ток, определяемый суммой:

$$I = I_d + I_{ph}, \quad (12)$$

а R - последовательное сопротивление диода (включая сопротивление пространственного заряда, если оно существенно). При выводе выражения (11) предполагалось, что $IR << V_b$. Для случая когда I_{d0} и I_d пренебрежимо малы по сравнению с I_{ph0} и I_{ph}, можно показать, что максимально достижимый коэффициент умножения описывается выражением:

$$M \approx M_{ph} \approx \sqrt{V_b / nI_{ph0}R}. \quad (13)$$

Лавинные фотодиоды чрезвычайно пригодны для использования в качестве детекторов не только потому, что способны обеспечивать высокие коэффициенты усиления, но и потому, что они могут работать на частотах выше 10 ГГц. Однако не всякий диод на основе p—n перехода или барьера Шоттки может работать в режиме лавинного умножения, будучи смещенным к лавинному пробою.

Например, для создания лавинного пробоя в GaAs требуется поле величиной приблизительно $4*10^5$ В/см. Следовательно, для типичной ширины обедненного слоя 3 мкм V_b=l20 В и многие диоды на GaAs будут пробиваться при значительно меньших напряжениях из-за наличия других механизмов, например краевого пробоя или генерации микроплазмы на локальных дефектах, никогда не достигнув, таким образом, условий лавинного пробоя. Чтобы изготовить лавинный фотодиод, необходимо соблюдать тщательную осторожность, начиная с заготовки полупроводникового материала для подложки, свободного от дислокаций. Лавинные фотодиоды относятся к сильнонапряженным устройствам. Следовательно, основное внимание необходимо уделить надежности. Возрастающий ток утечки, обусловленный плохой дезактивацией поверхности, или генерация внутренних дефектов в процессе работы в режиме сильноточных импульсов могут привести к ухудшению характеристик с увеличением срока службы прибора. Тем не менее для устройств, тщательно изготовленных и герметично упакованных в соответствующие корпуса, наблюдалось среднее время наработки на отказ до 10^5 ч при температуре 170°С, что предполагает около 10^9 ч работы при комнатной температуре.

Фотодиоды на основе МДП-структуры

МДП-фотодиод или фоточувствительная МДП-структура, используется в разных электрических режимах. Одна из таких возможностей характерна для фото-ПЗС, где излучение преобразуется в заряд неосновных носителей, который накапливается у границы раздела полупроводник - диэлектрик и при необходимости считывается (переносится вдоль поверхности) благодаря; механизму зарядовой связи. Тот же заряд модулирует проводимость приповерхностной области (МДП-канал), что ведет к изменению выходного сигнала МДП-транзистора или сопротивления МДП-фоторезистора. Зарядовый обмен между полупроводником и диэлектриком вызывает поляризацию диэлектрика, сохраняющуюся длительное время и после прекращения действия

излучения. Происходит «запоминание» светового сигнала, при оптических или электрических воздействиях записанную информацию можно считывать.

Во всех этих случаях МДП-фотодиод выступает как составная часть более сложного прибора. Фотодиодные режимы работы МДП-структуры также проявляются в нескольких формах. Фотогенерация и разделение электронов и дырок встроенным полем приповерхностной обедненной области вызывает фотовольтаический эффект. Наведенная фото-ЭДС, как и в случае вентильного режима обычного фотодиода, служит входным сигналом для соответствующей электронной схемы; отличие состоит в том, что, поскольку внутреннее сопротивление МДП-фотодиода бесконечно, входное сопротивление схемы также должно быть велико. Обычно используются схемы, изготавливаемые по МДП-технологии.

Тот же эффект фотогенерации носителей проявляется и иным образом, а именно в изменении ширины области пространственного заряда и соответственно емкости структуры, т. е. в форме фотоемкостного эффекта. Работающая в режиме фотоварикапа МДП-структура интересна тем, что характеризуется минимальным уровнем собственных шумов, как и любая почти идеальная реактивность, так как в этом режиме отсутствуют основные источники шума - темновой ток и фототок. Специальные МДП-структуры с очень тонким «туннельно-. прозрачным» диэлектриком пропускают электрический ток и по свойствам близки к фотодиодам с барьером Шотки. Если диэлектрик непрозрачен для носителей заряда, то фототок также может протекать через структуру, но лишь в течение времени, необходимого для установления нового равновесного состояния. Иными словами, при работе в импульсном, динамическом режиме МДП-фотодиод генерирует токовые сигналы точно так же, как любой другой фотодиод. Концепция динамического режима породила более своеобразный МДПДМ-фотодиод, в котором фото носители, генерируемые в полупроводнике (обычно близком к собственному), пространственно растягиваются под действием

приложенного напряжения, но из кристалла не выходят, а могут лишь рекомбинировать. Это позволяет получить значительное усиление фототока.

МДП-структура может работать в динамическом режиме лавинного умножения фототока: внешнее напряжение подается в виде прямоугольного импульса, а регистрируемый оптический сигнал поступает лишь во время действия импульса. Принципиальное отличие от обычных ЛФД заключается в самостабилизации развития лавины по всей площади фоточувствительной поверхности. Известно, что если повышать напряжение, все больше приближаясь к [/проб, то становятся существенными микроплазменные явления, приводящие к тому, что пробой локализуется в отдельных случайных точках плоскости р-/г-перехода. У МДП-структур при возникновении микроплазмы и увеличении тока напряжение перераспрселяется так, что большая его часть начинает падать на диэлектрике - это предохраняет от стягивания лавины в отдельные точки. Отсюда вытекают два важных следствия: возможность достижения очень высоких значений коэффициента умножения и создания больших и однородных фоточувствительных площадей (до нескольких квадратных сантиметров).

Параметры МДП-фотодиодов, а нередко и просто их работоспособность находятся в очень тесной зависимости от технологического совершенства структуры. Заметим, что несмотря на кажущуюся внешнюю простоту изготовить качественную МДП-структуру значительно сложнее, чем p-n переход, так как p-n переход создается внутри монокристалла, а в МДП-элементе взаимодействуют материалы структурно и химически разнородные. Как правило, диэлектрическая пленка получается несплошной, нестехиометрической, содержащей те или иные инородные включения. Особенно сильно влияют ионы щелочных металлов, обладающие заметной подвижностью и вызывающие нестабильность параметров.

Структурные и химические изменения в диэлектрике, миграция в него атомов металла верхнего электрода, диффузия газов через пленки металла и диэлектрика - все это порождает эффекты старения, деградации, гораздо более существенные, чем у обычных МДП-приборов, так как в фотоприборах должны использоваться наиболее тонкие металлические и диэлектрические пленки

Способы изменения спектральных характеристик

Основная проблема несовместимости по длинам волн, с которой мы уже встречались ранее конструировании и изготовлении монолитных структур лазер - волновод, также очень существенна в отношении волноводных детекторов. Идеальный волновод должен обладать минимальным поглощением на используемой длине волны. Однако эффективность детектирования зависит от межзонного поглощения, приводящего к генерации носителей. Следовательно, если детектор монолитно объединен с волноводом, необходимо предусмотреть некоторые меры, касающиеся увеличения поглощения фотонов, пропускаемых волноводом внутрь объема детектора. В этом отношении оказался эффективным целый ряд различных способов.

В случае монолитных волноводных детекторов на GaAs эффективную ширину запрещенной зоны можно локально уменьшить в активной части объема детектора путем добавки легирующих атомов кремния или германия. Кривые пропускания для GaAs, легированного Ge, представлены на рисунке 7.

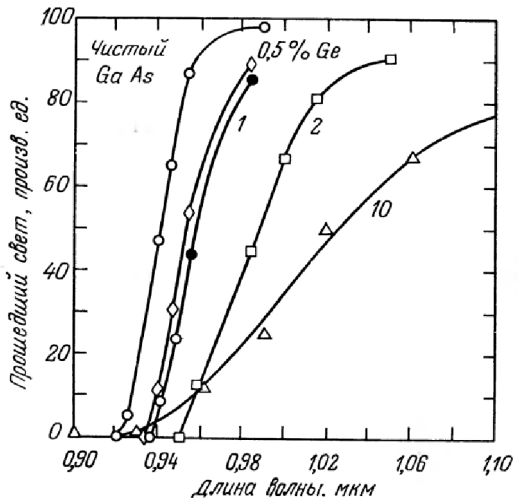


Рисунок 7 - Характеристики оптического пропускания GaAs, легированного германием.

Добавка лишь 0,5% Ge существенно сдвигает край поглощения, а введение 10% Ge передвигает его с длины волны 0,93 мкм для чистого GaAs приблизительно до 1,02 мкм. Следует заметить, что кривые, показанные на рисунке 7, дают процентное пропускание. Следовательно, поглощение на более длинных волнах возрастает с увеличением концентрации Ge. Поведение GaAs, легированного кремнием, по существу аналогично показанному на рисунке 8. Легирующие атомы кремния или германия можно добавить в основной волноводный материал GaAs (или GaAlAs) в том месте, где необходимо создать детектор путем диффузии через маску, ионной имплантации или эпитаксиального роста. Таким образом, диоды либо на p-n переходе, либо на основе барьера Шоттки формируются на поверхности легированной области, и для завершения структуры прибора добавляются контакты.

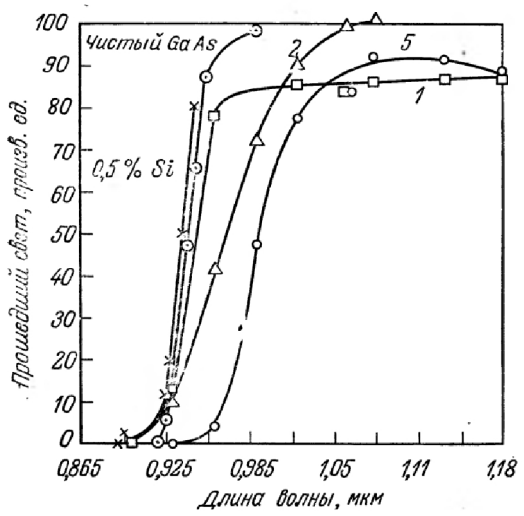


Рисунок 8 – Характеристики оптического пропускания GaAs, легированного кремнием.

Гибридные структуры

Один из наиболее очевидных подходов к получению совместимости по длинам волн заключается в использовании гибридных структур, в которых диодный детектор, сформированный в материале со сравнительно узкой запрещенной

зоной, связан с волноводом, изготовленным в материале с более широкой запрещенной зоной. Оба материала выбираются так, чтобы фотоны требуемой длины волны свободно пропускались волноводом, но сильно поглощались в материале детектора. Примером такого типа гибридного волноводного детектора является структура, состоящая из стеклянного волновода на кремнии, которая была продемонстрирована Островским с сотр. и представлена на рисунке 9.

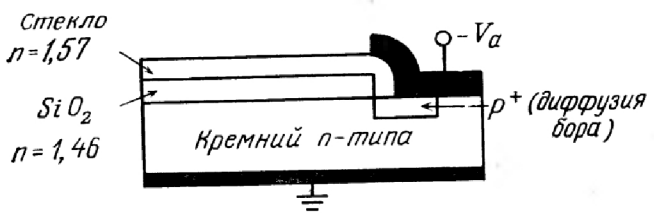


Рисунок 9 - Гибридный волноводный детектор, состоящий из стеклянного волновода, состыкованного с кремниевым фотодиодом.

Диод был сформирован путем диффузии бора на глубину около 1 мкм в кремниевую подложку n-типа сопротивлением 5 Ом*см. В качестве маски для диффузии использовался слой термически окисленного SiO_2 толщиной 1 мкм. Затем был напылен стеклянный волновод и изготовлены электроды из проводящей серебряной краски, как показано на рисунке. Измеренные полные потери в волноводе составляли 0,8 дБ/см ± 10% на длине волны 632,8 нм. Эффективность связи волновода с детектором составляла 80%. Однако из-за того, что свет входит в диод в направлении, перпендикулярном, а не параллельном плоскости перехода, данная геометрия волноводного детектора не обладает многими преимуществами. Тем не менее, можно ожидать хороших высокочастотных характеристик. Эти диффузионные диоды имели емкость лишь $3*10^{-9}$ Ф/см^2 при напряжении обратного смещения 10 В. Следовательно, диодный детектор радиусом около 10 мкм, используемый совместно с нагрузочным сопротивлением 50 Ом, будет иметь временную константу RC около 15 пс, что означает возможность регистрации частоты модуляции свыше 10 ГГц. Хотя

гибридные детекторы обеспечивают возможность выбора для детектора и волновода материалов с оптимальными характеристиками поглощения, большей эффективности связи можно достичь использованием методов монолитного изготовления. Монолитные структуры обладают преимуществом, что свет входит в плоскости перехода, а не перпендикулярно ей. Пример такой монолитной структуры приведен на рисунке 10, 11.

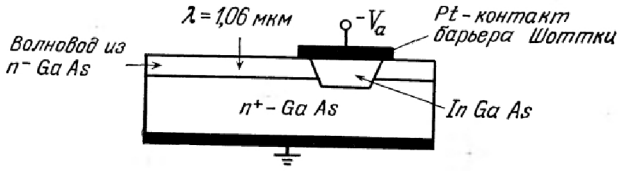


Рисунок 10 – Монолитно интегрированный InGaAs-детектор в GaAs волноводе

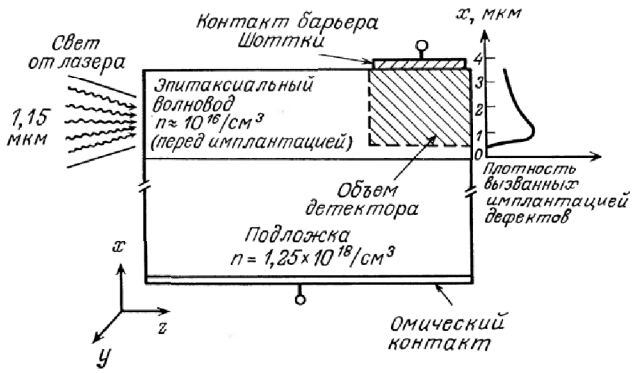


Рисунок 11 – Схема оптического детектора, полученного имплантацией протонов

Факторы, ограничивающие рабочие характеристики

В конструкции интегрально-оптических детекторов существует целый ряд механизмов, которые по-разному могут ограничивать рабочие характеристики. Не все из них важны в каждом случае применения. Однако конструктор (или пользователь) должен понимать причину ограничений, связанных с различными типами и геометриями устройств.

Высокочастотная отсечка

Благодаря малой площади волноводного фотодетектора, представленного на рисунке 4, б, постоянную времени RC, которая очень часто ограничивает характеристики обычных диодов, можно сделать достаточно малой, чтобы допустить работу на частотах, значительно превышающих 10 ГГц. В этом случае необходимо рассмотреть другие потенциально ограничивающие эффекты:

- постоянная времени RC, обусловленная последовательным объемным сопротивлением и емкостью перехода;
- время диффузии носителей из областей, расположенных вне обедненного слоя;
- время жизни носителей и длина диффузии;
- емкость и индуктивность корпуса;
- время дрейфа носителей через обедненный слой;
- захват носителей в глубокоуровневых ловушках.

Время дрейфа носителей сквозь обедненный слой можно минимизировать такой конструкцией прибора, чтобы поле в обедненном слое было достаточно большим для передвижения носителей со скоростью, ограниченной рассеянием. Например, в материале GaAs предельная скорость 10^7 см/с, ограниченная рассеянием, достигается при напряженности электрического поля $2*10^4$ В/см. Следовательно, время прохода через обычный обедненный слой шириной 3 мкм можно уменьшить вплоть да $3*10^{-11}$ с. Однако важно сконструировать детектор, чтобы обедненный слой простирался через весь волновод до подложки так, чтобы все носители создавались бы в самом обедненном слое. Если обеднение неполное, то носители, образованные в не обедненном материале, должны диффундировать в обедненную область относительно медленно, прежде чем попадут на электрод. Этот процесс происходит, грубо говоря, в течение минимального времени жизни неосновных носителей (около 10^{-8} с в слаболегированном материале GaAs).

Следовательно, при детектировании короткий оптический импульс будет появляться с длинным диффузионным хвостом.
Захват носителей на глубоких уровнях может также вызвать появление хвоста в форме детектируемых коротких импульсов, поскольку время освобождения может быть относительно длинным. Глубокоуровневые ловушки обычно связаны с дефектами, присутствующими в кристаллической решетке полупроводника. Следовательно, при выборе материала и изготовлении устройства следует особо позаботиться о минимизации числа дефектов.

Линейность

Фотодиод с обедненным слоем, смещенный в обратном направлении напряжением более 2 В, работает в режиме фотодиода (или фотопроводимости). Таким образом, он работает как источник тока, причем его ток пропорционален входной оптической мощности, обычно вплоть до уровней мощности 1 мВт. Следовательно, в большинстве случаев приложения этот прибор достаточно линеен. На более высоких уровнях мощности происходит насыщение, когда концентрация фотогенерированных носителей настолько велика, что поле в обедненном слое уменьшается из-за эффектов пространственного заряда. Это уменьшение поля особенно важно в высокочастотных приложениях, поскольку оно может вызвать перенос носителей со скоростью, меньшей скорости, ограниченной рассеянием.

Шум

Эффекты шума в волноводных детекторах по сути те же самые, что и в обычных фотодиодах. Основными компонентами шума являются тепловой шум, возникающий в объемных сопротивлениях устройства, дробовой шум, связанный с неоднородностями тока, например генерацией носителей и их рекомбинацией, а

также фоновый шум, вызванный фотонами, не являющимися частью оптического сигнала, входящего в детектор.

На основе относительно простой модели, разработанной Дидоменико и Звелто, можно показать, что отношение сигнал-шум в фотодиодах с обедненным слоем, связанное с эффектами теплового и дробового шума, задается выражением:

$$(S/N)_{мощн} = (\eta_q / 4B) M^2 \varphi_0 A \left(1 + (2kT/q)\left((\omega RC)^2 R I_S\right)\right)^{-1} \quad (14)$$

где η_q - квантовая эффективность, B - ширина полосы, M - коэффициент модуляции (в случае луча света, модулированного по интенсивности), φ_0 — плотность потока падающих фотонов, A — площадь входного торца, R - объемное сопротивление диода, ω - частота модуляции оптического сигнала, I_S - ток обратного (темнового) насыщения, C - емкость. В случае лавинного фотодиода существуют дополнительные источники шума, связанные со статистической природой лавинного процесса, которые не учитываются выражением (14). Волноводному детектору присущи преимущества перед обычными детекторами в отношении фонового шума, поскольку волновод работает как фильтр, устраняющий большую часть фонового света. Хорошее согласование с длиной волны сигнала также уменьшает фоновый шум. Например, если свет длиной волны 850 нм от излучателя на GaAlAs принимается Si-детектором, то будут детектироваться также фоновые фотоны с длинами волн короче 1,2 мкм. Однако, если вместо Si-детектора используется GaAs-детектор, эти фотоны с длиной волны в диапазоне 1,2— 0,9 мкм не будут вносить вклада в фоновый шум.

p-i-n фотодиоды

P-i-n фотодиод (рисунок 12) представляет собой трехслойную структуру из p- и n-полупроводников, разделенную сравнительно протяженной i – областью слаболегированного полупроводника. Поглощение квантов света происходит в

обедненной i - области, при этом в результате внутреннего фотоэффекта в объеме этою слоя образуются электроны и дырки, время жизни которых намного превышает их время жизни в p- и n- слоях (где они быстро рекомбинируют). Во внешнем электрическом поле носители заряда дрейфуют, обусловливая электрический ток в замкнутой электрической цепи фотоприемного устройства (рисунок 13).

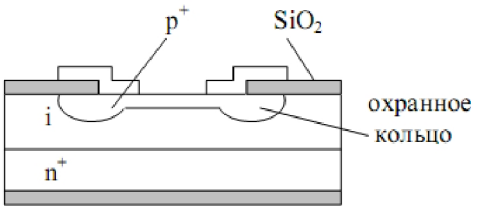


Рисунок 12 - Структура p-i-n фотодиода

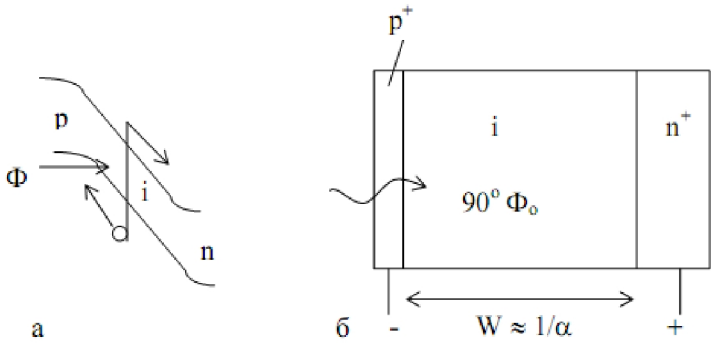


Рисунок 13 – Энергетическая диаграмма p-i-n фотодиода при освещении под обратным смещением (а) и относительные размеры p-i-n областей структуры (б)

Спектральная чувствительность p-i-n фотодиодов определяется типом полупроводниковой структуры, при этом красная граница фотоэффекта составляет около 0.9 мкм для арсенида галлия, 1.1 мкм для кремния и 1.7 мкм для германия; сегодня разработаны и практически используются более сложные фотоприемные полупроводниковые структуры, в том числе квантоворазмерные, обладающие высокой квантовой эффективностью как в узкой, так и в широкой полосах спектра длин волн. В лучших образцах p-i-n фотодиодов чувствительность фотодиода (отношение величины фототока к мощности падающего излучения) достигает уровня свыше 0.9 А/Вт. Временные характеристики p-i-n фотодиодов определяются толщиной

обедненного i-слоя, скоростью дрейфа носителей заряда и собственной емкостью фотодиода и подводимых электродов. Высокоскоростные p-i-n фотодиоды характеризуются временем нарастания-спада регистрируемого оптического сигнала величиной 10 - 100 пикосекунд. Величина обратного напряжения смещения в p-i-n фотодиодах обычно от 10 до 20 В.

К достоинствам p-i-n фотодиода относятся:

- высокая чувствительность (для λ=0,9 мкм S=0,7 А/Вт) и быстродействие (10^{-9} – 10^{-10} с);
- высокая чувствительность в длинноволновой части спектра, обусловленная широкой i-областью;
- малая барьерная емкость;
- высокая эффективность при малых обратных напряжениях.

Недостатками p-i-n фотодиодов являются:

- малая фотоЭДС (<0,35-0,45 В);
- повышенные токи утечки;
- невысокая воспроизводимость;
- несовместимость с интегральными схемами, вследствие чего приходится использовать волоконный ввод излучения.

Лекция 5

Волноводные усилители. Выходная мощность сигнала, эффективность накачки. Ширина и равномерность полосы усиления. Полупроводниковые усилители. Эрбиевые усилители.

Наиболее распространены два типа оптических усилителей – полупроводниковый квантовый усилитель и волоконно-оптический (например, эрбиевый).

В полупроводниковом квантовом усилителе усиливающей свет средой является тонкий «активный» волноводный слой на границе полупроводников p- и n-типов, через который проходит электрический ток накачки. Конструкция усилителя подобна конструкции обычного полупроводникового лазера, коэффициент отражения зеркал резонатора которого близок к нулю. Усиливаемое излучение обычно вводится в активный волновод волновод при его прямом согласовании (стыковке) с подводимым оптическим волокном; похожим образом усиленное излучение выводится из активного волновода.

Оптический усилитель (ОУ) увеличивает мощность оптических сигналов в волоконно-оптической линии связи без их промежуточною преобразования в электронные сигналы. Такой способ усиления особенно удобен в линиях связи со спектральным уплотнением, благодаря возможности одновременною усиления нескольких информационных сигналов с различными длинами волн. В волоконно-оптическом усилителе также используется лазерный принцип усиления, при этом свет проходит через оптическую инверсную среду, в которой населенность верхних электронных состояний выше, чем нижних, что обеспечивается благодаря оптической накачке.

Эрбиевые волоконно-оптические усилители

Коэффициент усиления G оптического усилителя определяется как отношение мощности оптического сигнала на выходе и на входе усилителя:

$$G = P_{вых} / P_{вч}.$$

Часто коэффициент усиления исчисляется в дБ, при этом G[дБ] =10*lg(G).

При относительно небольшой мощности входного сигнала полный коэффициент усиления G равен:

$$G = \exp(Fgl),$$

где g - коэффициент усиления активной среды, $см^{-1}$, l - длина усиливающей среды, см, F – оптический коэффициент ограничения, характеризующий относительную долю мощности вводимою излучения, распространяющегося в усиливающей среде.

Средняя длина волны и ширина полосы частот усиления зависит от типа активной среды полупроводника и ширины запрещенной зоны. В ближнем инфракрасном диапазоне в окрестности длин волн 1,3 и 1,55 мкм данная ширина составляет 30 – 40 нм. При этом в усилителях используются те же методы создания активных сред, что и в полупроводниковых лазерах. Типичный коэффициент усиления современных устройств составляет 30 дБ.

В волоконно-оптических усилителях спектральные характеристики полностью определяются структурой энергетических уровней рабочего материала, например, эрбия, неодима, германия и др. В однородно уширенной среде волоконно-оптического усилителя с оптической накачкой из-за явления насыщения усиления величина коэффициента g изменяется вдоль оси оптического волокна. При этом может иметь место нелинейность полного усиления G от мощности входного и выходного излучения:

$$G = \exp(Fg_0 l)\exp\left(\left(\frac{1}{G} - 1\right)\frac{P_{вых}}{P_н}\right),$$

где g_0 – ненасыщенный коэффициент усиления среды; $P_н$ – мощность насыщения, при достижении которой коэффициент усиления уменьшается в 2 раза (рисунок 1).

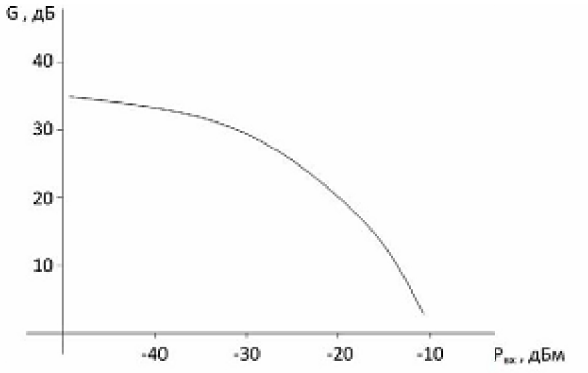


Рисунок 1 – Типичная зависимость коэффициента усиления от мощности входного сигнала

Принципиальная схема волоконно-оптического усилителя на основе кварцевого оптоволокна, легированного ионами эрбия, показана на рисунке 2.

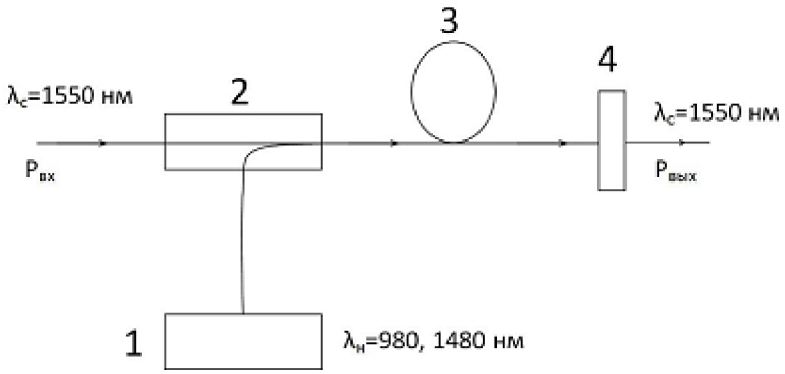


Рисунок 2 – Схема волоконно-оптического усилителя: 1 – источник накачки (мощный п/п лазер с длиной волны 980 или 1480 нм), 2 – мультиплексор, 3 – эрбиевое волокно, 4 – полосовой оптический фильтр на длину волны 1550 нм.

Структура энергетических уровней трехвалентного иона эрбия в кварцевом стекле и схема поглощения излучения накачки, безызлучательного и рабочего переходов в диапазоне длин волн 1520-1570 нм показаны на рисунке 3.

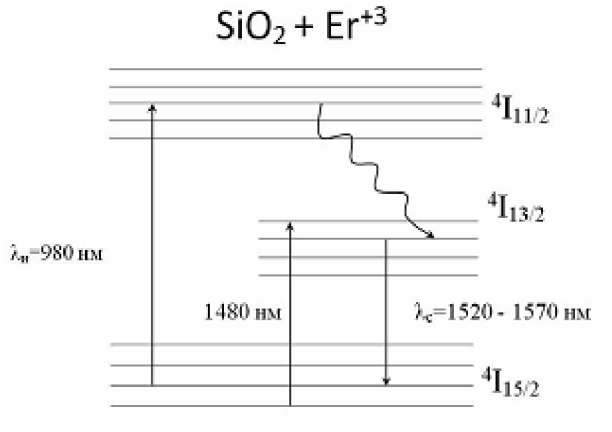


Рисунок 3 – Энергетические уровни иона эрбия

Шумовые характеристики оптических усилителей описываются с помощью NF-показателя (noise figure), равного отношению сигнал/шум на входе и выходе усилителя:

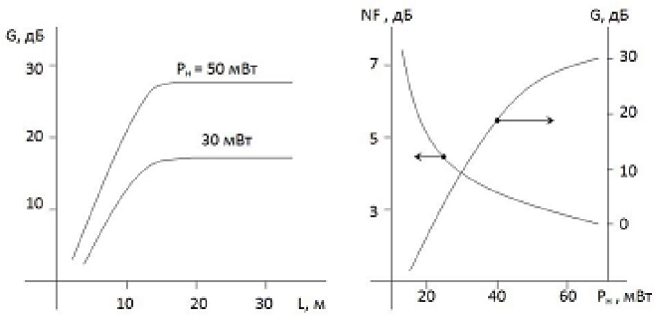
$$NF = \frac{(S/N)_{вх}}{(S/N)_{вых}},$$

где S – мощность сигнала, мВт, N – мощность шума, мВт. Также применяется логарифмический показатель NF[дБ]=10*lg(NF). Основной причиной собственных шумов усилителя являются спонтанные переходы между рабочими уровнями, для которых с помощью оптической накачки создана инверсия населенностей. При большой величине усиления и относительно узкой полосе частот выполняется равенство:

$$NF = \frac{1}{\eta G} + \frac{2(G-1)\eta_u}{\eta G} \approx \frac{2\eta_u}{\eta} \to 2,$$

где $\eta_и$ – параметр инверсии населенности, $\eta_и = N_2/(N_2\text{-}N_1)$, η – коэффициент потерь при вводе излучения в усилитель. Таким образом, минимальное значение фактора шума NF = 2.

Необходимо отметить, что длины волн усиления в эрбиевых волокнах хорошо совпадает с областью минимального поглощения кварцевого оптического волокна. Кроме этого, эрбиевые усилители обладают высокими параметрами при сравнительно небольших мощностях излучения накачки и длине активного участка оптоволкна. Типичные рабочие характеристики волоконно-оптических эрбиевых усилителе приведены на рисунке 4.



а. *б.*

Рисунок 4 – Зависимость коэффициента усиления G от длины волокна (а) и зависимость NF-показателя и коэффициента усиления G от мощности накачки.

Полезным свойством волоконно-оптических усилителей является возможность двунаправленной оптической накачки, т.е. возможность ввода излучения накачки с обоих концов оптического волокна. В этом случае спектральные мультиплексоры, с помощью которых в усиливающее волокно вводится усиление накачки, располагаются с обеих сторон волокна. В лучших образцах волоконно-оптических усилителей на длине волны достигнуто усиление 45 дБ при NF-показателе 3-4 дБ.

Оптическое усиление в диодах с гетеропереходами

Полупроводниковые оптические усилители имеют хорошие перспективы практических применений во многих указанных областях в качестве усилителей мощности выходного излучения; линейных усилителей для компенсации потерь в линиях связи при передаче информации по волокну; оптических предусилителей для повышения чувствительности приемников, кольцевых низкокогерентных источников для гироскопов, эффективных оптических модуляторов в диапазоне 1–10 ГГц; для съема информативного сигнала в его электрической цепи по электрическому отклику на оптический сигнал, проходящий через активную среду, в качестве компенсаторов дисперсии, оптических коммутаторов.

Ранее было показано, что в том случае, в полупроводниках под действием оптической или электронной накачки становится возможным достижение инверсии заселенности между электронами в зоне проводимости и дырками в валентной зоне. Тогда даже сумели получить условие прозрачности, т. е. концентрацию заряженных носителей $n_{пр}$ в зонах, необходимую для того, чтобы материал был эффективно прозрачен для излучения, и при превышении которой материал проявляет усиление. Относительно тока электрической инжекции это означает, что материал становится прозрачным, когда он возбуждается плотностью тока прозрачности $J_{пр}$, определяемой:

$$\widetilde{F}_{1/2}\left(\frac{J_{nt}t}{qdN_C}\right)+\widetilde{F}_{1/2}\left(\frac{J_{nt}t}{qdN_V}\right)=0,$$

где N_C и N_V - эффективные плотности состояний в зоне проводимости и валентной зоне. При возбуждении выше этого порога прозрачности усиление будут испытывать фотоны с энергией hν, удовлетворяющей критерию:

$$E_g < h\nu < E_{Fc} - E_{Fv}.$$

Условие усиления дается формулой:

$$\gamma(h\nu) = \alpha_0(\omega(f_c(h\nu) - f_v(h\nu))),$$

где f_c, f_v – функции Ферми-Дирака, α_0 - поглощение пустой зоны проводимости (в условиях нулевого тока), определяемое выражением:

$$\alpha_0(\omega) = \frac{2\pi q^2 x_{vc}^2}{\lambda_0^2 \varepsilon_0 h n_{пр}} \left(\frac{4\pi m_{эфф}}{h} \right)^{3/2} \sqrt{\omega - 2\pi E_g / h},$$

где x_{vc} – матричный элемент, определенный ранее, $m_{эфф}$ - приведенная эффективная масса, $n_{пр}$ - показатель преломления полупроводника, а λ_0 - длина электромагнитной волны в вакууме.

Таким образом, по мере того, как плотность тока J, протекающего через диод возрастает, увеличивается и концентрация носителей. Это приводит к тому, что квазиуровни Ферми расщепляются, что приводит уширению кривой усиления. Рисунок 5 показывает

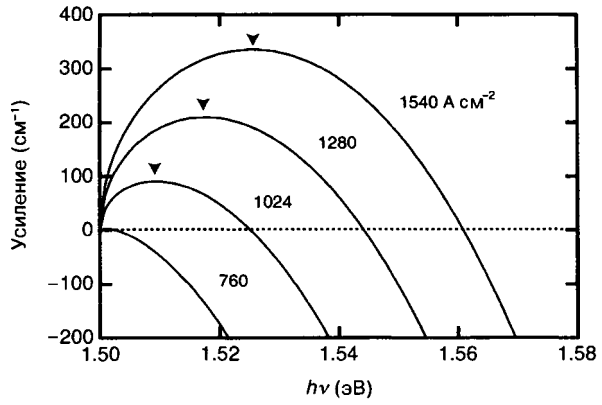


Рисунок 5 - Зависимости усиления от плотности тока инжекции для гетероструктуры

GaAs/AlGaAs толщиной 100 нм в предположении, что излучательное время жизни составляет 5 нс при внутренней квантовой эффективности 0,5, что соответствует $t_{полн}$ = 2,5 нс. В данном случае плотность тока прозрачности составляет 760 А $см^{-2}$.

результирующие кривые усиления для разных уровней тока инжекции. Максимальное усиление γ_{max} возрастает практически линейно с увеличением концентрации носителей n в области выше порога прозрачности. Общепринятым является использование для сложной зависимости γ_{max} от n линейной аппроксимации:

$$\gamma = \gamma_0 \left(\frac{n}{n_{прозр}} - 1 \right).$$

Здесь параметр γ_0 выбирают методом оптимального подбора. При концентрации n в диапазоне от 0 до $n_{прозр}$ среда является поглощающей. При $n > n_{прозр}$ материал становится прозрачным и зависимость $n_{прозр}$ от n обеспечивает достаточную точность расчета. Можно переписать данное выражение в виде зависимости от плотности тока J. В этом случае максимальное усиление γ_{max} дается выражением:

$$\gamma = \gamma_0 \left(\frac{J}{J_{прозр}} - 1 \right),$$

Пример. Рассмотрим усиление, обеспечиваемое лазерным диодом на основе двойного гетероперехода GaAs/AlGaAs. Из экспериментов установлено, что время жизни излучательной рекомбинации составляет 5 нс. Следовательно, пороговая концентрация носителей, соответствующая прозрачности $n_{прозр}$, составляет $1{,}2 * 10^{18}$ $см^{-3}$. Предположим, что диод обладает квантовой эффективностью 50%, толщина его активной области составляет 100 нм, ширина и длина соответственно равны 10 и 300 мкм. В этом случае пороговая плотность тока прозрачности

$J_{прозр}$=1,6*10^{-19} Кл х 1,2*10^{18} $см^{-3}$/ (0,5*5*10^{-9} с) = 760 А/$см^2$ или 23 мА. При концентрации носителей n= 2,4*10^{18} $см^{-3}$ (что соответствует уровню, вдвое большему порога, и току инжекции 46 мА) максимальное усиление составляет 300 $см^{-1}$. Это приводит к усилению G=8100 в указанной структуре.

Полупроводниковые оптические усилители используются в качестве усилительных каскадов на выходе лазерных диодов (в этом случае их обычно называют монолитными оптическими усилителями). В рассматриваемом случае одна из проблем заключается в том, чтобы они находились в режиме лазерной генерации. С этой целью для устранения обратной оптической связи на выходные окна наносятся просветляющие покрытия. Более того, для максимального извлечения энергии из этих приборов их резонаторам придается клиновидная форма. Это делает резонатор нестабильным, что препятствует формированию стационарных оптических мод в спектральном диапазоне усиления.

В настоящее время разработаны полупроводниковые усилители с центральными длинами волн в диапазонах 770-790 нм, 810-840 нм, 910-990 нм, 1030-1070 нм, 1270-1330 нм, 1510-1570 нм, с шириной полосы усиления 30-40 нм, обладающие коэффициентом усиления 20..30 дБ при входной чувствительности около 10 дБм.

Лекция 6

Принцип работы волноводов. Методы и основные этапы изготовления. Тонкопленочные, диффузионные, имплантированные, эпитаксиальные, электрооптические волноводы. Измерение толщины, затухания, эффективных показателей преломления волноводных мод.

Основы теории и виды волноводов

Трехслойный плоский волновод

Рассмотрим основную структуру трехслойного плоского волновода, представленную на рисунке 1. Будем считать, что ограничивающие свет слои с показателями преломления n_1 и n_3 распространяются до бесконечности в направлениях +x и -x соответственно. Главное достоинство подобного предположения состоит в том, что в направлении x нет необходимости учитывать отражение, за исключением отражения, существующего на границах n_1 - n_2 и n_2 - n_3.

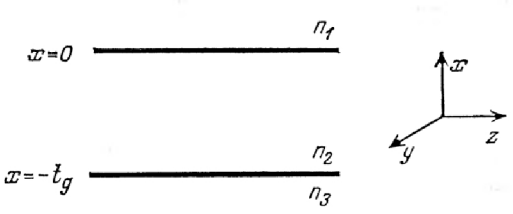


Рисунок 1 - Структура трехслойного плоского волновода.

Для плоской ТЕ-волны, распространяющейся в направлении & с постоянной распространения p, соответствующее уравнение Максвелла сводится к уравнению:

$$\nabla^2 E_y = \frac{n_i^2}{c^2}\frac{\partial^2 E_y}{\partial t^2}, \quad i = 1,2,3,\ldots, \qquad (1)$$

имеющему решение в виде

$$E_y(x,z,t) = \xi_y(x)e^{i(\omega t - \beta z)}, \qquad (2)$$

Индекс i в уравнении (1) естественно соответствует одному из слоев волноводной структуры. Напомним, что у ТЕ-волны компоненты E_x и E_z равны 0. Отметим также, что в выражении (2) функция $\xi_y(x)$ не зависит от y и z в силу предположения неограниченности слоев в этих направлениях, что исключает возможность отражения и образования стоячих волн.

Функция $\xi_y(x)$ имеет общий вид:

$$\xi_y(x) = \begin{cases} A\exp(-qx), 0 \le x \le \infty, \\ B\cos(hx) + C\sin(hx), -t_g \le x \le 0, \\ D\exp(p(x + t_g)), -\infty \le x \le -t_g. \end{cases} \qquad (3)$$

где A, B, C, D, q, h, p - постоянные, которые могут определяться условиями согласования на границах, требующими сохранения непрерывности функций $\xi_y(x)$ и $H_z = (i/\omega\mu)\partial\xi_y(x)/\partial x$. Так как по предположению магнитная проницаемость μ и частота ω - константы, второе условие сводится к требованию, чтобы производная $\partial\xi_y(x)/\partial x$ была непрерывной функцией. Таким образом, постоянные A, B, C, D можно определить, если сделать $\xi_y(x)$ и $\partial\xi_y(x)/\partial x$ непрерывными на границе между областями 1 и 2 (x=0), а $\xi_y(x)$ - непрерывной при x = - t_g. При этом имеем три уравнения с четырьмя неизвестными, и $\xi_y(x)$ можно выразить только через одну-единственную постоянную C':

$$\xi_y(x) = \begin{cases} C'\exp(-qx), 0 \le x \le \infty, \\ C'[\cos(hx) - (q/h)\sin(hx)], -t_g \le x \le 0, \\ C'[\cos(ht_g) + (q/h)\sin(ht_g)]\exp(p(x + t_g)), -\infty \le x \le -t_g. \end{cases} \quad (4)$$

Чтобы определить постоянные q, h и p, подставим выражение (4) в (2), используя результирующее выражение (1) для $E_y(x, z, t)$ в каждой из трех областей, и получим:

$$\begin{aligned} q &= \left(\beta^2 - n_1^2 k^2\right)^{1/2}, \\ h &= \left(n_2^2 k^2 - \beta^2\right)^{1/2}, \\ p &= \left(\beta^2 - n_3^2 k^2\right)^{1/2}, \\ k &= \frac{2\pi}{\lambda}. \end{aligned} \quad (5)$$

Отметим, что в выражении (5) величины q, h и p выражены лишь через единственную неизвестную величину β, которая является постоянной распространения волны в направлении z. Из условия непрерывности $\partial E_y / \partial x$ при x= -tg получим выражение для β. Взяв выражение для $\partial E_y / \partial x$ из (4) и используя ее непрерывность при x= -tg, приходим к условию:

$$-h\sin(-ht_g) - q\cos(-ht_g) = p[\cos(ht_g) + (q/h)\sin(ht_g)], \quad (6)$$

или после упрощения

$$tg(ht_g) = (p + q)/(h(1 - pq/h^2)). \quad (7)$$

Трансцендентное уравнение (7) совместно с выражением (5) можно решить либо графически, если начертить правую и левую стороны в виде кривой, зависящей от β и отметить точки пересечения, либо численно на ЭВМ. Независимо от метода решения результатом является набор разрешенных дискретных значений p,

соответствующих разрешенным модам. Для каждого значения β_m соответствующие ему значения q_m, h_m и p_m можно определить из выражения (5).

Одна оставшаяся неизвестная константа C выбирается произвольно. Однако удобно провести нормировку таким образом, чтобы функция $\xi_y(x)$ представляла собой поток мощности, выраженный в ваттах на единицу ширины в направлении y.

Симметричный волновод

Частный случай основной структуры трехслойного плоского волновода, представляющий особый интерес, встречается, когда $n_1=n_3$. Такие симметричные волноводы часто используются в оптических интегральных схемах, например когда направляющий слой с показателем преломления n_2 ограничивается на обеих поверхностях одинаковыми слоями с несколько меньшими показателями преломления n_1. В многослойных оптических интегральных схемах на основе GaAlAs часто используются волноводы такого типа. Выражения, выведенные в предыдущем разделе, применимы и к волноводам такого типа, однако можно сделать весьма серьезное упрощение при определении того, какие моды будут поддерживаться в таких волноводах. Во многих случаях не требуется знать значения β для различных мод. Возникает вопрос: может ли волновод направлять конкретную моду или нет?

В этом случае аналитическое выражение для условия отсечки для TE-мод можно вывести, имея в виду, что при отсечке (т. е. в точке, в которой электрическое поле становится осциллирующим в областях 1 и 3) величина β задается соотношением:

$$\beta = kn_1 = kn_3 . \qquad (8)$$

Подставляя выражение (8) в (5) находим величины p, q, h:

$$p = q = 0, \quad h = k\left(n_2^2 - n_1^2\right)^{1/2} = k\left(n_2^2 - n_3^2\right)^{1/2}. \qquad (9)$$

Подставляя выражение (9) в (7), приходим к условию:

$$ht_g = m_s\pi, \quad m_s = 0,1,2,\ldots. \qquad (10)$$

Объединяя выражения (9), (10) получим

$$k\left(n_2^2 - n_1^2\right)^{1/2} t_g = m_s\pi. \qquad (11)$$

Таким образом, чтобы волна в виде данной моды могла распространяться в волноводе, необходимо выполнение следующего условия:

$$\Delta n = (n_2 - n_1) > m_s^2\lambda_0^2/[4t_g^2(n_2 + n_1)], \quad m_s = 0,1,2,3,\ldots, \qquad (12)$$

в котором используется равенство $k=2\pi/\lambda_0$. Условие отсечки, заданное выражением (3.1.25), определяет, какие моды могут поддерживаться в волноводе при заданных значениях Δn и отношении λ_0/t_g. Интересно отметить, что мода низшего порядка ($m_s = 0$) в симметричных волноводах не обычна в том смысле, что для нее нет отсечки, как для всех других мод. В принципе этой моде может отвечать любая длина волны с бесконечно малым значением приращения Δn. Однако для малых значений Δn и (или) больших значений λ_0/t_g ограничение обычно незначительно с относительно большими «хвостами» моды, заходящими в подложку.

Если $n_2 \approx n_1$, то условие отсечки (12) принимает вид:

$$\Delta n = (n_2 - n_1) > m_s^2\lambda_0^2/8t_g^2 n_2, \quad m_s = 0,1,2,\ldots, \qquad (13)$$

а если $n_2 >> n_1$, задается выражением:

$$\Delta n = (n_2 - n_1) > m_s^2 \lambda_0^2 / 4 t_g^2 n_2, \quad m_s = 0,1,2,\ldots, \qquad (14)$$

Асимметричный волновод

Другим важным частным случаем структуры трехслойного плоского волновода является асимметричный волновод, в котором $n_3 >> n_1$. Конечно, чтобы имело место каналирование оптической энергии, необходимо, чтобы выполнялось условия $n_2 > n_3$. Волноводы такого типа часто используют в оптических интегральных схемах, в которых тонкая световедущая пленка осаждается или формируется каким-либо способом на материале подложки, обладающем несколько меньшим показателем преломления, а верхняя: поверхность световедущего слоя либо остается открытой на воздухе, либо для обеспечения электрического контакта покрывается слоем металла.

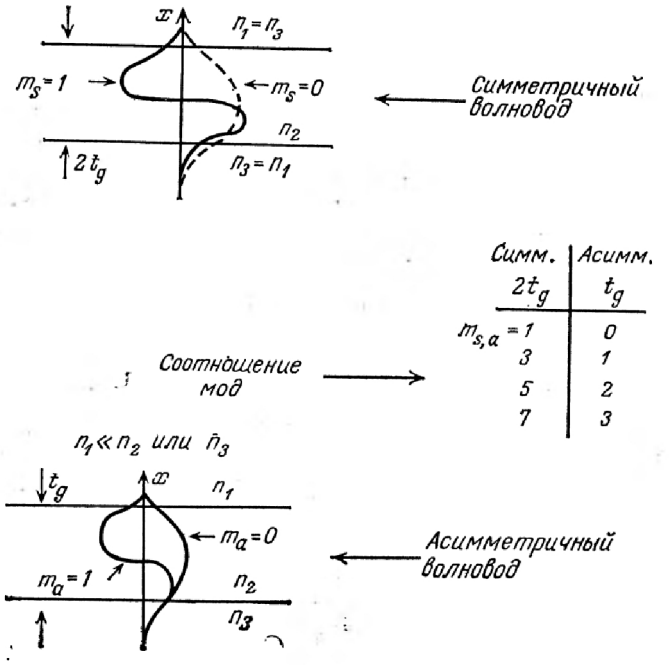


Рисунок 2 – Форма мод симметричного и ассиметричного плоских волноводов.

Рассмотрим асимметричный волновод, как показано на рисунке 2, который имеет толщину, равную половине толщины соответствующего ему симметричного волновода. На рисунке представлены две ТЕ-моды низшего порядка симметричного (m_s=0, 1) и асимметричного (m_a = 0, 1) волноводов. Отметим, что для хорошо ограниченных мод нижняя половина моды с индексом m_s=l симметричного волновода достаточно хорошо соответствует моде с индексом m_a=0 асимметричного волновода половинной ширины. Этим фактом можно воспользоваться в качестве математического аргумента для получения аналитического выражения, описывающего условия отсечки в случае асимметричного волновода.

Решая уравнение (7) для случая симметричного волновода с шириной $2t_g$ получим:

$$\Delta n = (n_2 - n_1) > m_s^2 \lambda_0^2 / 4(n_2 + n_3)(2t_g)^2, \quad m_s = 0,1,2,\ldots, \qquad (15)$$

Однако асимметричный волновод поддерживает только моды, соответствующие нечетным модам симметричного волновода шириной, в два раза большей ширины асимметричного волновода. Поэтому условие отсечки для симметричного волновода задается следующим неравенством:

$$\Delta n = (n_2 - n_3) > m_a^2 \lambda_0^2 / \left[16(n_2 + n_3)t_g{}^2\right], \quad m_s = 0,1,2,\ldots, \qquad (16)$$

где m_a - элементы подмножества, состоящего из нечетных значений m_s. Это удобно выразить следующим образом:

$$m_a = 2m + 1, m = 0,1,2,\ldots \qquad (17)$$

В предположении $n_2 \approx n_3$ неравенство (16) принимает вид:

$$\Delta n = (n_2 - n_3) > (2m+1)^2 \lambda_0^2 / 32 n_2 t_g^2, \quad m_s = 0,1,2,..., \qquad (18)$$

Хотя условия отсечки (18) и (14) справедливы лишь для частных случаев, их удобно использовать для оценки количества мод, которые могут поддерживаться в конкретном волноводе. В общем случае для ответа на данный вопрос или для определения значений β_m для разных мод необходимо решить трансцендентное уравнение (7).

Несмотря на то что выражение (18) выведено для случая ТЕ-волн, оно также верно и для ТМ-волн при условии $n_2 \approx n_3$. Следовательно, в асимметричном волноводе очевидно, что отсечка имеет место для всех мод в отличие от симметричного волновода, в котором моду TE_0 нельзя отсечь. Это свойство дает возможность, в частности, использовать такие волноводы в качестве оптических переключателей.

Прямоугольные волноводы

Плоские волноводы, рассмотренные в предыдущем разделе, находят достаточно широкое применение в различных прикладных областях интегральной оптики, несмотря на то. что с их помощью можно обеспечить ограничение оптических полей только в одном измерении. Даже достаточно сложные интегрально-оптические системы, например высокочастотный анализатор спектра Мергеряна и Маларки, можно изготовить е использованием плоских волноводов. Однако другие области применения требуют ограничения оптического поля в двух измерениях. При использовании волноводов полосковой геометрии прямоугольного сечения можно реализовать лазер с меньшей величиной порогового тока и одномодовой генерацией или электрооптический модулятор с меньшей потребляемой мощностью. В некоторых случаях двумерное ограничение оптического поля требуется для того, чтобы доставить свет от одной точки на поверхности ОИС к другой, соединив два элемента схемы.

Канальные волноводы

Общая схема прямоугольного волновода включает в себя волноводную область с показателем преломления n_1 окруженную со всех сторон средой с меньшим показателем преломления n_2, как это показано на рисунке 3.

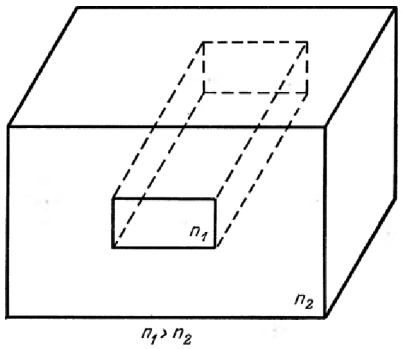


Рисунок 3 – Структура прямоугольного диэлектрического волновода.

Подобные волноводы часто называют канальными, полосковыми или трехмерными. При этом нет необходимости в том, чтобы показатель преломления среды, ограничивающей волноводную область, был бы одинаковым во всех точках этой среды. Для этой цели можно использовать несколько материалов с разными показателями преломления, меньшими n_1. Однако в таких случаях моды волновода не будут строго симметричными. Точное решение волнового уравнения для такого общего случая очень сложно и до сих пор еще не получено.

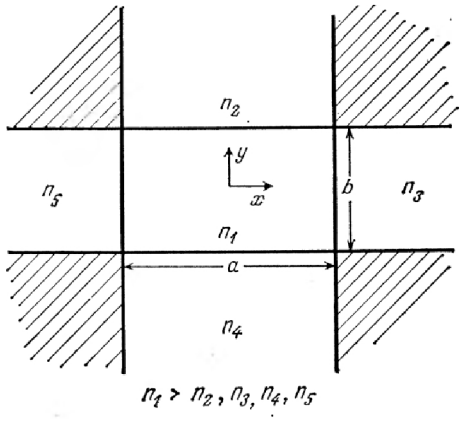


Рисунок 4 - Поперечное сечение прямоугольного диэлектрического волновода, ограниченного областями с меньшим значением показателя преломления.

Волноводы с наложенной полоской

Можно получить трехмерные волноводы с ограничением излучения по обоим направлениям x и y без фактического окружения волновода материалами с меньшим показателем преломления. Этого можно достичь за счет формирования полоски из диэлектрика с меньшим показателем преломления n_3 на верхней грани плоского волновода с показателем преломления n_1 (рисунок 5).

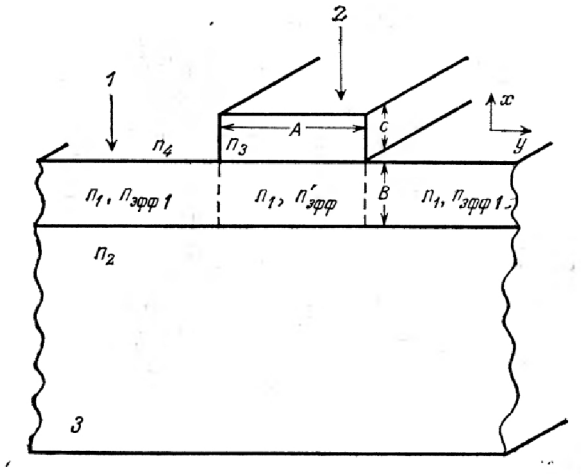


Рисунок 5 – Волновод с наложенной полоской из диэлектрика: 1 – плоский волновод, 2 – наложенная полоска, 3 – подложка.

$$n'_{эфф1} > n_{эфф1}, \quad n_1 > n_2 > n_3 > n_4 \quad или \quad n_1 > n_3 > n_2 > n_4$$

Подобные волноводы обычно называют либо волноводами с наложенной полоской, либо оптическими полосковыми линиями. Присутствие полосы, наложенной на волновод, характеризуется тем, что эффективный показатель преломления света $n'_{эфф1}$ в области под ней больше эффективного показателя $n_{эфф1}$ преломления в соседних областях y и x. Это приводит к ограничению в направлениях y и x. Физическую природу такого явления легче понять,

основываясь на приближениях лучевой оптики. Рассмотрим конкретную моду, распространяющуюся в направлении z, но состоящую из плоской волны, идущей по зигзагообразному пути в светопроводящем слое волновода. Так как $n_3 > n_4$, волна проникает несколько глубже на границе n_1-n_3, чем на границе n_1-n_4; поэтому эффективная высота волновода больше под полоской, чем в областях с любой стороны от нее. Это означает, что зигзагообразный путь плоской волны немного длиннее под наложенной полоской, в результате чего:

$$n'_{эфф1} = \beta' / k > n_{эфф1} = \beta / k. \qquad (19)$$

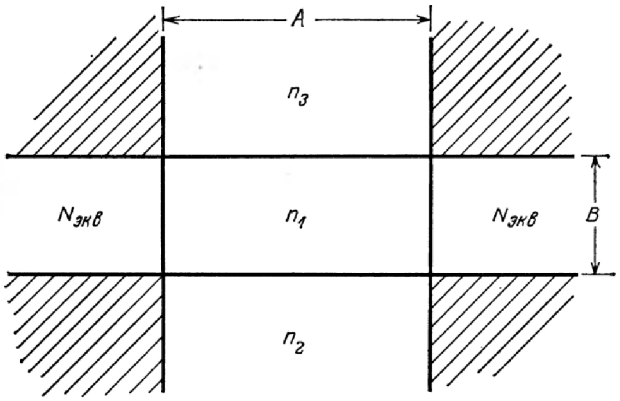


Рисунок 6 - Поперечное сечение прямоугольного диэлектрического волновода, эквивалентного волноводу с наложенной полоской из рисунка 5.

С использованием метода эффективного показателя преломления для анализа волновода с наложенной полоской, подобного изображенному на рисунке 5, можно показать, что его проводящие характеристики эквивалентны характеристикам диэлектрического волновода, подобного представленному на рисунке 6, у которого эквивалентный показатель преломления ограничивающих слоев задается выражением:

$$n_{экв} = \left(n_1^2 - n_{эфф1}'^2 + n_{эфф1}^2\right)^{1/2}. \qquad (20)$$

Можно изготовить волноводы с наложенной полоской, имеющей показатель преломления n_3 равный показателю преломления n_1 материала светопровода. Волноводы такого типа обычно называют гребенчатыми или ребристыми. На основе метода эффективного показателя преломления для волноводов такого типа было показано, что распространение волн характеризуется фазовой постоянной

$\beta = kN$, где

$$N^2 = n_{эфф1}^2 + b\left(n_{эфф1}^{'2} - n_{эфф1}^2\right). \quad (21)$$

В данном случае $n_{эфф1}^{'}$ - эффективный показатель преломления волновода, имеющего высоту, равную сумме толщин световодного слоя и наложенной полоски. Параметр b - нормализированный показатель преломления волновода и определяется выражением:

$$b = \left(N^2 - n_2^2\right)/\left(n_1^2 - n_2^2\right). \quad (22)$$

Для изготовления оптических полосковых линий можно также использовать металлические наложенные полоски. В таком случае две металлические наложенные полоски помещаются на поверхности волноводного слоя по обеим сторонам области, в которой требуется удержание световой энергии, как это показывается на рисунке 7.

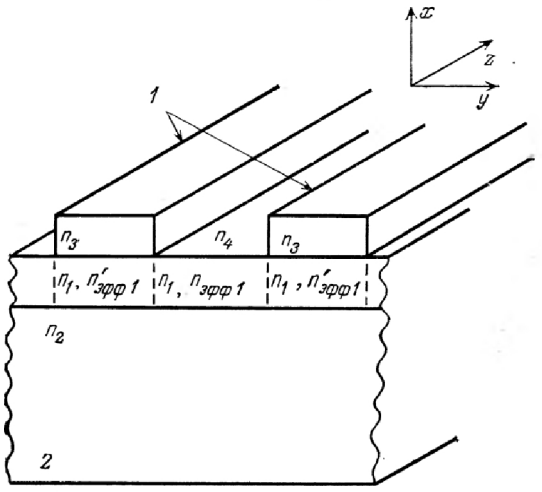


Рисунок 7 - Волновод с наложенными металлическими полосками. Эффективный показатель преломления и показатель преломления материала в объеме указаны в волноводном слое $n'_{эфф1} < n_{эфф1}$

Так как проникновение распространяющихся по волноводу волн глубже на поверхности раздела n_1-n_4, чем на поверхности раздела n_1-n_3, то требуемое ограничение в направлении у получается так же, как и в случае волноводов с наложенной полоской из диэлектрического материала. Волноводы с металлическими наложенными полосками особенно ценны в таких случаях применения, как электрооптические модуляторы, у которых на поверхности интегральной схемы необходимо иметь электроды; их роль могут играть металлические наложенные полоск

В принципе следует ожидать, что волноводы с наложенными полосками будут обладать меньшими оптическими потерями по сравнению с прямоугольными канальными волноводами из диэлектрика, так как рассеяние на шероховатостях стенок в этом случае меньше. Наибольшее беспокойство с точки зрения пригодности волноводов с наложенной полоской вызывает то, что небольшая разница в эффективном показателе преломления в направлении оси у, связанная с действием наложенной полоски, будет служить препятствием для уменьшения потерь излучения, возникающих на изгибах светопровода. Однако

при исследовании волноводов с наложенной полоской из стекла, изогнутых на 90° с радиусами изгиба до 2,5 мм, дополнительных потерь не обнаружено.

Диффузионные волноводы

Технология создания волноводов путем диффузии титана достаточно проста. Хотя существуют некоторые другие варианты такой технологии, на рисунке 8 приведены основные этапы литографического процесса, приводящие к формированию структуры полосковых волноводов необходимой конфигурации.

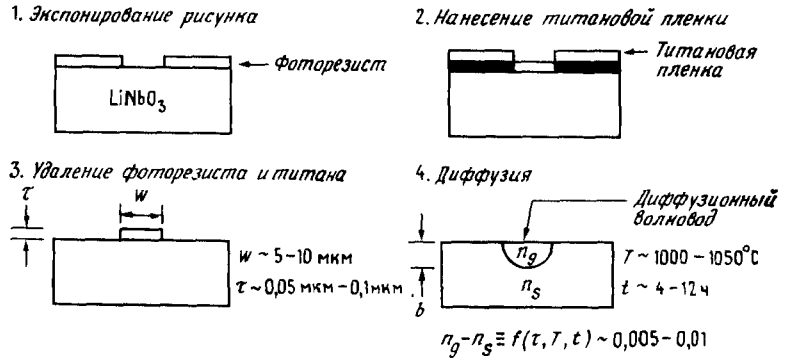


Рисунок 8 – Основные этапы создания диффузионных волноводов в ниобате лития

Первоначально на очищенную поверхность полированного кристалла наносится фоторезист, который засвечивается ультрафиолетовым излучением через маску, имеющую конфигурацию создаваемого волновода. После проявления экспонированного фоторезиста на его поверхности образуются окна, повторяющие конфигурацию маски. Далее на поверхность образца наносится пленка титана, для чего обычно используется высокочастотное распыление, распыление электронным пучком или термическое напыление. После этого образец опускается в растворитель, смывающий фоторезистивную маску с напыленным на нее титаном, в результате чего на поверхности кристалла остается требуемый рисунок, образованный титановой пленкой. Вместо этого процесса

«взрывной» литографии рисунок можно формировать нанесением титана на всю поверхность подложки и селективного травления требуемых участков

Далее образец помещается в печь для проведения диффузии при температурах 980-1050 С на время, составляющее в среднем 4-10 ч. При минимальных в указанном интервале значениях температуры требуется длительное время диффузии, а верхний предел температуры определяется ее значением в точке Кюри ниобата лития (~ 1125 С), превышать который нежелательно во избежание деполяризации кристалла. Для успешного проведения диффузии требуется соблюдение и ряда других условий. В первых опытах процесс иагрева и диффузии проводился в потоке аргона, пропускаемого через воду. При этом образцы охлаждались после диффузии в атмосфере кислорода с целью окисления кристаллов и компенсации потери кислорода при диффузии. Для уменьшения фоторефракции первоначально применялся водяной пар. Позднее было установлено, что этот процесс снижает диффузию Li_2O из образца, которая может привести к нежелательному образованию планарного волновода для необыкновенной поляризации. Обычно относительная влажность в потоке газа 80% является достаточной для избежания образования такого волноводного слоя при температуре диффузии ниже 1000 С. Однако этот метод успешно применялся и при температурах 1050 С.

Другим способом, позволяющим снизить нежелательную диффузию из кристалла, является помещение в потоке газа до образца нагреваемого источника Li_2O, насыщающего незамкнутый объем, такого, как $LiCO_3$ или $LiNbO_3$. Аналогичный результат может быть достигнут в замкнутом объеме печи без специального источника. Другой сходный метод заключается в запрессовке подложки в порошок $LiNbO_3$ идентичного состава. Насыщение объема газа оксидом лития достигается за счет спрессованного порошка, и Li_2O, первоначально содержащийся в подложке, сохраняется в ней в процессе диффузии. Другим подходом является компенсация изменения показателя преломления за счет диффузии в области, прилегающие к полосковому волноводу, вещества, снижающего коэффициент преломления, например, такого,

как оксид магния. Хотя ни один из этих методов не исключает диффузии из образца Li_2O полностью, благодаря им в большинстве случаев удается избежать образования нежелательного пленарного волновода

Помимо легкости изготовления волноводные структуры $Ti:LiNbO_3$ обладают тем преимуществом, что важные параметры волноводов - ширина полоски, эффективная глубина и максимальный перепад показателей преломления - могут быть выбраны независимо друг от друга в данном технологическом процессе. В упрощенном представлении глубина волновода d зависит от времени диффузии t, температуры процесса T и в меньшей степени от ширины титановой полоски. Ширина волновода задается выбранными параметрами маски при фотолитографическом процессе, однако может несколько увеличиваться за счет боковой диффузии. Наиболее важный параметр - максимальный перепад показателя преломления Δn - зависит при фиксированных T и t от толщины и плотности титановой пленки. Эти соображения могут быть отражены численно за счет использования простой математической модели диффузии. При условии, что титановая плеика полностью продиффундировала в кристалл, относительная концентрация титана как функция глубины y запишется в виде:

$$C(y) = \frac{2}{\sqrt{\pi}} \frac{\tau}{d} \exp\left(-\frac{y^2}{d^2}\right), \qquad (23)$$

где τ - толщина исходной титановой пленки, а d - эффективная глубина диффузии. Последняя величина зависит от температуры диффузии T, температуры активации T_0 и коэффициента объемной диффузии D_0 следующим образом:

$$d = 2(Dt)^{1/2}, \qquad (24)$$

где $D = D_0 \exp(-T_0 / T)$. Величины D_0 и T_0 определяются составом кристалла ниобата лития. Для наиболее часто используемой конгруэнтной комбинации $LiNbO_3$ указанные значения составляют $T_0 \sim 2{,}5*10^4$ К и $D_0 = 2{,}5*10^{-4}$ см^2/с.

Механизм изменения показателя преломления за счет диффузии титана к настоящему времени до конца не выяснен, однако известно, что фотоупругий эффект, вызванный напряжениями за счет проникновения Ti, играет важную роль в этом процессе. Увеличение показателя преломления, конечно, зависит от концентрации Ti, ио эта зависимость несколько отличается для обыкновенного n_0 и необыкновенного n_e показателей преломления. Экспериментальные зависимости приведены на рисунке 9.

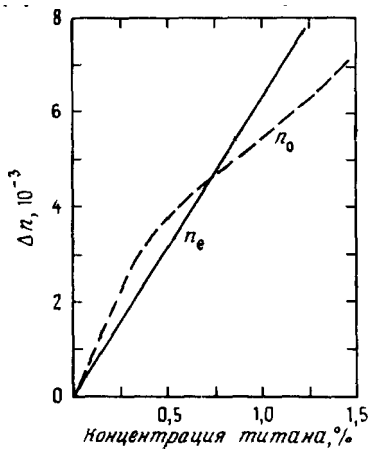


Рисунок 9 – Изменение обыкновенного и необыкновенного показателей преломления в зависимости от концентрации титана на λ=0,63 мкм

Значение n_e зависит от концентрации строго линейно, в то время как n_0 изменяется линейно (с иным коэффициентом) только при малых значениях концентрации титана и входит в насыщение при больших ее значениях.

Волноводы на основе электрооптического эффекта

GaAs и $Ga_{(1-x)}Al_xAs$ обладают достаточно сильно выраженным электрооптическим эффектом, выражающимся в том, что присутствие электрического поля приводит к изменению у них показателя преломления. Из этого следует, что если ни подложку из GaAs нанести металл в виде полоски, подобно тому как это делают

для получения контакта для барьера Шоттки, и подать обратное напряжение смещения, то электрическое поле в обедненном слое приведет к значительному изменению показателя преломления, которое обусловит формирование волноводного слоя с показателем преломления n_2, большим показателя преломления n_1 подложки. К счастью, очень просто получить барьер Шоттки на полупроводниковых соединениях GaAs и $Ga_{1-x}Al_xAs$. Почти любой металл, за исключением серебра, если его просто без термообработки нанести на материал n-типа, образует на этих материалах барьер Шоттки, а не омический контакт. При подаче на барьер Шоттки обратного напряжения смещения образуется обедненный носителями слой, как при p-n переходе. Показатель преломления света в этом слое больше показателя преломления материала подложки, что обусловливается двумя механизмами. Во-первых, он увеличивается за счет обеднения слоя носителями заряда. Во-вторых, увеличение показателя преломления вызывается присутствием электрического поля.

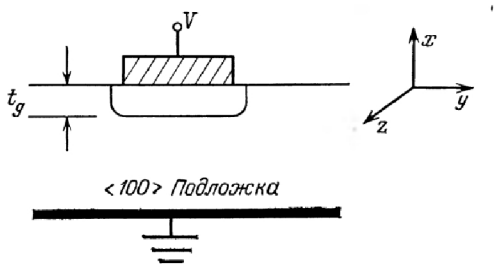


Рисунок 10 - Схема волновода на основе электрооптического эффекта.

Изменение показателя преломления для конкретной ориентации кристаллической решетки, представленной на рисунке 10, и для волны TE дается следующим выражением:

$$\Delta n = n^3 r_{41}\left(V / 2t_g\right), \quad (25)$$

где n - показатель преломления материала в отсутствие электрического поля, V - величина приложенного напряжения, t_g - толщина обедненного слоя, а r_{41} -

компонента электрооптического тензора, соответствующая выбранной ориентации кристалла и электрического поля. Электрооптический эффект является анизотропным, и, следовательно, другие типы ориентации не обязательно приводят к такому же изменению показателя преломления. Например, для случая такого кристалла, ориентированного, как на рисунке 10, но для волны ТМ (с вектором Е, поляризованным в направлении x), изменение показателя преломления равно нулю. Поэтому при проектировании волноводов на электрооптическом эффекте следует учитывать как ориентацию кристаллической решетки подложки, так и поляризацию направляемой волны.

Толщина волновода, т. е. толщина обедненного слоя, зависит от концентрации носителей материала подложки, а также от величины приложенного электрического поля. Если предположить разумной концентрацию носителей материала подложки порядка 10^{16} $см^{-3}$, а напряжение смещения V равным 100 В, т. е. максимально возможным, чтобы избежать лавинного пробоя полупроводника, то получим толщину обедненного слоя $t_g = 3,6$ мкм. Такой результат основан на использовании для расчета резкого перехода. В этом случае изменение показателя преломления составляет $\Delta n = 8,3*10^{-4}$. Можно видеть, что в таком случае в волноводе едва будет удерживаться лишь мода наинизшего порядка (для этого требуется $\Delta n = 5,7*10^{-4}$).

Из рассмотренного выше примера очевидно, что весьма серьезным ограничивающим фактором, который следует учитывать при разработке таких волноводов, является предотвращение лавинного пробоя.

Самым значительным преимуществом электрооптического волновода по сравнению с волноводами описанных ранее типов является то, что его можно ввести в действие электрическим способом и поэтому использовать в качестве переключающих и модулирующих элементов. Изменяя приложенное напряжение смещения, можно изменить не только показатель преломления волноводной области, но и толщину волновода, создавая таким образом для определенной моды желаемые условия либо до отсечки, либо после отсечки. Металлическая пластинка для контакта Шоттки может иметь, форму прямой или изогнутой узкой

полоски. В таком случае, под полоской при напряжении смещения выше уровня отсечки образуется прямоугольный канальный волновод.

Оптические потери, обусловленные рассеянием

Имеются два вида потерь, обусловленных рассеянием в оптических волноводах: объемное рассеяние и поверхностное рассеяние. Объемное рассеяние вызывается несовершенствами внутри объема волновода, например незаполненными местами, атомами примесей дефектами кристаллической структуры. Потери на единицу длины, обусловленные объемным рассеянием, пропорциональны числу несовершенств (центров рассеяния) на единице длины. Кроме того, объемные потери очень сильно зависят от относительных размеров этих несовершенств по сравнению с длиной волны излучения в материале. Во всех, кроме самых грубых, волноводах объемные дефекты настолько малы по сравнению с длиной волны и столь немногочисленны, что потери за счет объемного рассеяния пренебрежимо малы по сравнению с потерями за счет поверхностного рассеяния.

Потери, обусловленные поверхностным рассеянием

Потери за счет поверхностного рассеяния могут быть существенными даже для относительно ровных поверхностей, особенно в случае мод более высокого порядка, так как распространяющиеся волны сильно взаимодействуют с поверхностью волновода. Этот эффект можно лучше всего представить, если

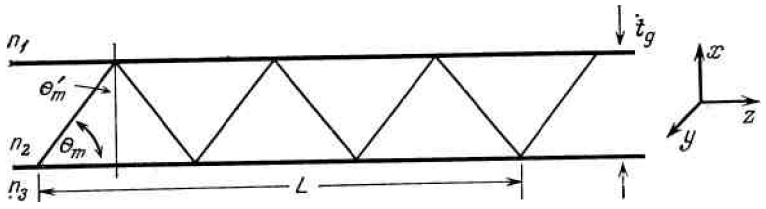


Рисунок 11 - Схема, иллюстрирующая оптико-лучевой подход к определению оптических потерь за счет рассеяния.

использовать понятия лучевой оптики, как это показывается на рисунке 11. Волна, распространяющаяся по волноводу, испытывает большое количество отражений. На длине L число отражений от каждой поверхности описывается выражением

$$N_R = L / 2t_g ctg\Theta_m, \qquad (26)$$

Для количественного описания величины оптических потерь обычно используют коэффициент экспоненциального ослабления. В таком случае интенсивность (мощность на единицу площади) в любой точке по длине волновода задается выражением

$$I(z) = I_0 e^{-\alpha z}, \qquad (27)$$

где I_0 — начальная интенсивность, т. е. интенсивность при z=0. Можно показать, что оптические потери в дБ/см связаны с коэффициентом ослабления α следующим соотношением:

$$L[дБ / см] = 4{,}3\alpha\left[см^{-1}\right], \qquad (28)$$

В общем случае поверхностное рассеяние вносит преобладающий вклад в потери в волноводах на основе диэлектрических (стеклянных или оксидных) пленок; этот вклад составляет около 0,5 - 5 дБ/см для моды низшего порядка и еще большую величину для мод более высокого порядка. Эти потери соответствуют поверхностным неоднородноетям размером около 0,1 мкм, которые обычно наблюдаются у волноводов, полученных на основе напыления тонких пленок. В полупроводниковых волноводах изменения толщины обычно поддерживаются на уровне 0,01 мкм; кроме того, абсорбционные потери гораздо больше, так что поверхностное рассеяние не столь важно.

Абсорбционные потери

В общем случае абсорбционные потери в тонких аморфных: пленках и кристаллических сегнетоматериалах, например $LiTaO_3$ или $LiNiO_3$, пренебрежимо малы по сравнению с потерями за счет рассеяния, если только нет загрязняющих атомов. Однако в полупроводниках имеют место существенные потери, связанные как с межзонным поглощением или поглощением на краю полосы, так и с поглощением на свободных носителях.

Межзонное поглощение

Фотоны оптического излучения, энергия которых больше величины запрещенной зоны материала, испытывают сильное поглощение, при этом они отдают свою энергию электронам, которые переходят из валентной зоны в зону проводимости. Обычно воздействие этого эффекта очень сильно, что приводит к коэффициентам поглощения в полупроводниках с прямой запрещенной зоной, большим 10^4 $см^{-1}$. Чтобы избежать межзонного поглощения, необходимо использовать длину волны, которая значительно длиннее волны, соответствующей краю полосы поглощения материала волновода. Ранее рассматривался метод добавления в бинарное соединение третьего элемента с целью получения такого трехкомпонентного соединения, которое имеет широкую запрещенную зону, как это имеет место при добавлении Al к GaAs до образования $Ga_{1-x}Al_xAs$. В общем случае подобный метод эффективен для соединений групп III—V и II—VI, поскольку проявляется забота о надлежащем выборе третьего элемента, чтобы достичь должного согласования кристаллических решеток эпитаксиальных слоев. Чтобы обеспечить достаточный сдвиг границы полосы поглощения и чтобы при этом длина волны используемого излучения находилась за хвостом кривой поглощения, третий элемент должен иметь приемлемую концентрацию. Этот эффект для случая $Ga_{1-x}Al_xAs$ иллюстрируется рисунком 12,

на котором приводятся экспериментально полученные кривые поглощения для длин волн, лишь немного больших длины волны, соответствующей :краю полосы поглощения.

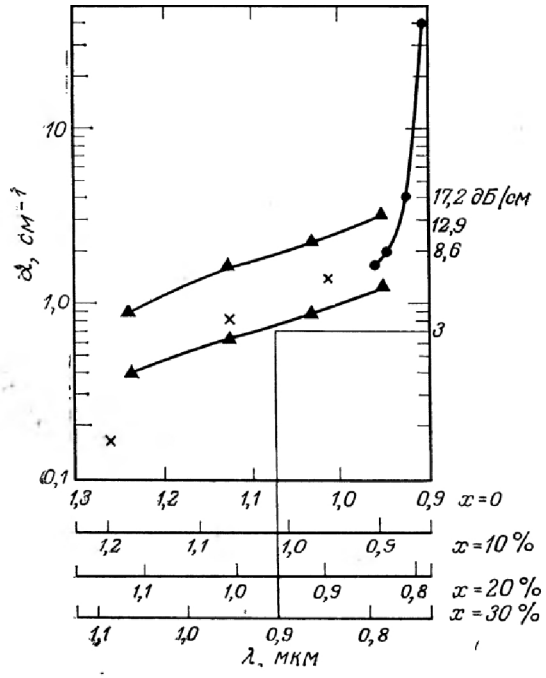


Рисунок 12 – Абсорбция излучения в области длинноволновой границы полосы собственного поглощения для $Ga_{1-X}Al_XAs$

Для иллюстрации сдвига края полосы поглощения в сторону более коротких волн (по мере увеличения концентрации Al) на рисунке использовано несколько шкал по оси абсцисс. Например, можно видеть, что концентрация атомов A1 (x=30%) приводит к уменьшению потерь за счет поглощения до 3 дБ/см для излучения с длиной волны 900 нм. В GaAs без добавления атомов A1 эти потери могут составлять около 215 дБ/см, или 50 $см^{-1}$. Допустимый уровень оптических потерь несомненно зависит от конкретного использования волноводов. Однако, поскольку большинство ОИС имеет размеры порядка сантиметра, во многих случаях следует признать приемлемыми потери на уровне 3 дБ/см.

Можно предотвратить межзонное поглощение, используя гибридные схемы, в которых применяется лазер, работающий на длине волны, значительно большей

длины волны, соответствующей краю полосы поглощения материала волновода. Независимо от того, каким методом избавляются от межзонного поглощения в полупроводниковых волноводах, для реализации практически приемлемых волноводов необходимы дополнительные меры, исключающие поглощение на свободных носителях.

Поглощение на свободных носителях

Поглощение на свободных носителях, которое иногда называют внутризонным поглощением, представляет собой процесс, при котором фотоны отдают свою энергию электронам, находящимся в зоне проводимости, или дыркам, находящимся в валентной зоне, переводя их на более высокий энергетический уровень. Обычно к поглощению на свободных носителях относят и поглощение, при котором электроны переходят с неглубоких донорных уровней вблизи дна зоны проводимости или при котором образуются дырки в валентной зоне за счет переходов электронов на неглубокие акцепторные уровни, находящиеся около потолка валентной зоны.

Излучательные потери

Распространяющаяся по волноводу оптическая мода может испытывать потери энергии за счет такого процесса, при котором фотоны уходят из волновода в окружающую среду и, следовательно, не вносят своего вклада в распространение волны по волноводу. Излучение может иметь место как в плоских, так и в канальных волноводах.

Излучательные потери в плоских и прямых канальных волноводах

Излучательные потери как для плоских, так и для прямых канальных волноводов обычно незначительны для хорошо ограниченных мод, которые

находятся в условиях, далеких от условия отсечки. Однако при отсечке вся энергия переходит в радиационные моды подложки, как это было показано ранее. Так как в волноводе моды более высокого порядка всегда находятся либо вне отсечки, либо по крайней мере ближе к отсечке, чем моды более низкого порядка, то излучательные потери больше у мод более высокого порядка. В идеальном волноводе моды являются ортогональными, поэтому передачи энергии из мод более низкого порядка в моды более высокого порядка не происходит. Однако при наличии неоднородностей и нерегулярностей в реальных волноводах может происходить модовое преобразование, что приводит к перекачке оптической энергии от мод более низкого порядка к модам более высокого порядка. В подобном случае, даже если какая-то мода хорошо ограничена, она может испытывать потери энергии за счет передачи ее оптической энергии модам более высокого порядка с последующим излучением. Подобная трудность не является характерной для типичных волноводов довольно хорошего качества, и потери, обусловленные этим процессом, как правило, пренебрежимо малы по сравнению с потерями за счет рассеяния или поглощения излучения. Имеется лишь одно важное исключение - это случай изогнутых канальных волноводов.

Излучательные потери в изогнутых канальных волноводах

В таких канальных волноводах потери за счет выхода излучения в окружающую среду сильно возрастают, что связано с дисторсией оптического поля, которую испытывают волны, проходя в волноводе по искривленному пути. В самом деле, минимально допустимый радиус кривизны волноводов определяется, как правило, не технологическими возможностями изготовления, а величиной радиационных потерь. Так как изгибы в волноводах являются необходимой частью большинства ОИС (исключая лишь самые простые), то потери на излучение света из искривленного волновода следует учитывать при разработке схем.

Тонкопленочные покрытия

Одним из первых и наиболее эффективных методов изготовления волноводов является осаждение тонких пленок диэлектрических материалов. В этом разделе слово «осаждение» используется в широком смысле, чтобы учесть методы нанесения пленок из вещества, находящегося в жидкой фазе, например, центрифугированием или погружением, а также испарением в вакууме и распылением.

Распыление диэлектрических пленок

Метод термического испарения диэлектрика в вакууме, который широко применяется в технологии получения тонких пленок для различных прикладных целей, например просветляющих покрытий, сравнительно редко используется для изготовления волноводов, поскольку при этом получают пленки с относительно высокими потерями в видимом диапазоне (10 дБ/см). Столь высокие оптические потери обусловлены включением атомов инородных веществ, которые действуют как центры рассеяния и поглощения оптического излучения. Вместо этого используют распыление молекул вещества, находящегося в твердом состоянии. Распыление - это процесс, в котором атомы или молекулы удаляются в вакууме с поверхности материала источника (мишени) за счет бомбардировки ионами с избыточной энергией примерно от 30 эВ и примерно до 2 кэВ. Такие атомы (или молекулы), вырываемые с поверхности материала мишени, осаждаются на поверхность подложки и образуют тонкий пленочный слой. Напыляемая тонкая пленка медленно растет за счет отдельных частиц, достигающих поверхности с достаточной кинетической энергией. При подобном процессе получают слои с высокой степенью однородности, поскольку осаждаемые атомы за счет движения равномерно распределяются по поверхности.

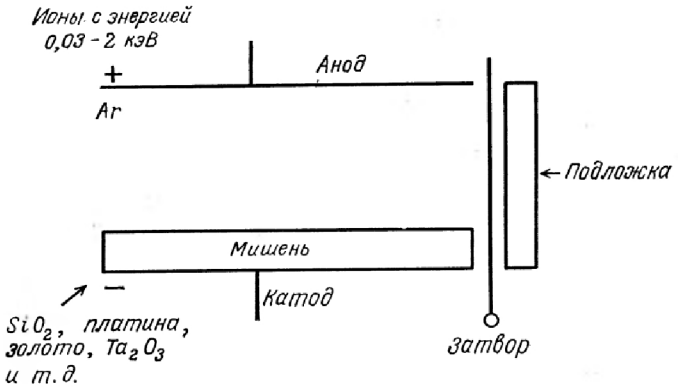


Рисунок 13 – Система для получения тонких пленок с помощью плазменного разряда

Вероятность внедрения инородных атомов крайне мала, так как процедура осуществляется при относительно низких температурах по сравнению с теми, которые необходимы для осаждения методом испарения вещества в вакууме, а материал мишени перед использованием можно хорошо очистить. В результате такой процесс дает возможность получать высококачественные оптические пленки с потерями излучения порядка 1 дБ/см.

Одним из методов нанесения пленок является плазменный разряд; соответствующая блок-схема представлена на рисунке 13. Мишень и подложка помещаются в вакуумную систему, затем подается газ под давлением $(2\text{-}20)*10^{-3}$ мм. рт .ст. Между анодом и катодом прикладывается высоковольтное напряжение смещения, достаточное для зажигания между ними плазменного разряда. Ионы, образующиеся при разряде плазмы, ускоряются по направлению к катоду. Они ударяют о мишень, отдавая при этом свой импульс атомам приповерхностного слоя материала мишени, которые при этом разлетаются в разные стороны, а затем осаждаются на подложке.

Чтобы напыленные атомы материала мишени прочно сцеплялись с подложкой (до получения однородных пленок с малыми оптическими потерями), последнюю перед помещением в вакуум следует тщательно обработать соответствующим растворителем или травителем. Подобная обработка подложки

является весьма важной операцией в любом процессе осаждения тонких пленок, причем способы чистки определяются материалами подложки и пленки.

Положение подложки на рисунке 13 является лишь одним из возможных среди многих геометрических конфигураций, так как атомы, распыляемые из мишени, стремятся осесть на поверхность любого предмета, находящегося внутри вакуумной системы. Для большей однородности напыляемого слоя расстояние между подложкой и мишенью должно быть больше размеров подложки. Однако при увеличении этого расстояния уменьшается скорость осаждения пленки, что приводит к увеличению времени, требуемого для получения пленки заданной толщины. Часто подложку для получения параллельной геометрии напыления крепят к аноду, что способствует получению более однородных слоев. Но в таком случае следует озаботиться об электрической изоляции подложки, если на нее подается отдельное напряжение смещения. Ее нагрев за счет падающих электронов может быть существенным. Затвор, представленный на рисунке 13, используется для экранирования подложки от осаждения в первые моменты времени, т. е. до установления режима плазменного разряда, так как в это время выбрасывается большая часть адсорбированных инородных: атомов.

Для исключения загрязнения осаждаемого слоя инородными атомами активных веществ в камеру, в которой производят плазменное напыление, обычно вводят один из инертных газов (аргон, неон или криптон). Однако в некоторых случаях выгоднее использовать реактивное напыление, при котором атомы,, вырываемые из мишени, вступают в реакции с бомбардирующими ионами, образуя оксидные или нитридные осажденные пленки. Например, можно использовать кремниевую мишень в присутствии аммиака до образования нитрида кремния.

Напряжение смещения, подаваемое на анод и катод, может быть как постоянным, так и высокочастотным с постоянной составляющей. Высокочастотное распыление, как правило, позволяет получать пленки лучшего качества, и при его использовании нет трудностей, связанных с накоплением

заряда в мишениях из материалов с низкой проводимостью, но, конечно, при этом требуется более сложный источник питания.

Кроме того, плазменное распыление можно применить для получения металлических пленок, а также диэлектрических слоев. Поэтому при необходимости его можно использовать для напыления электрических контактов на верхнюю поверхность волноводов или примыкающие к ней поверхности.

В качестве альтернативы плазменному получению ионов для бомбардировки мишени можно использовать дискретный источ-

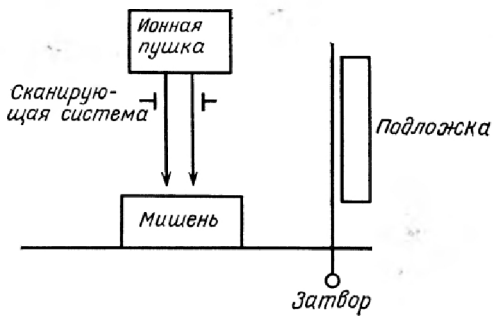


Рисунок 14 - Система для ионно-лучевого распыления.

ник или ионную пушку (рисунок 14), который дает коллимированный пучок ионов. К преимуществам ионно-лучевого распыления по сравнению с плазменным следует отнести то, что процесс напыления производится в высоком вакууме ($<10^{-6}$ мм. рт. ст.) и что сфокусированный ионный пучок взаимодействует лишь с мишенью. При этом загрязняющие атомы, выбиваемые из стенок камеры, не распыляются. Так как сфокусированный ионный пучок часто меньше мишени, то для обеспечения равномерного напыления используют электронное сканирование.

Кроме того, оксидные волноводы, например, из пятиокиси тантала и пятиокиси ниобия можно изготовлять с помощью распыления тантала и ниобия с последующим термическим окислением.

Осаждение из растворов

Для создания волноводов можно использовать большое число материалов путем нанесения на поверхность раствора, который при высыхании образует диэлектрическую пленку, центрифугирования подложек на центрифугах, как в случае работы с фоторезистом, для получения равномерного слоя или погружения подложек в раствор и их медленного вытягивания. Ниже приводится перечень некоторых общепринятых материалов для изготовления волноводов центрифугированием, которые можно использовать в данной технологии, и соответствующих растворителей для каждого из этих материалов (в скобках).

- фоторезисты (ацетон)
- эпоксидная смола (коммерческие растворители)
- полиметилметакрилат (хлороформ, толуол)
- полиуретан (ксилол)

Преимуществом таких волноводов является то, что они не дороги и для их изготовления не требуется сложного оборудования. Однако чистота материала хуже по сравнению с чистотой, достигаемой в пленках, полученных методом распыления, а неоднородность относительно выше. Тем не менее в ряде случаев получены удивительно хорошие результаты. Сообщается о пленках из эпоксидной смолы (потери всего 0,3 дБ/см на длине волны 632,8 нм), изготовленных усовершенствованным методом погружения, при котором подложка в горизонтальном положении сначала покрывалась смолой, а затем поворачивалась в вертикальное положение, чтобы лишняя смола могла стечь.

Методы, основанные на использовании замещающих атомных примесей

Хотя технология нанесения тонких пленок доказала свою эффективность при изготовлении волноводов из стекла и других аморфных материалов, ее принципиально нельзя использовать для изготовления оптических водноводов из

кристаллических материалов, например полупроводников или сегнетоэлектриков, подобных $LiNbO_3$ и $LiTaO_3$. Обычно это связано с тем, что разные составные элементы этих материалов конгруэнтно не осаждаются; относительная концентрация различных элементов не сохраняется при переходе вещества от источника на подложку. Даже когда имеется возможность подобрать такие условия, при которых имеет место конгруэнтное распыление, выращенный слой в принципе не является монокристаллическим и эпитаксиальным (с той же самой кристаллической структурой, как у подложки, на которой он выращен). Чтобы избежать подобных трудностей, было разработано несколько подходов для получения волноводов на основе кристаллических материалов без серьезного повреждения кристаллической решетки. Ряд методов использует внедрение атомов примесей, которые замещают некоторые атомы в решетке вещества, что приводит к увеличению показателя преломления этого вещества.

Диффузия атомов примесей

Так как волноводы для видимого и близкого инфракрасного диапазонов излучения обычно имеют толщину лишь несколько микрометров, то плодотворным для их изготовления является подход, основанный на диффузии примесных атомов с поверхуости подложки. При этом используются стандартные диффузионные методы, в которых подложка помещается в печь (обычно при 700-1000 °С), где находится источник примесных атомов. Такой источник может быть потоком газа, жидкости или твердой поверхностной пленкой. Например, для получения волноводов на основе $LiNbO_3$ и $LiTaO_3$ обычно используют диффузию атомов из металлических пленок титана или тантала, или металлорганический слой, полученный из раствора. Волноводы, получаемые с помощью диффузии металла, могут обладать потерями не более 1 дБ/см.

Диффузию металлов можно также использовать для получения волноводов на основе полупроводниковых материалов. Обычно примесь p-типа диффундирует в подложку n-типа или наоборот, что приводит к образованию p-n

перехода, обеспечивая, таким образом, электрическую изоляцию, а также распространение оптического излучения. Например, Мартин и Холл использовали диффузию Cd и Se в ZnS, а также Cd в ZnSe для изготовления как плоских, так и канальных волноводов. Было найдено, что оптические потери в канальных волноводах шириной 10 мкм и глубиной 3 мкм, образованных диффузией Cd в ZnSe, составляют менее 3 дБ/см.

Использование диффузии замещающих примесных атомов для изготовления волноводов не ограничивается полупроводниковыми материалами; ее можно также использовать в стекле. Например, Золатов с сотр. исследовали волноводы, полученные с помощью диффузии серебра в стекло. Отмеченное изменение в показателе преломления, обусловленное присутствием серебра, составило $\Delta n \approx 0{,}073$ на длинах волн 546,1- 632,8 нм. Из-за того что диффузия должна происходить при сравнительно низких температурах, ниже температуры плавления стекол, она является относительно медленным процессом. С этим фактом связано то, что для введения атомов примесей часто используют другие методы, в частности методы, основанные на ионном обмене и миграции.

Методы ионного обмена и миграции

Типичная блок-схема установки для изготовления волноводов методом ионного обмена и миграции приведена на рисунке 15. Материалом подложки является стекло, легированное натрием. Если подать электрическое напряжение, как показано на рисунке, и нагреть стекло примерно до 300 °С, то ионы Na^+ будут мигрировать в направлении катода. Поверхность стекла помещается в расплавленный нитрат таллия. Часть ионов Na^+ замещается на Ti^+, что приводит к образованию поверхностного слоя с большим показателем преломления.

Изава и Накагоме использовали метод ионного обмена и миграции для изготовления заглубленного волновода, находящегося в приповерхностном слое подложки. Их установка похожа на установку, приведенную на рисунке 15, за

исключением того, что использовалась подложка из боросиликатного стекла при температуре 530 °С. Процесс происходил в две стадии.

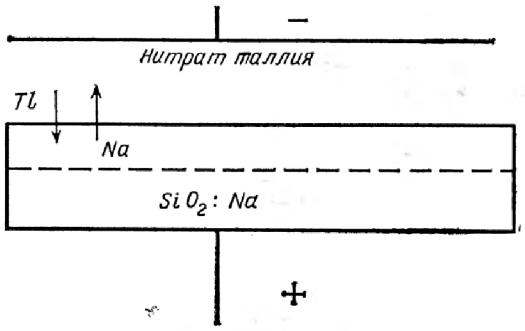


Рисунок 15 - Система для изготовления волноводов методом ионного обмена и миграции.

На первой стадии в состав солей, окружающих подложку (анод), входят нитрат таллия, нитрат натрия и нитрат калия, поэтому ионы Na+ и К+ в стекле обмениваются на ионы Т1+, что приводит к образованию на поверхности стекла слоя с большим показателем преломления. На второй стадии на аноде находились лишь нитрат натрия и нитрат калия. Следовательно, ионы Na+ и К+ диффундируют назад к «поверхности, тогда как ионы Ti+ перемещаются еще глубже внутрь подложки. Конечная структура имеет больший показатель преломления, а также легированный Ti+ слой, расположенный на некоторой глубине под поверхностью и закрытый снизу и сверху обычным боросиликатным стеклом. Таким образом был изготовлен заглубленный волновод.

Имплантация ионов

Замещающие примесные атомы можно также вводить путем имплантации ионов. При использовании такого метода легирования ионы требуемой примеси генерируются, затем ускоряются обычно до энергий 20-200 кэВ и после этого направляются на подложку. Указанный процесс, естественно, должен проводиться в вакууме. Основными элементами такой системы являются источник ионов, ускоряющие электроды, электростатические или магнитные дефлекторы

для сепарации ионов по массе и устройство сканирования коллимированным пучком поверхности подложки. Глубина проникания имплантированных ионов зависит от массы, энергии, а также типа материала подложки и ее ориентации. По тормозным характеристикам ионов и дозе имплантируемых ионов на 1 $см_2$, которую можно сравнительно точно определить по измерениям плотности тока в пучке и времени имплантации, можно рассчитать профиль распределения концентрации имплантированных ионов по глубине.

В большинстве случаев после проведения ионной имплантации необходимо провести отжиг материала подложки при повышенных температурах для удаления наведенных дефектов кристаллической решетки и для стимулирования замещения введенными ионами собственных ионов материала. После отжига волноводы, полученные таким методом, обладают почти такими же характеристиками, что и волноводы, полученные методом диффузии. Однако имплантационный метод представляет большие возможности для управления распределением концентрации введенных примесных атомов благодаря изменению энергии и дозы ионов с целью получения заданного профиля. В процессах диффузии всегда приходят либо к гауссовому профилю распределения примесных атомов, либо к профилю, описываемому дополнительной функцией ошибок в зависимости от типа источника диффузии

Помимо воздействия, связанного с замещением атомами примеси, ионную имплантацию можно также использовать для таких волноводов, в которых увеличение показателя преломления достигается в результате нарушений в структуре кристаллической решетки материала подложки, вызываемых ускоренными ионами, не принимающими участия в замещениях атомов.

Третий тип оптических волноводов можно получить методом ионной имплантации в некоторых полупроводниках, в которых могут возникать центры для захвата электрических носителей. Облучение протонами можно использовать для генерации повреждений в кристаллической решетке, приводящих к образованию компенсирующих центров, с целью получения области со сравнительно низкой концентрацией носителей в подложке, обладающей относительно

большой концентрацией носителей. Показатель преломления в области с более низкой концентрацией носителей несколько больше, так как свободные носители обычно уменьшают показатель преломления. Имплантированные протоны сами непосредственно не являются причиной заметного увеличения показателя преломления; его изменение связано с захватом носителей ловушками, которые понижают концентрацию носителей.

Волноводы, изготовленные путем уменьшения концентрации носителей

В полупроводниках любые свободные носители действуют «отрицательно» на показатель преломления, т. е. они уменьшают его значение по сравнению с тем значением, которое имел бы полностью обедненный носителями полупроводниковый материал. Таким образом, если каким-то способом удалить носители из определенной области подложки, то в ней показатель преломления будет иметь большее значение по сравнению с окружающей средой и она может действовать как волновод.

Гребенчатые волноводы, получаемые с помощью травления

Обычный метод изготовления гребенчатых волноводов состоит в нанесении на образец с плоским волноводом фоторезиста, облучении его ультрафиолетовым или рентгеновским излучением через контактно наложенную маску, которая и определяет форму волновода и проявлении фоторезиста, после чего на поверхности образца остается картина требуемой структуры, как это показано на рисунке 16.

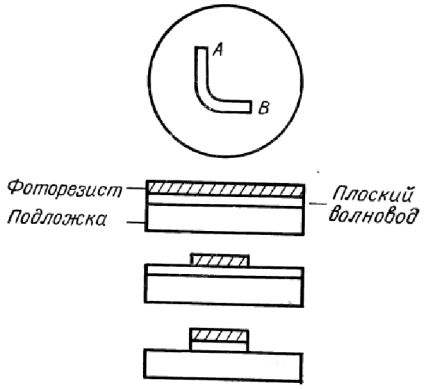


Рисунок 16 – Метод фоторезистивного маскирования для изготовления канальных волноводов

Фоторезист используют в качестве маски как в случае обычного химического травления, так и при использовании травления с помощью ионного пучка. Последний метод, который иногда называют тонкой размерной обработкой с помощью ионного пучка, дает более ровные кромки, особенно на изгибах, но при его использовании образуются радиационные дефекты в кристаллической решетке материала подложки, которые необходимо удалить методом отжига, если желательно получить волноводы с минимальными оптическими потерями.

Для получения структур канальных волноводов можно использовать и влажное химическое травление в сочетании с фоторезисторным маскированием. Такое травление не приводит к образованию радиационных дефектов, однако при его использовании очень трудно получить структуры с заданной глубиной и формой профиля. Действие большинства травителей чувствительно к ориентации кристаллической решетки протравливаемого материала, что приводит к образованию зазубрин на кромках в областях изгибов волноводов. Однако имеются случаи, когда данную анизотропию в действии травителей можно выгодно использовать, как это имеет место при получении прямоугольных мезаструктур полупроводниковых лазеров. По-видимому, заслуживает внимания метод, который основан на сочетании преимуществ ионно-лучевой размерной обработки и химического травления, а именно метод ионно-лучевого усиления травления, примененный Кавабе с сотр. к $LiNbO_3$. Они получили гребенчатые

волноводы шириной 2 мкм на основе Ti-диффузионного плоского волноводного слоя, используя полиметилметакрилат (резистивную маску) и электронно-лучевую литографию. Травление выполнялось сначала облучением образца ионами Ar+ с энергией 60 кэВ и с дозой $3*10^{15}$ $см^{-2}$ при комнатной температуре. Так как энергия ионов Ar+ была сравнительно высока, то при таком облучении образовывалось большое количество радиационных дефектов, тогда как с поверхности материала распылялось незначительное количество атомов. После облучения проводилось химическое травление разбавленной кислотой HF, которая в основном удаляет поврежденный слой. При использовании ионов Ar+ с энергией 60 кэВ глубина удаленного слоя составляла 700 А, но, повторяя эту процедуру (облучение с последующим травлением) несколько раз, можно осуществить более глубокое травление материала. Такое травление с усилением эффекта ионным облучением сочетает в себе более четкие границы рисунка, характерные для ионно-лучевой размерной обработки, и отсутствие радиационных дефектов, как это имеет место при химическом травлении.

Волноводы с наложенной полоской

Вместо формирования боковых стенок канальных волноводов путем травления волноводов с плоской структурой для достижения ограничения с боков можно использовать наложенную полоску из диэлектрика, которая наносится на верхнюю поверхность, как это было описано в ранее. Форму такой полоски

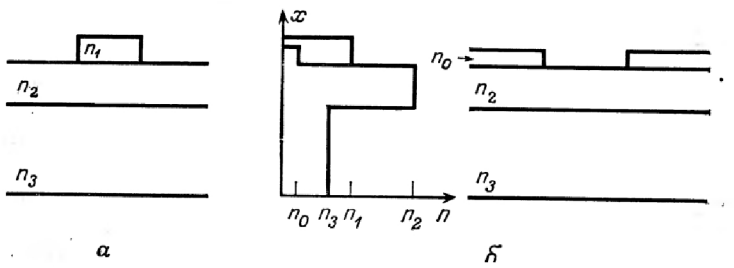


Рисунок 17 - Профили показателя преломления в волноводах с наложенными полосками из диэлектрика (а) и металла (б).

можно получить с помощью того же фотолитографического процесса, который применяется для получения гребенчатых волноводов. Для этих целей можно использовать либо ионно-лучевое, либо химическое травление в сочетании с фоторезистивным маскированием. Изготовлены оптические полосковые линии такого типа на основе GaAs, используя наложенные полоски из сильнолегированного GaAs с показателем преломления n_1, незначительно меньшим показателя преломления n_2 волноводного слоя, как показано на рисунке 17. Они использовали химическое травление с помощью NaOH и H_20_2, а также маскирование фоторезистом. На рисунке 17 приведен также волновод с наложенной полоской из металла и с показателем преломления n_0 в металле. Волноводы подобного типа можно также изготовлять, используя маскирование фоторезистом, а также либо ионно-лучевое, либо химическое травление.

Метод ионной имплантации или диффузии в сочетании с маскированием

Не все методы изготовления канальных волноводов основаны на начальном использовании волноводов с плоской структурой. Имеются такие методы изготовления канальных волноводов, которые основаны на использовании либо диффузии, либо ионной имплантации соответствующих легирующих атомов непосредственно в подложку, но через маску. Такие волноводы часто называют погруженными канальными волноводами, так как они находятся под поверхностью. В данном случае маскирование с помощью фоторезиста не является эффективным, так как не может выдерживать воздействие высоких температур, необходимое для проведения диффузии, и не обладает достаточной молекулярной массой для экранирования подложки от падающих на нее высокоэнергетических ионов. Обычно при диффузии в качестве маскирующих веществ используют пленки оксидов SiO_2 или $A1_2O_3$, а при имплантации - благородные металлы Au или Pt. Однако фоторезист используют для определения границы рисунка на маскирующем слое оксида или металла.

При изготовлении волноводов непосредственно имплантацией или диффузией атомов в подложку крайне важно оптическое качество материала подложки. Он должен обладать малыми оптическими потерями, а также иметь гладкую поверхность для предупреждения рассеяния. Одним из самых значительных преимуществ изготовления заглубленных волноводов методом имплантации или диффузии через маску является то, что это является планарным процессом; поверхность ОИС не нарушается выступами и впадинами, и поэтому не осложняется процесс оптического согласования при вводе и выводе излучения, а проблемы, связанные с загрязнением поверхности пылью или влагой, сводятся к минимуму. Другим достоинством этого процесса является то, что он не столь сложен по сравнению с процессами, которые были описаны в предыдущих разделах, так как нет необходимости в изготовлении плоской волноводной структуры по всей поверхности чипа ОИС перед формированием полосковых волноводов.

Рельефные оптические волноводы

В заключение целесообразно рассмотреть подход, затрагивающий скорее поточное производство, а не лабораторные исследования. Все методы изготовления волноводов, рассмотренные до сих пор, требуют использования относительно сложных, а следовательно, и дорогостоящих технологических процессов. Они подходят для лабораторных исследований или сложных ОИС, например анализаторов спектра радиочастот, которые будут выпускаться небольшими партиями и продаваться по высокой цене. А что делать с простыми ОИС, например ответвителями оптических линий передачи или делителями оптического сигнала, которые необходимы в больших количествах при малой стоимости? К настоящему времени для развития технологии изготовления схем такого рода сделано мало. Однако примером метода, ориентированного на массовое производство волноводов, является метод изготовления рельефных волноводов, разработанный Ульрихом с сотр. Последние изготовили ди-

электрические волноводы путем нанесения канавок в подложках из термопластика и заполнения их жидкостью, которая при отвердевании образовывала волноводные участки с большим показателем преломления. Материалом для подложек, который они использовали, служил полиметил-метакрилат (ПММА) (плексигласе), показатель преломления которого равен 1,49. В качестве заполнителя использовался циклогексилметакрилат.

Процесс изготовления таких волноводов достаточно прост. Штамп, имеющий форму необходимого волновода, вдавливался в подложку из ПММА при температуре около 100 °С. Подложке дают остыть, а затем штамп удаляют. Сперва выдавленный канал заполняют жидким циклогексилметакрилатом, просто разливая его по подложке, а затем ставят подложку вертикально, чтобы стек излишек жидкости. Облучая ультрафиолетом ртутной лампы мощностью 450 Вт, проводят полимеризацию» жидкости до тех пор, пока она не затвердеет и не превратится в полициклогексилметакрилат (ПЦГМА) с показателем преломления 1,505. Было найдено, что изготовленные таким образом волноводы при использовании в качестве штампа стеклянного волокна диаметром 7 мкм имели оптические потери 2- 4 дБ/см на длине волны 632,8 нм. В тех областях использования ОИС, в которых оптические потери не являются критическим фактором, подобная штамповка волноводов может привести к снижению издержек производства без неприемлемого ухудшения качества.

Лекция 7

Потери в оптических волноводах. Абсорбционные потери. Потери на излучение. Ввод-вывод излучения волноводов. Основы оптического согласования. Фокусирующие, торцевые, призменные, решеточные, сужающиеся элементы.

Каждый тип элементов связи обладает целым рядом достоинств и недостатков; не один из них не является наилучшим для всех случаев использования интегральной оптики. Поэтому как потребителям оптических интегральных схем, так и их разработчикам необходимо знать особенности каждого вида элементов связи.

Для изготовления элементов волноводной связи в общем случае используют такие методы, как фоторезистивное маскирование, распыление, эпитаксиальное выращивание слоев, которые были рассмотрены ранее. Однако в данной лекции рассматриваются некоторые специальные методы изготовления элементов связи, например экспонирование фоторезиста голографическими методами для изготовления решеточных элементов связи.

Основы оптического согласования

Основными характеристиками любого элемента связи являются его эффективность и медовая избирательность. Эффективность сопряжения, как правило, определяется частью общей энергии оптического пучка, введенной в волновод или выведенной из него. Кроме того, ее можно определять, исходя из потерь сопряжения (в дБ). Для элементов связи с модовой избирательностью эффективность сопряжения можно выразить независимо для каждой моды, тогда как в случае многомодовых элементов связи они, как правило, характеризуются общей эффективностью. Однако имеются случаи, когда можно определить относительные модовые эффективности и для многомодовых элементов связи.

Поэтому самое общее определение эффективности сопряжения дается следующим отношением:

$$\eta_{Cm} = \frac{энергия,\ \ введенная\ \ в\ \ моду\ \ m-го\ \ порядка}{полная\ \ энергия\ \ оптического\ \ пучка}, \qquad (1)$$

а потери сопряжения равны:

$$L_{Cm} = -10\log\left(\frac{энергия,\ \ введенная\ \ в\ \ моду\ \ m-го\ \ порядка}{полная\ \ энергия\ \ оптического\ \ пучка}\right). \qquad (2)$$

Если нет возможности определить энергию в каждой моде отдельно, то используют усредненное значении η_{Cm} и L_{Cm}. Эффективность сопряжения очень сильно зависит от согласования между полями оптического пучка и волновой моды. Этот принцип можно хорошо проиллюстрировать при рассмотрении случая поперечных элементов связи.

Поперечные элементы связи

Поперечные элементы связи представляют собой такие элементы, в которых пучок фокусируется непосредственно на специально подготовленное поперечное сечение волновода. Для случая пучка в свободном пространстве (воздухе) такую фокусировку можно выполнить с помощью линзы. Поперечное сопряжение двух твердотельных волноводов можно осуществить, соединяя встык отполированные или сколотые торцы волноводов.

Прямое фокусирование

Самым простым методом поперечного сопряжения лазерного пучка с волноводом является прямое фокусирование, или метод светящегося конца, представленный на рисунке 1.

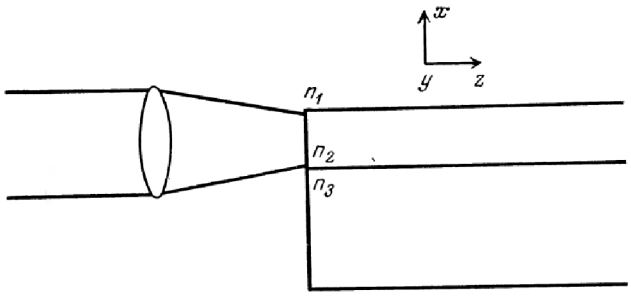


Рисунок 1 – Поперечное сопряжение методом светящегося конца.

Можно использовать как плоские, так и канальные волноводы, однако в данном случае допустим, что мы имеем дело с плоским волноводом. Передача оптической энергии лазерного пучка достигается за счет сопряжения полей пучка и волноводной моды. Эффективность сопряжения можно рассчитать из интеграла перекрытия картин полей падающего пучка и волноводной моды, который задается следующим выражением:

$$\eta_{Cm} = \left(\int A(x)B_m(x)dx\right)^2 / \int A(x)A^*(x)dx \int B_m(x)B_m^*(x)dx, \qquad (3)$$

где A(x) - амплитудное распределение входного лазерного пучка, а B(x) - амплитудное распределение моды m-го порядка.

Метод светящегося конца особенно полезен для сопряжения пучков от газовых лазеров и основной моды в волноводе, что -связано с относительно хорошим согласованием между гауссовым профилем лазерного пучка и формой TE_0-моды волновода. Конечно, для оптимального сопряжения диаметр пучка должен быть достаточно близким к толщине волновода. В принципе если контуры распределения полей тщательно согласованы, то можно достичь 100%-ной эффективности сопряжения. Однако на практике обычно достигают эффективности 60%. Это связано с тем, что толщины пленок волноводов имеют порядок 1 мкм, поэтому юстировка весьма критична. Такое сопряжение сравнительно часто используется в лабораторных условиях благодаря. Однако

трудности, связанные с монтажом оптической аппаратуры для столь точной юстировки без использования оптических скамей, существенно ограничивают его практическое использование.

Торцевое сопряжение встык

Поперечное сопряжение используется на практике в приложениях интегральной оптики в тех случаях, когда волновод сопрягается с полупроводниковыми лазерными источниками или с другим волноводом. При этом можно использовать стыковое соединение, как показано на рисунке 2. В таком случае можно достичь очень высокой эффективности сопряжения, что определяется тем, что ширину волновода можно приблизительно подогнать под ширину излучающего слоя лазера, и тем, что распределение поля в основной моде лазера достаточно хорошо согласуется с полем TE_0-моды волновода. Этот метод особенно полезен для сочленения лазерных диодов с плоскими волноводами, так как достичь эффективного сочленения инжекционного лазера с тонкопленочными волноводами трудно при использовании таких элементов связи, как призменные, решеточные элементы или сужающийся край пленки. Причина этого в том, что инжекционные лазеры дают сравнительно плохо коллимированные пучки с угловой полушириной расходимости обычно 10-20°. Как будет показано дальше в данной главе, призменные, решеточные элементы связи и элементы связи на утончающихся пленках крайне чувствительны к углу падения пучка излучения и требуют для эффективного сопряжения коллимации пучков менее 1°.

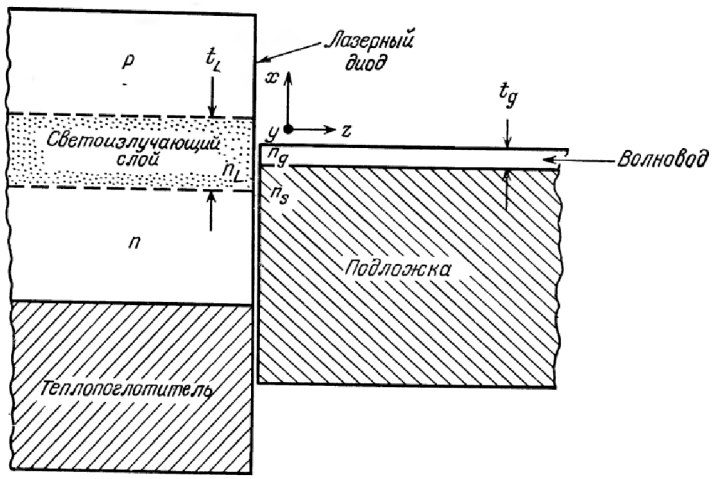


Рисунок 2 – Параллельное соединение лазерного диода с тонкопленочным волноводом

Для случая сочленения лазерного диода, работающего на основной моде TE_0, и плоского волновода, как это показано на рисунке 2, эффективность сопряжения на TE_0-моде дается выражением:

$$\eta Cm = \frac{64}{(m+1)^2\pi^2}\frac{n_L n_g}{(n_L+n_g)^2}\cos^2\left(\frac{\pi t_g}{2t_L}\right)\frac{1}{\left(1-\frac{t_g}{(m+1)t_L}\right)^4}\frac{t_g}{t_L}\cos^2\left(\frac{m\pi}{2}\right), \quad m = 0,1,2,\ldots . \qquad (4)$$

Это выражение основано на предположении, что все волноводные моды хорошо ограничены и $t_g \le t_L$. Интересно отметить отсутствие сопряжения для волноводных мод нечетного порядка, что определяется последним сомножителем в выражении (4). Это связано с тем, что распределения полей имеют компенсирующие лепестки, когда берутся интегралы перекрывания с четной (m = 0) модой лазерного излучения. Первый сомножитель в выражении (4) является нормировочным членом, второй обусловлен отражением, которое имеет место на границе лазер - волновод. Остальные сомножители связаны с рассогласованием распределений полей в лазере и волноводе.

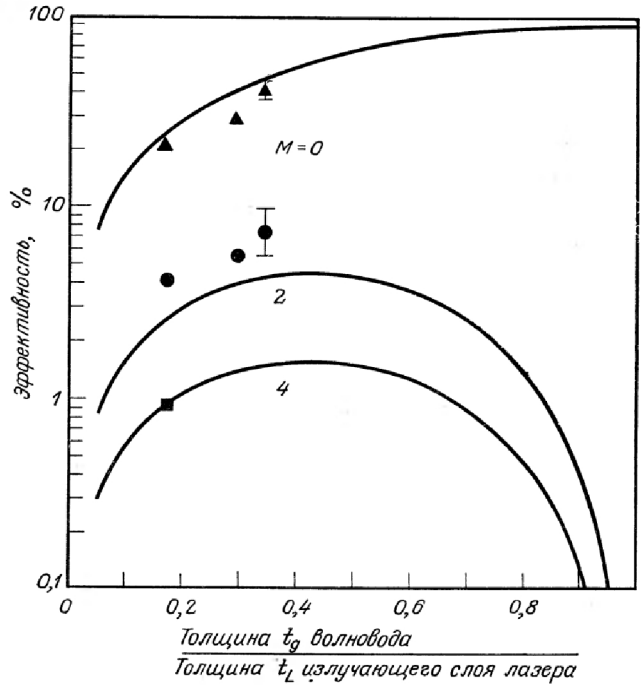


Рисунок 3 - Сравнение экспериментальных и теоретических значений эффективности связи в зависимости от толщины волновода: n_L=3,6, n_g=2,0, t_L=5,8 мкм.

Рассчитанные кривые зависимости η_{Cm} от относительной толщины волноводов представлены на рисунке 3 наряду с экспериментальными данными, полученными для сочленения лазерного диода на GaAs с волноводом из Ta_2O_5 на стеклянной подложке. Если $t_g \approx t_L$, эффективность сопряжения для моды наинизшего порядка теоретически приближается к 100%. В таком случае сопряжение для волноводных мод более высокого порядка почти равно нулю. Данные, касающиеся эффективности сопряжения на рисунке 3, являются оптимальными значениями, соответствующими идеальной настройке лазера и волновода. Эффективность сопряжения очень чувствительна к поперечным боковым рассогласованиям в направлении x. Смещение x волновода относительно лазера, как показано на рисунке 4, уменьшает эффективность сопряжения в соответствии с выражением:

$$P / P_0 = \cos^2(\pi X / t_L), \qquad (5)$$

где P_0 - энергия сопряжения при X=0. В этом выражении предполагают, что $t_g<t_L$ и $X \le (t_L - t_g)/2$. Штриховая линия на рисунке 4 соответствует теоретически вычисленным значениям P/P_0 для t_L=5,8 мкм и t_g=2 мкм, в то время как сплошные кривые представляют собой экспериментально полученные данные. Для определения относительной энергии, которая вводится при стыковом соединении для каждой из трех наблюдаемых мод, использовался призменный элемент связи на выходе волновода.

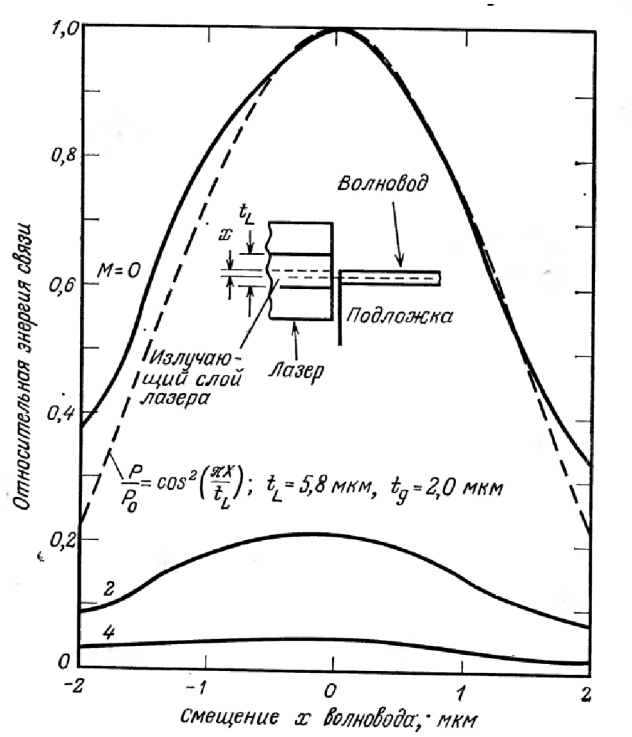


Рисунок 4 – Сравнение экспериментальных данных по эффективности связи (сплошная линия) с теоретическими (штриховая линия) в зависимости от поперечного смещения лазера и волновода.

Величина зазора между лазером и волноводом в направлении оси z также является очень критичной, и для достижения оптимального сопряжения она должна контролироваться с точностью порядка длины волны излучения. На рисунке 5 представлена зависимость экспериментально полученных изменений введенной мощности от смещения в направлении оси z. Осциллирующая форма

данной кривой связана с модуляцией эффективной отражательной способности выходной поверхности лазера за счет резонансных явлений в эталоне Фабри-Перо, образованном параллельными плоскостями границ раздела лазер—волновод. В принципе от этого эффекта можно избавиться, если использовать жидкость с промежуточным значением показателя преломления между лазером и волноводом. Если сделать это, то введенная мощность изменяется плавно в зависимости от смещения до оси z, как показано на рисунке 5 штриховой линией.

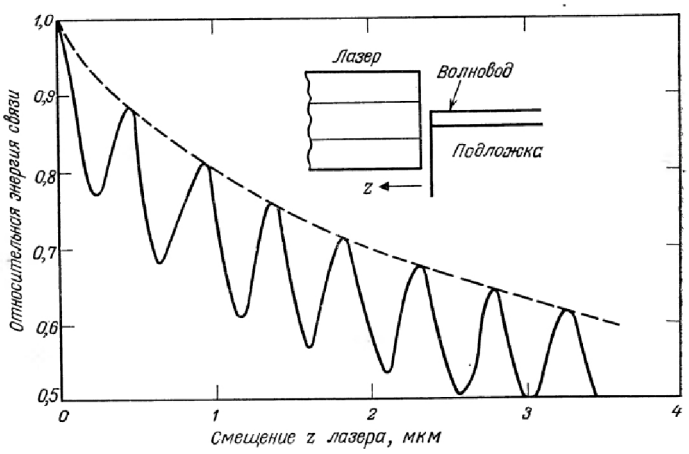


Рисунок 5 – Экспериментальная зависимость эффективности связи от величины зазора между лазером и волноводом

Представленные до сих пор результаты показывают, что стыковое соединение может быть достаточно эффективным для сопряжения лазерных диодов с волноводами. Кэмпбелл получил подобные результаты для стыкового соединения лазера и оптического волокна. Однако очевидно, что для получения оптимального сопряжения требуется субмикрометровая настройка. Настройку с такими допусками можно реализовать, если использовать микрометрические головки с пьезоэлектрическим приводом, которые отличаются тем, что имеют маленький пьезоэлектрический кристалл-держатель, закрепленный на конце обычного микрометрического винта. Грубая юстировка проводится за счет вращения микрометрического винта, а окончательная подстройка производится при подаче напряжения на пьезоэлектрическую насадку для передвижения лазера

(или волновода), в то время как вводимая оптическая мощность сопряжения измеряется с помощью соответствующего элемента связи и фотодетектора. Имеются в продаже пьезоэлектрические микрометры чувствительностью выше 40 А на вольт, работающие в диапазоне до 2000 В. Таким образом, настройку на уровне лучше 0,1 мкм можно сравнительно легко реализовать. После того как один раз -проведена подобная настройка, радиатор лазера можно прочно скрепить с держателем волновода с помощью эпоксидной смолы или металлического крепежа. Так как размер и масса лазера и ОИС сравнительно малы, чувствительность к вибрациям не является серьезной проблемой и настройка не нарушается. В самом деле, планируется, что стыковое сочленение будет использоваться для соединения лазерных источников с волноводами в интегрально-оптических анализаторах спектров высокой частоты.

Призменные элементы связи

Поперечное сочленение можно использовать лишь тогда, когда обнажена плоскость поперечного сечения торцевого конца волновода. Во многих случаях задача состоит в том, чтобы ввести излучение в волновод, находящийся внутри ОИС, когда доступна лишь одна поверхность. Если пытаться проводить фокусирование излучения на поверхность волновода при косом падении, как показано на рисунке 6, то это связано с принципиальными трудностями

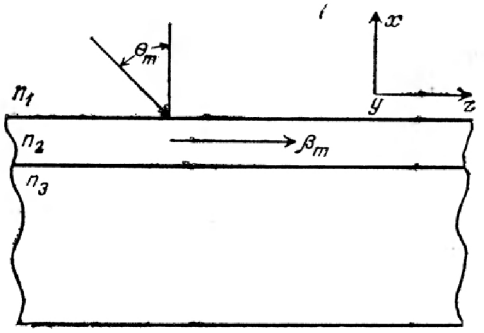


Рисунок 6 - Схема пробного эксперимента по наклонному введению света в волновод через его поверхность.

Чтобы сопряжение имело место, необходимо равенство составляющих фазовых скоростей волн в направлении оси z как для волновода, так и для пучка. Таким образом, надо удовлетворить условие согласования фаз, которое требует:

$$\beta_m = kn_1 \sin\Theta_m = (2\pi / \lambda_0) n_1 \sin\Theta_m . \qquad (6)$$

Однако в ранее было показано, что для мод в волноводе

$$\beta_m > kn_1 . \qquad (7)$$

Объединяя выражения (6) и (7), получаем $\sin\Theta_m > 1$, что, конечно, невозможно.

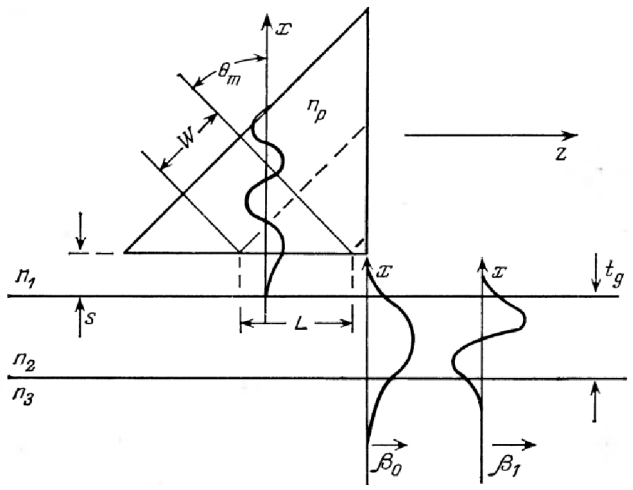


Рисунок 7 - Схема призменного элемента связи.

Одним из решений этой проблемы согласования фаз является использование призмы, как показано на рисунке 7. Пучок излучения шириной W направляется на основание призмы, которая имеет $n_p > n_1$. Пучок испытывает полное внутреннее отражение на поверхности раздела $n_p - n_1$, что приводит к установлению моды стоячей волны в призме (рисунок 7). Эта мода является стационарной в направлении оси x, но движется в направлении оси z с фазовой

постоянной β_m. В волноводе может существовать несколько мод, движущихся в направлении оси z с фазовыми постоянными β_m. Все такие моды имеют экспоненциальные «хвосты», немного заходящие за поверхность раздела $n_1 - n_2$. Если зазор s между призмой и волноводом настолько мал, что «хвосты» волноводных мод перекрываются хвостом моды призмы, то при выборе Θ_m таким образом, что $\beta_p = \beta_m$ имеет место когерентная перекачка энергии моды призмы в волноводную моду m-го порядка. Условие для согласования значений β определяется следующим соотношением:

$$\left(2\pi n_p / \lambda_0\right)\sin\Theta_m = \beta_m . \qquad (8)$$

Хотя возбуждение определенной моды будет иметь место при определенном значении угла Θ_m, с помощью одной-единственной призмы можно осуществить возбуждение нескольких различных мод просто за счет изменения угла падения оптического пучка. Падающий пучок ,не обязательно должен быть перпендикулярным к поверхности призмы. Но если направление падения пучка не перпендикулярно поверхности призмы, то условия преломления на поверхности раздела требуют изменения выражения (8) для угла Θ_m.

Процесс перекачки оптической энергии через перекрывающиеся «хвосты» мод, в то время как падающий пучок имеет тенденцию к полному внутреннему отражению в призме, иногда называют оптическим туннелированием по аналогии с туннельным эффектом в квантовой механике при прохождении частиц через энергетические барьеры. Моды волновода очень слабо связаны с модой призмы. Поэтому имеет место сравнительно небольшая деформация форм мод. Разумеется, условие:

$$\Theta_m > \Theta_C = \arcsin\left(n_1 / n_p\right), \qquad (9)$$

также должно выполняться для достижения полного внутреннего отражения, где Θ_C — критический угол полного внутреннего отражения.

Длина L, на которой происходит взаимодействие между модами призмы и волновода, определяется размерами призмы. Из теории слабосвязанных мод следует, что полный обмен энергией между фазово-согласованными модами имеет место, если длина взаимодействия в направлении оси z удовлетворяет соотношению

$$\chi L = \pi / 2, \qquad (10)$$

где χ — коэффициент связи. Коэффициент χ зависит от n_p, n_1 и n_2, определяющих форму «хвостов» мод, а также от зазора s между призмой и волноводом. Из выражения (10) можно определить длину, необходимую для полной перекачки:

$$L = W / \cos\Theta_m = \pi / 2\chi. \qquad (11)$$

Таким образом, для заданной длины L значение коэффициента связи, требуемое для полной перекачки, имеет вид

$$\chi = \pi \cos\Theta_m / 2W. \qquad (12)$$

При выводе этого условия для полной перекачки предполагается, что амплитуда электрического поля равномерно распределена по всей ширине W пучка излучения. На практике это никогда не реализуется. Для пучка гауссовой формы это приводит к максимальной эффективности связи, равной примерно 80%. Кроме того, видно, что для достижения 100%-ной связи, даже в случае равномерного распределения энергии по ширине пучка, правая граница пучка должна точно проходить через вершину прямого угла призмы. Если она проходит справа от этой точки, то часть падающей оптической энергии будет либо отражаться, либо непосредственно проходить в волновод и будет исключаться из моды призмы. Если же она будет проходить левее этой точки, то часть энергии, введенной в волновод при сопряжении, будет выводиться за счет сопряжения обратно в призму.

Призменные элементы связи довольно часто используются в различных областях применения интегральной оптики благодаря их универсальности. Их можно использовать в качестве как входных, так и выходных элементов. При использовании их в качестве выходных элементов призма располагается точно так, как изображено на рисунке 7, за исключением того, что направление распространения каналированного в волноводе излучения совпадает с отрицательным направлением оси z. Если в волноводе распространяется несколько мод, то вывод света происходит под своим определенным углом для каждой определенной моды. В связи с данной особенностью призменные элементы связи можно использовать в качестве приборов для анализа относительной мощности в каждой из волноводных мод. Кроме того, можно использовать призму для определения волноводных потерь, передвигая ее относительно волновода. Однако при этом следует обратить особое внимание на то, чтобы механическое давление, оказываемое на призму, было одинаково во всех измерениях, чтобы зазор, а следовательно, и коэффициент связи были бы постоянными.

Одним из недостатков призменных элементов связи является то, что величина n_p должна быть больше не только величины n_1 но и величины n_2. Это так, потому что показатель преломления n_2 волновода вообще близок к показателю преломления n_3 подложки, что приводит к выводу

$$\beta_m \approx kn_2 = (2\pi / \lambda_0)n_2 . \qquad (13)$$

Так как $\sin\Theta_m \leq 1$, то из выражения (13), связанного с (8), следует, что $n_p > n_2$. В случае стеклянных волноводов с показателями преломления ~1,5 сравнительно легко подобрать подходящий материал для призмы с $n_p > n_2$. Однако для полупроводниковых волноводов, материалы которых обычно имеют значения показателей преломления ~3 или 4, значительно труднее подобрать подходящий материал для призмы. При под боре материала для призм следует учитывать как показатель преломления, так и его .пропускание излучения на рабочей длине

волны. В таблице 1 перечислены материалы, пригодные по своему оптическому качеству для изготовления призменных элементов связи, а также их показатели преломления в различных волновых диапазонах.

Таблица 1 - Материалы, пригодные для изготовления призменных элементов связи

Материал	Приблизительное значение пока-	Диапазон длин волн
Титанат стронция	2,3	Видимая — ближняя ИК
Рутил	2,5	Видимая — ближняя ИК

Другим недостатком призменных элементов связи является то, что падающий пучок должен быть очень хорошо коллимирован из-за острой зависимости от угла эффективности связи для определенной моды. По этой причине призменные элементы связи нельзя эффективно использовать в сочетании с полупроводниковыми лазерами, которые дают пучки с угловой полушириной расходимости излучения 10 - 20°.

Призменные элементы связи достаточно хорошо подходят для проведения исследований в лабораторных условиях, где требуется большая маневренность относительно положения падающего пучка излучения. Однако необходимость постоянного механического давления для удержания призмы на месте делает их менее полезными для прикладных областей интегральной оптики при наличии вибраций и изменений температуры. Для преодоления этих трудностей без потери преимуществ модовой избирательности можно использовать решеточные элементы связи.

Решеточные элементы связи

Решеточные элементы, подобно призменным, используются для обеспечения фазового согласования между определенной волноводной модой и

оптическим пучком, падающим наклонно к поверхности волновода, как показано на рисунке 8. Необходимо помнить, что без решетки условия фазового согласования задаются выражением (6) и их нельзя удовлетворить при любом угле Θ_m.

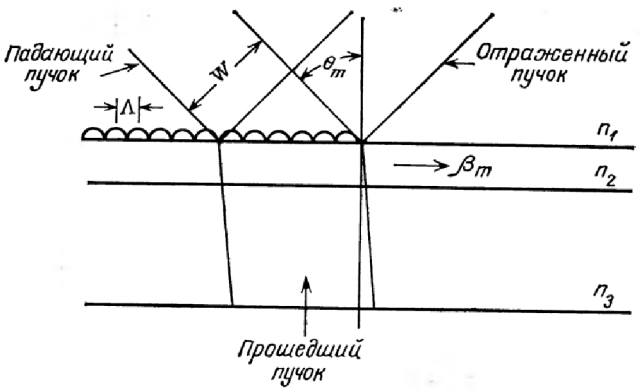


Рисунок 8 - Схема решеточного элемента связи.

Вследствие своей периодической структуры решеточный элемент связи возмущает волноводные моды в области, находящейся под решеткой, вызывая тем самым разложение каждой моды в ряд пространственных гармоник с постоянными распространениями в направлении оси z, определяемыми выражением

$$\beta_v \approx \beta_0 + (2\pi v / \Lambda), \qquad (14)$$

где v=0, ±1, ±2,..., а Λ — период решетки. Слагаемое β_0 приблизительно равняется β_m для определенной моды в области данного волновода, не закрытой решеткой. Так как величина v может быть отрицательной, то в таком случае условие согласования фаз (6) может выполняться так, чтобы:

$$\beta_v \approx k n_1 \sin\Theta_m, \qquad (15)$$

даже при $\beta_m > kn_1$. Так как все пространственные гармоники каждой моды согласованы для образования законченного поверхностного волнового поля в области решетки, энергия, введенная из пучка в любую пространственную гармонику, очевидно, переходит в основную (v=0) гармонику по мере ее распространения вправо и дальше за пределы области решетки. Эта основная гармоника очень близка к β_m-моде (и в конце концов становится ею) вне области решетки. Поэтому решеточный элемент связи можно использовать для избирательной передачи энергии от оптического пучка определенной моде волновода за счет соответствующего подбора угла падения пучка. Решеточный элемент связи можно использовать и для выведения энергии из волновода, так как в силу принципа взаимности энергия волноводной моды будет согласованно выводиться при определенных углах Θ_m, соответствующих конкретным модам. В предыдущем разделе описывалась работа решеточного элемента связи в обычных и относительно простых терминах. Однако детали явления такого сопряжения достаточно сложны и очень сильно зависят от формы поперечного сечения штрихов решетки, а также расстояния между ними. Как и в случае призменных элементов связи, теоретически можно достичь оптимальной эффективности связи около 80%, когда с помощью решетки вводят пучок света с гауссовой формой распределения. Однако типичные решетки «без блеска» (с симметричными профилями) обычно имеют эффективность 10 - 30%. Основной причиной такого положения является то, что большая часть падающей оптической энергии, как правило, проходит через волновод и теряется в подложке, так как в отличие от призменных элементов связи решетка работает в условиях, далеких от условий полного внутреннего отражения. Кроме того, энергия может уходить в дифракционные пучки более высокого порядка, производимые решеткой, несмотря на то, что отношение периода решетки к длине волны в волноводе составляет приблизительно 1.

Эффективность решеточного элемента связи можно значительно повысить, если придать асимметричный профиль его штрихам, чтобы иметь «блеск» решетки для получения оптимальных условий ввода света на рабочей длине

волны. Например, теоретический максимум эффективности связи моды TE_0 или TM_0 с пучком, выходящим из воздуха, для решетки с симметричным профилем составляет приблизительно 50%, тогда как для решетки с асимметричными пилообразными профилями штрихов - более 95%. Такие теоретические предсказания подтверждаются экспериментальными результатами, полученными на решетках «с блеском», и дальнейшими теоретическими работами.

Основное преимущество решеточных элементов связи заключается в том, что после изготовления они являются составной частью волноводной структуры. Следовательно, их эффективность сопряжения остается постоянной и заметно не изменяется при воздействии вибраций или окружающих условий. Кроме того, решеточные элементы можно использовать на волноводах из полупроводников с высокими показателями преломления, к которым трудно подобрать подходящий материал для изготовления призменных элементов. Однако из-за сильной зависимости эффективности решеточных элементов от угла их нельзя успешно использовать с относительно расходящимися пучками полупроводниковых лазеров. Может быть, самый большой недостаток решеточных элементов состоит в том, что они очень трудны в изготовлении, требующем использования сложных методов маскирования и травления.

Решеточную структуру можно сформировать либо маскированием с последующим травлением поверхности волновода, либо маскированием поверхности и нанесением на нее тонкопленочного рельефа решетки. Как при том, так и при другом методе самой трудной частью технологии является процесс получения четкой картины близко расположенных штрихов решетки. Расстояние между штрихами должно составлять порядка длины волны излучения в материале волновода. Поэтому для материалов с показателем преломления 1,4 - 4, используемых в видимой и ближней инфракрасной областях излучения, расстояния между штрихами решетки обычно находятся в пределах 100 - 300 нм. Обычные фоторезисты, используемые в микроэлектронной промышленности, имеют адекватное разрешение, однако практический предел обычных фотомасок составляет около 1 мкм. Поэтому обычно решеточные элементы связи

изготовляют с использованием оптического интерференционного процесса, который иногда называют голографическим (рисунок 9).

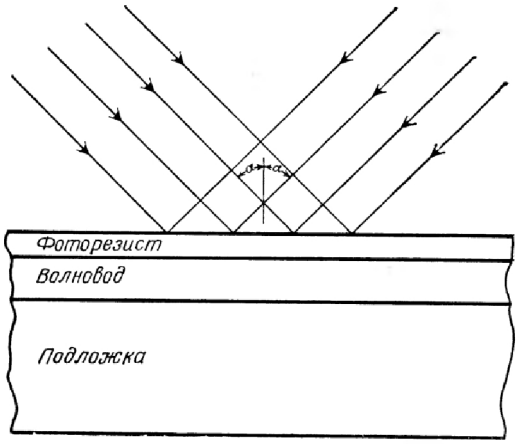


Рисунок 9 - Голографическое экспонирование фоторезиста,

При проведении такого процесса подложка с волноводом сначала покрывается фоторезистом. Затем данный фоторезист экспонируют под интерференционной картиной, которая образуется на поверхности от когерентных лучков лазера, как показано на рисунке 9 (Оба пучка обычно получают от одного лазера за счет использования расщепителя пучка). Простые геометрические соображения показывают, что соотношение между периодом решетки Λ и углом падения пучка α имеет вид:

$$\Lambda = \lambda_0 / 2\sin\alpha, \qquad (16)$$

Из выражения (16) очевидно, что Λ ограничено значением большим $\lambda_0/2$. Однако если использовать прямоугольную призму, как показано на рисунке 10, то при экспонировании можно получить решетки с меньшим расстоянием между штрихами, чем, при обычном использовании данного лазерного источника.

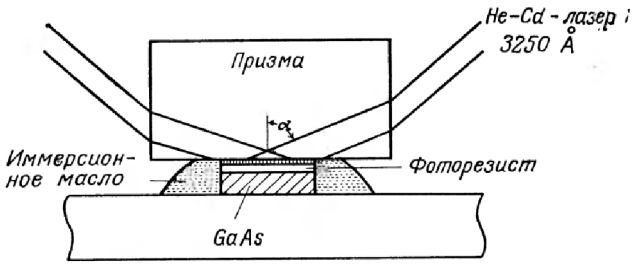


Рисунок 10 - Схема использования призмы в голографическом методе изготовления дифракционных решеток первого порядка в волноводе из GaAs.

В этом, случае соотношение (16) принимает вид

$$\Lambda = \lambda_0 / 2n\sin\alpha, \qquad (17)$$

где n - показатель преломления материала призмы. Используя призму из кварца и He-Cd лазер, получают решетки с расстояниями между штрихами до 115 нм. После облучения фоторезист обрабатывается с помощью обычных методов до получения необходимой маски на поверхности волновода. Далее для формирования структуры решетки используют либо химическое, либо ионное травление. Обычно считают, что при ионном травлении получаются решетки с более однородной структурой, а при химическом травлении материал волновода меньше повреждается. Кроме того, химическое травление при тщательном подборе травителя дает возможность получать более глубокие канавки, без побочной обработки материала. Например, при использовании раствора NH_4OH можно стравливать GaAs до глубины более 100 мкм. Другим преимуществом химического травления является то, что соответствующим подбором травителя и ориентации материала подложки можно достичь условий анизотропного травления, которое приводит к асимметричным решеткам «с блеском» на нужной длине волны.

В качестве варианта метода получения дифракционных решеток травлением служит метод получения решеток нанесением на волноводную поверхность тонкопленочных штрихов с использованием фоторезистивной маски для

ограничения формы штрихов. Фактически сам фоторезист можно использовать в качестве материала, наносимого в виде штрихов решетки. Обычно такие методы изготовления приводят к решеточным элементам связи с большими потерями (за счет рассеяния и поглощения излучения) по сравнению с элементами связи, изготовленными с использованием травления, так как часть пучка при выходе из волновода задерживается.

Сужающиеся элементы связи

Принцип действия сужающихся элементов связи основан на том факте, что волноводы толщиной, меньшей требуемой условием отсечки, передают энергию в излучательные моды.

Толщина волновода уменьшается в области связи, тем самым образуя волновод клинообразной формы с уменьшающейся критической длиной волны, как показано на рисунке 11. Механизм связи удобнее представить себе, если использовать оптико-лучевое приближение. Направляемая волна, подходя к данному элементу связи, испытывает зигзагообразные отражения с постоянно уменьшающимся углом падения на плоскость раздела волновод — подложка, который отсчитывается от нормали к этой плоскости. Когда значение этого угла падения станет меньше критического для полного внутреннего отражения, оптическая энергия будет переходить в подложку волновода. Энергия от последующих лучей также выходит из волновода аналогичным образом, поэтому можно достичь эффективности связи вплоть до 70%. Остальные 30% приходятся на потери, связанные в основном с излучением в воздух, когда толщина волновода достигает значения, соответствующего отсечке. После достижения точки отсечки мода волновода полностью выводится на расстоянии, составляющем около восьми длин волн излучения в вакууме. Главное преимущество сужающихся элементов связи состоит в том, что они просты в изготовлении и достаточно хорошо действуют в качестве выходных элементов связи, но формируют расходящийся луч (рисунок 11), который имеет угловые размеры 1-20° в зависимости от

величины сужения. Подобный расходящийся луч иногда неудобно использовать, однако такое положение допустимо во многих случаях использования, когда форма выходного луча не является критическим параметром.

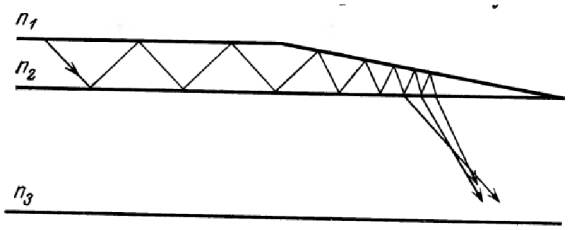


Рисунок 11 – Сужающийся элемент связи

В принципе сужающиеся элементы связи можно также использовать в качестве входных элементов. Однако, чтобы достичь высокой эффективности связи, необходимо сформировать сходящийся входной пучок излучения, который был бы обратным по отношению к расходящемуся пучку, представленному на рисунке 11. Так как это практически неосуществимо, то эффективность использования таких элементов связи на входе волновода крайне низка. Практическое использование сужающихся элементов связи возможно при сочленении тонкопленочных волноводов с оптическим волокном, так как конец волокна можно расположить очень близко к волноводу, и этой торцевой поверхности волокон можно придать форму, необходимую для увеличения эффективности связи. Такой тип использования подробно рассматривается в следующем разделе, посвященном связи волноводов с волокном.

Элементы связи волокно - волновод

Ключевым элементом интегральных оптических систем является элемент связи, передающий оптические волны между волоконным световодом, используемым для передачи сигналов на большие расстояния, и волноводом, входящим в состав ОИС, используемых для обработки сигналов. Исследования последних лет привели к разработке ряда различных видов элементов связи

волокно - волновод. Хотя часть из них обладает достаточной эффективностью, нельзя сказать, что найдено окончательное решение проблемы, в частности более трудной задачи сопряжения с одномодовым оптическим волокном, имеющим малый диаметр светопроводящей жилы. Однако работы в данном направлении продолжаются на обнадеживающем уровне, и не видно серьезных теоретических ограничений для достижения эффективной связи; необходимо преодолеть лишь чисто практические трудности.

Волокно можно непосредственно состыковать с волноводом без какого-либо промежуточного устройства. Если поперечные сечения сердцевины волокна и волновода достаточно близки по размерам, то можно достичь высокой эффективности соединения, как и в случае стыкового соединения канальных волноводов. Для уменьшения потерь на границе раздела можно использовать жидкость с промежуточным показателем преломления. Главная проблема при использовании подхода, основанного на применении стыкового соединения, состоит в том, что крайне трудно достичь и поддерживать правильную юстировку, поскольку сердцевина волокна и волновод, как правило, имеют размеры порядка 1 мкм.

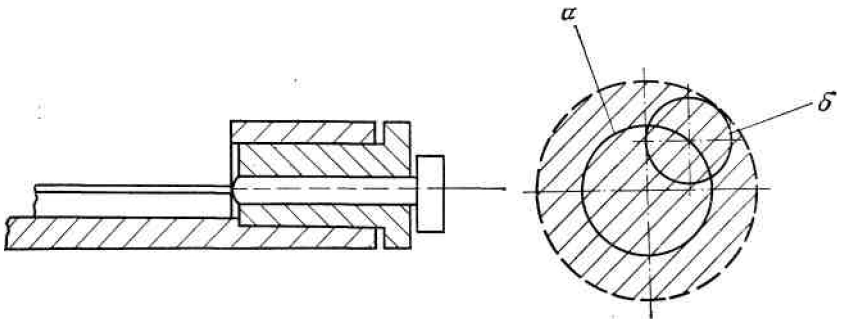


Рисунок 12 - Устройство связи волокно – волновод. Вращением внутренней втулки достигается перемещение волокна по окружности (б), а вращением большой втулки — по окружности (а). Таким образом, волокно можно перемещать по всей заштрихованной площади.

Был предложен механический соединитель, который облегчает юстировку волокна с волноводом. Волокно эксцентрично закрепляется в соединительной

втулке, которая в свою очередь монтируется эксцентрично во второй большей втулке, как показано на рисунке 12. Вращательное движение обеих втулок дает возможность перемещать волокно по значительной площади в пределах соединителя, обеспечивая тем самым его юстировку с волноводом.
Совсем недавно для решения проблемы юстировки канального волновода с одномодовым волокном были использованы методы V-образных канавок в кремнии и монтажа с переворотом пластины кремния, которые уже ряд лет применяются в технологии промышленного изготовления электрических интегральных схем. Шим с сотр. использовали подобный элемент связи для соединения диффузионного волновода на основе $LiNbO_3$ шириной 3 мкм с жилой одномодового волокна диаметром 4,5 мкм. Они достигли эффективности связи 75% как на TE-, так и на TM-модах.

Другой метод стыкового соединения, позволяет избежать трудностей, связанных с юстировкой, так как при этом используется стыковое соединение конца волокна с верхней поверхностью, а не с торцом волновода. Капля жидкости с большим показателем преломления используется для получения получить утечки оптической энергии. Величина этой утечки зависит от угла, однако при соответствующей ориентации волокна была достигнута эффективность связи порядка 50%. Этот метод хорошо себя зарекомендовал в лабораторных условиях, но он мало полезен в прикладных областях интегральной оптики из-за трудностей, связанных с сохранением юстировки.

Элементы связи сужающейся пленки с волокном

Можно использовать сужающуюся часть волновода для связи с волоконным световодом, как показано на рисунке 13, а.

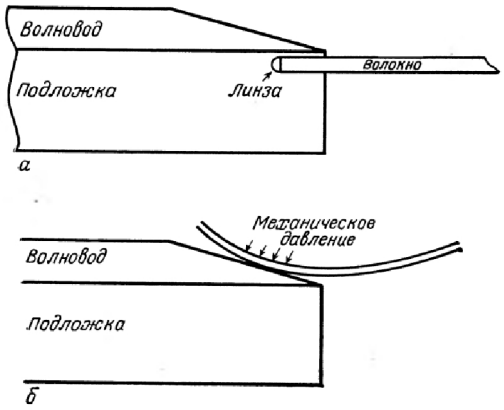


Рисунок 13 - Элементы связи сужающейся пленки с волокном.
а - связь через подложку; б - связь через поверхность.

Волокно вставляется в цилиндрическое отверстие, просверленное в подложке чуть ниже волновода. Полусферическое дно этого отверстия заполняется материалом с высоким показателем преломления, чтобы сформировалась линза. Световые волны, выведенные из тонкопленочного волновода в подложку, собираются волокном. В случае многомодового волоконного световода с достаточно большим значением углов ввода достигнута приемлемая эффективность связи. Например, в одном эксперименте, в котором показатели преломления пленки, подложки, волокна и линзы равнялись 1,53, 1,49, 1,51, 1,55 соответственно, свет длиной волны 632,8 нм от He-Ne лазера был введен с эффективностью 60%.

Другой подход к сопряжению волокно - волновод с использованием сужающейся пленки был продемонстрирован Техом и Стегманом. Вместо того чтобы вставлять волокно в подложку под сужающейся частью волновода, они сняли оболочку волоконного световода, а затем прижали небольшой участок обнаженной жилы к поверхности сужающейся части волновода, как показано на рисунке 13, 6. Так как при этом использовалось многомодовое волокно диаметром 60 мкм, то оно собирало около 90% излучаемого из волновода света.

Решеточные элементы связи для соединения волокно – волновод

Хаммер с сотр. показали, что решеточные элементы связи можно использовать для связи волокна малого показателя преломления с волноводом из материала высокого показателя преломления. Вдавливанием неэкранированного многомодового волокна из пластика (ПММА) (n≈1,5) в решетку, полученную травлением диффузионного волновода на основе $LiNb_xTa_{1-x}O_3$ (n=2,195), они достигли эффективности связи 6%. Однако их расчеты показывают, что при оптимальном «блеске» дифракционной решетки можно достичь эффективности 86%.

Балмер и Уильсон также использовали решетку для связи одномодового волоконного световода со стеклянным волноводом, полученным методом распыления. В этом случае защитная оболочка не удалялась, а волокно нагревалось и вытягивалось для уменьшения поперечного сечения волокна диаметром 18 мкм на одну треть (диаметр светопроводящей жилы 3 мкм). Такое уменьшение толщины оболочки допускает туннелирование оптической энергии, если решетка обеспечивает фазовое согласование. Эффективность составляла лишь 0,4%, однако такая малая величина обусловливалась в основном крайне малой площадью согласования между волокном и используемым широким плоским волноводом. Если провести расчет поправки на рассогласование за счет такой разницы в размерах, то с ее учетом теория предсказывает эффективность связи в 30% для канальных волноводов тех же размеров, что и сердцевины волоконного световода.

Связь между волноводами

Явление оптического туннелирования можно использовать не только для связи волоконного световода или пучка света с волноводом, но и для связи между волноводами. Элементы связи подобного типа, как правило, называются направленными ответвителями, поскольку в данном случае энергия передается когерентно с сохранением направления -распространения. Направленные ответвители изготовляются в виде двух основных геометрических структур: многослойной панарной структуры и двух параллельных канальных волноводов, расположенных достаточно близко друг от друга.

Многослойные плоские волноводные элементы связи

Хотя стыковое соединение можно использовать для двух плоских волноводов, как упоминалось в ранее, более обычным методом является тот, при котором волноводы располагают в непосредственной близости друг от друга и создают условия, необходимые для связи за счет фазово-когерентного перехода энергии (оптического туннелирования), как показано на рисунке 14.

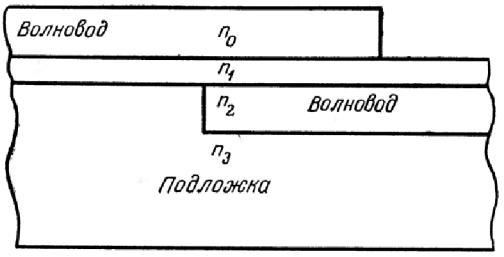


Рисунок 14 - Связи между двумя плоскими волноводами за счет оптического туннелирования.

Переход энергии осуществляется при наличии фазового синхронизма через изолирующий слой с показателем преломления n_1.

При этом показатели преломления n_0 и n_2 волноводных слоев должны быть больше n_1 и n_3, а толщина промежуточного слоя *l* должна быть настолько мала, чтобы перекрывались экспоненциальные «хвосты» мод, распространяющихся в волноводах. Чтобы имел место переход энергии из одного волновода в другой, они должны иметь одинаковые постоянные распространения. Таким образом, показатели преломления и толщины волноводных слоев должны тщательно контролироваться, чтобы обеспечить согласование постоянных распространения. Использование такого труднодоступного метода дает прекрасные результаты. Для достижения оптимального согласования, как и в случае устройств, использующих синхронизм связанных волн, например призменный элемент связи, должна тщательно выбираться длина взаимодействия. Условия для полной передачи энергии снова даются выражением (10), однако значение к для пары перекрывающихся плоских волноводов, очевидно, будет отличаться от значения для призменного элемента связи.

Данный метод связи трудно использовать для волноводов, полученных нанесением тонких пленок, так как такая технология не позволяет столь жестко контролировать получаемые величины показателей преломления и толщины волноводов. Однако в случае эпитаксиально выращенных волноводов при наличии очень точного контроля толщин и показателей преломления (за счет контроля состава) проблему согласования постоянных распространения мод решить гораздо легче. Очевидно, этот метод особенно хорошо подходит для связи интегрально-оптических лазерных диодов с волноводами. Например, с использованием многослойной планарной структуры для связи CdS_xSe_{1-x} лазера с волноводом, изготовленным из того же материала. В этом случае промежуточный изолирующий слой был образован из GdS. Данный метод также был использован авторами работы для связи лазеров и волноводов на основе GalnAsP, разделенных тонким слоем из InP. Такую структуру сдвоенных волноводов можно использовать для изготовления эффективных одномодовых лазеров с распределенной обратной связью, обеспечиваемой решетками. Можно изготовить лазеры с очень большой эффективностью, если использовать два различных, но связанных друг с

другом волновода, чтобы изолировать активную область, в которой идет генерация фотонов от области распределенной обратной связи.

Двухканальные направленные ответвители

Двухканальные направленные ответвители, являющиеся аналогами двухволновых дырочных микроволновых ответвителей, в принципе состоят из параллельных канальных оптических волноводов, расположенных настолько близко друг от друга, что энергия излучения переходит из одного волновода в другой путем оптического туннелирования, как показано на рисунке 15. Такой переход энергии обеспечивается за счет связи в условиях когерентности и синхронизма между перекрывающимися экспоненциальными «хвостами» мод, каналированных в каждом из волноводов. Фотоны моды, распространяющейся в одном из волноводов, например в волноводе 0, переходят в моду, которая распространяется в волноводе 1, сохраняя фазовую когерентность во время такого перехода. Этот процесс происходит кумулятивно на значительной длине; следовательно, для достижения такой синхронной связи свет должен распространяться с одинаковой фазовой скоростью в каждом канале.

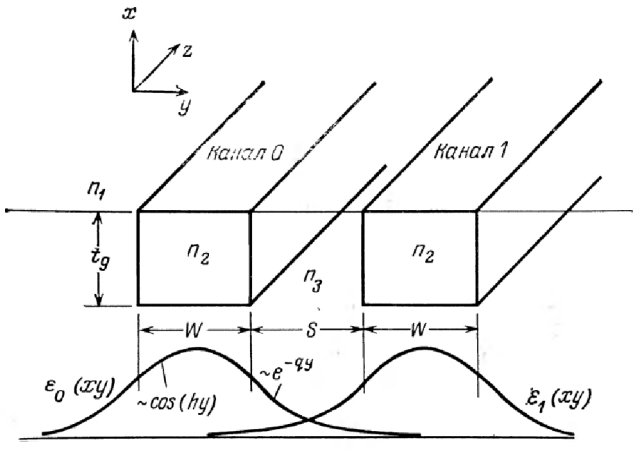


Рисунок 15 - Схема двухканального направленного ответвителя. Внизу представлено характерное для них распределение амплитуд электрической составляющей.

Часть переданной энергии на расстоянии в единицу длины определяется перекрытием мод в отдельных каналах. Поэтому она зависит от расстояния между каналами s, длины взаимодействия L, а также проникновения моды в пространство между каналами, которое можно охарактеризовать коэффициентами ослабления p и q, характерные дл волноводных мод.

В двухканальных ответвителях энергия попеременно переходит из одного волновода в другой и обратно, если длина взаимодействия между ними является достаточно большой. Если бы было проведено измерение плотности распространяющейся оптической энергии в направлении оси z (вдоль одного из каналов направленного ответвителя), то модно было бы наблюдать зависимость синусоидального распределения от расстояния.

Согласование светоизлучающих диодов (СИД) с волоконным световодом с помощью микролинз

Непосредственная стыковка поверхностного светодиода с волоконным световодом (рисунок 16) приводит к значительным потерям мощности излучения. Если излучение с диаграммой направленности, пропорциональной $\cos^m\Theta$ ($m \geq 1$), вводится в волоконный световод с площадью сердцевины, большей или равной A, и невысокой числовой апертурой NA, то при учете только меридиональных лучей потери при стыковке составят:

$$\alpha = 10\lg\left[2/\left((m+1)NA^2\right)\right],$$

Следовательно, при m=1 (ламбертов источник) и NA = 0,1 потери составят 20 дБ.

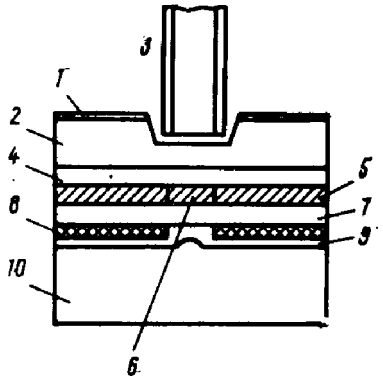


Рисунок 16 - Непосредственная стыковка ВС с СИД на основе двойной гетероструктуры GalnAsP/InP: 1 - контактный слой n-InP; 2 - подложка n-InP; 3 – волоконный световод; 4 - n-InP; 5 - p-GaInAsP; 6 - активная область; 7 - p-InP; 8 - защитный слой SiO_2; 9 - контактный слой p-InP; 10 - теплоотвод.

Для уменьшения потерь при стыковке можно использовать микролинзы, нанесенные на излучающую площадку светодиода или торец волоконного световода. По сравнению с непосредственным торцевым соединением применение волоконного световода со сферическим торцом увеличивает эффективность ввода излучения в 3-5 раз. Использование светодиода со сферическими линзами (рисунок 17) дает рост эффективности ввода излучения в 15-20 раз. Максимум эффективности ввода излучения в световоде при использовании сферических линз и поверхностных светодиодах имеет место при отношении диаметров световода и излучающей площадки светодиода, не превышающем 5 (для световода с NA=0,2) и 2,5 (для ВС с NA = 0,4).

Гораздо проще решается проблема стыковки при использовании излучателей, торцевого типа. Конструкция их весьма схожа с конструкцией мезаструктурного инжекционного лазера без резонансной обратной связи. Активная область торцевого светодиода сверху и снизу ограничена слоями полупроводника с меньшим показателем преломления, образующим, двойную гетероструктуру, а с боков - канавками, расстояние между которыми выбирается под диаметр сердцевины световода. Образованный вокруг активной области волновод достаточно эффективно каналирует свет к излучающей поверхности

светодиода, площадь которой много меньше площади активной области. Поэтому если перепоглощение излучения в активном слое мало, то яркость излучения на выходе торцевого светодиода может существенно превысить яркость поверхностного источника.

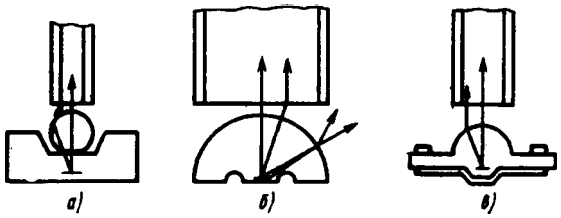


Рисунок 17 - Методы увеличения эффективности ввода излучения с помощью микролинзы, наклеенной на излучающую площадку (а), придания выходной поверхности СИД сферической формы (б) и интегральной линзы, выращенной на подложке СИД (в)

Вытравив сзади активной области канавку, параллельную торцу, или придав ей конфигурацию уголкового отражателя, можно еще дополнительно увеличить яркость торцевых светодиодов.

Расходимость по уровню половинной мощности в плоскости p-n перехода у торцевых СИД составляет 60...80° (в зависимости от длины СИД), а в перпендикулярной плоскости 20...60°. Управляющие токи для СИД составляют обычно 100 -..200 мА, а выходная мощность 1... 6 мВт. Диаграмму направленности торцевых СИД можно улучшать, используя структуру большого оптического резонатора, в котором очень тонкий активный слой помещаете» внутрь достаточно толстого волновода с малым коэффициентом поглощения излучения. Наличие раздельного ограничения излучения и носителей позволяет уменьшить расходимость излучения и увеличить его эффективность ввода в ВС. Кроме того, увеличение плотности носителей в активной области сокращает их спонтанное время жизни и увеличивает тем самым полосу частот модуляции

излучения. Использование микролинзового ввода излучения торцевого СИД в ВС позволило увеличить эффективность связи примерно в 5 раз.

Сравнение эффективности ввода в многомодовый световод мощности от поверхностного и торцевого СИД показывает, что при малой числовой апертуре световода (NA<=0,2) торцевой СИД на 20..50% эффективнее поверхностного (при диаметре сердцевины световода, который больше диаметра излучающей площадки поверхностного СИД и ширины активной области торцевого СИД). При числовой апертуре световода NA=0,3 эффективность ввода излучения для обоих типов СИД сравнивается, а при NA=0,4 поверхностный СИД обеспечивает большую введенную мощность, чем торцевой (до 0,5 мВт при диаметре сердцевины световода 50 мкм).

Выпускаемые торцевые СИД на GaAsP/InP с длиной волны излучения 1300 нм дают на выходе отрезка одномодового световода, состыкованного с ним в торец, мощность до 100 мкВт при токе до 100 мА. Поверхностные СИД, изготовленные из того же материала, дают в тех же условиях в несколько раз меньшую мощность при стыковке с одномодовым световодом. Суперлюминесцентный диод благодаря усилению сигнала в активной области и лучшей диаграмме направленности имеет на выходе отрезка одномодового световода, состыкованного с ним в торец, мощность до 0,5 мВт при собственной мощности 5 мВт.

Лекция 8

Геометрическая оптика планарных волноводов. Классификация мод планарного волновода. Волноводные моды тонкопленочного волновода. Эффективная толщина волновода. Градиентные планарные волноводы.

Лучевое описание плоских диэлектрических волноводов. Направляемые лучи.

Опишем плоский трехслойный диэлектрический волновод на языке геометрической оптики. Для этого рассмотрим однородный диэлектрический слой с показателем преломления n_1 и толщиной 2a, расположенный между двумя полубесконечными диэлектрическими слоями с более низкими показателями преломления n_2 и n_3, как показано на рисунке 1. На этом рисунке изображена также зигзагообразная траектория луча, распространяющегося под углом φ к оси z такого волновода. Для определенности предположим, что

$$n_1 > n_2 \geqslant n_3, \quad (1)$$

т. е. при $n_2 \neq n_3$ имеем асимметричную, а при $n_2=n_3$ - симметричную волноводную структуру. Поскольку у этой структуры две границы раздела, уже рассмотренные ранее, то понят-

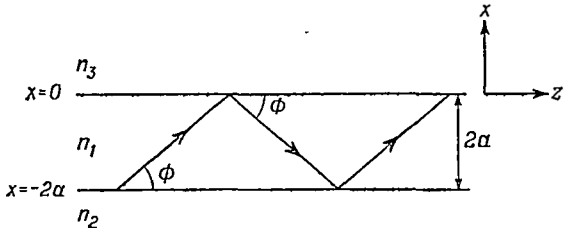


Рисунок 1 - Зигзагообразная траектория луча в плоском диэлектрическом волноводе.

но, что поведение луча в таком волноводе существенным образом зависит от угла φ. На рисунке 1 изображен случай, когда углу φ соответствует угол падения, величина которого превышает критический угол для каждой границы раздела. При $\varphi=\pi/2-\Theta_1$ это означает, что мы имеем дело с углами, для которых выполняется соотношение

$$\cos\phi \geqslant n_2/n_1 \geqslant n_3/n_1. \quad (2)$$

При таких величинах угла φ луч претерпевает полное внутреннее отражение на каждой границе раздела и, следовательно, весь свет распространяется в центральном слое волноводной структуры, что соответствует распространению направляемой моды волновода. Было показано, что распространение волн в направлении оси z описывается величиной exp(iβz), причем продольная постоянная распространения β связана с углом φ следующим образом:

$$\beta = n_1 k \cos\phi. \quad (3)$$

Следовательно, согласно (2) и (3), постоянная распространения заключена в интервале:

$$n_1 k \geqslant \beta \geqslant n_2 k \geqslant n_3 k. \quad (4)$$

Характеристическое уравнение

Мы видим, что минимальное значение постоянной распространения β, при котором возможно волноводное распространение (т. е. условие отсечки моды), имеет вид $\beta=n_2k$. Однако изменение величин β и φ между соответствующими

пределами происходит не непрерывно, как предполагалось до сих пор, а дискретно. Чтобы учесть дискретный характер изменения этих величин, рассмотрим условие поперечного фазового резонанса. Согласно этому условию, полное изменение фазы за один зигзаг траектории луча должно быть кратно 2π. Поскольку длина оптического пути за один такой цикл равна 4a*sin(φ), то, учитывая изменения фазы, происходящие при отражении от границ, получаем условие фазового резонанса в виде

$$4akn_1 \sin\phi - \delta_{12} - \delta_{13} = 2N\pi \quad (N = 0,\ 1,\ 2,\ \ldots), \tag{5}$$

где δ_{12} и δ_{13} - сдвиги фазы, происходящие на границах раздела 1-2 и 1-3 соответственно. Для луча, электрическое поле которого перпендикулярно плоскости падения (соответствующего TE-моде), справедливо равенство:

$$\delta_{12\perp} = 2\,\mathrm{arctg}\left[\frac{(\cos^2\phi - (n_2/n_1)^2)^{1/2}}{\sin\phi}\right] = 2\,\mathrm{arctg}\left[\left(\frac{\beta^2 - n_2^2k^2}{n_1^2k^2 - \beta^2}\right)^{1/2}\right], \tag{6}$$

где β и k введены с помощью соотношения (3). Аналогичное выражение имеет место для сдвига фазы на границе раздела 1-3 для волны с такой же поляризацией. Введем следующие обозначения:

$$\begin{aligned} p^2 &= \beta^2 - n_2^2k^2, \\ q^2 &= n_1^2k^2 - \beta^2, \\ r^2 &= \beta^2 - n_3^2k^2. \end{aligned} \tag{7}$$

Тогда уравнение (5) принимает вид

$$4aq - 2\,\mathrm{arctg}\,(p/q) - 2\,\mathrm{arctg}\,(r/q) = 2N\pi \quad (N = 0,\ 1,\ 2,\ \ldots), \tag{8}$$

и его можно переписать в более удобной форме:

$$\operatorname{tg}(2aq - N\pi) = \frac{(p+r)\,q}{q^2 - pr} \quad (N = 0,\ 1,\ 2,\ \ldots). \tag{9}$$

Уравнение (9) называется характеристическим уравнением для лучей в плоском трехслойном диэлектрическом волноводе. Аналогичным образом, используя введенные выше обозначения, можно представить выражение для лучей с противоположной поляризацией:

$$\delta_{12\parallel} = 2\operatorname{arctg}\left[(n_1/n_2)^2\, p/q\right]. \tag{10}$$

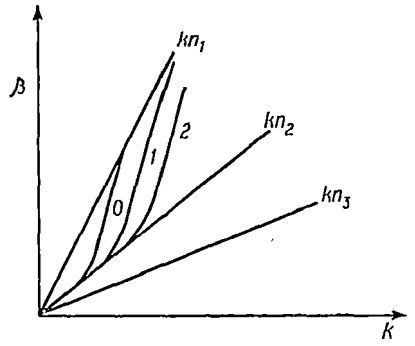


Рисунок 2 - Схематическая зависимость постоянной распространения β от волнового числа k для мод (или лучей) низшего порядка в плоском диэлектрическом волноводе.

Это выражение для сдвига фазы на границе 1-2 и аналогичное выражение для границы раздела 1-3 позволяют записать условие резонанса в виде:

$$\operatorname{tg}(2aq - N\pi) = \frac{\left(n_3^2 p + n_2^2 r\right) n_1^2 q}{n_2^2 n_3^2 q^2 - n_1^4 pr} \quad (N = 0,\ 1,\ 2,\ \ldots). \tag{11}$$

Уравнение (11) называется характеристическим уравнением для ТМ-лучей в плоском трехслойном диэлектрическом волноводе. Хотя для решения уравнений (9) и (11) можно использовать различные аналитические приближения решения

этих уравнений в наиболее общем виде можно получить только с помощью численных методов. На рисунке 2 схематически изображены решения этих уравнений в виде зависимости постоянной распространения β от к для нескольких мод волновода.

Согласно (3) область допустимых значений для β, соответствующих направляемым модам, ограничена прямыми $\beta=kn_1$ и $\beta=kn_1$. Различным значениям величины N в уравнениях (9) и (11) соответствуют различные зависимости β от k, характеризующие определенную направляемую моду.

Эффективная толщина волновода

Приведенный выше вывод характеристического уравнения впервые был проведен Тьеном и Ульрихом. Несколько отличные выводы этого уравнения, основанные на лучевом описании, даны в работах Маркузе, Лава и Снайдера. Понятно, что, зная решение характеристического уравнения, можно найти фазовую скорость моды, а продифференцировав его, - групповую скорость. Однако последнюю величину можно получить непосредственно из рассмотрения эффекта Гуса - Хэнхена на границах волновода. Чтобы продемонстрировать это, вспомним, что, при отражении луча от диэлектрической границы раздела существует определенная глубина проникновения луча. Физически это явление обусловлено наличием затухающего поля, проникающего в материал с более низким показателем преломления. Если мы учтем эту глубину проникновения и связанный с ней продольный сдвиг луча, то зигзагообразная модель распространения луча примет вид, изображенный на рисунке 3. Сдвиг луча на границе раздела 1-2 обозначен через z_{s12} и описывается выражением:

$$z_s = \frac{1}{2kn_1 \cos\theta_1} \frac{\partial\delta}{\partial\theta_1}, \qquad (12)$$

а глубина проникновения, обозначенная x_{s12}, задается выражением:

$$x_s = z_s \operatorname{ctg} \theta_1 . \quad , \qquad (13)$$

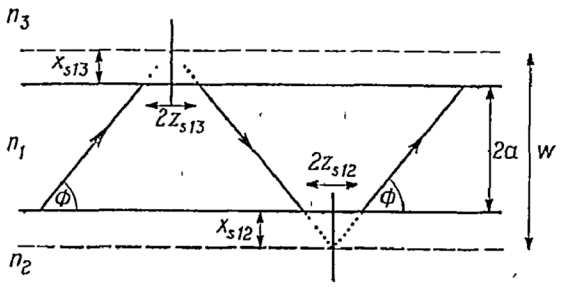


Рисунок 3 - Зигзагообразная траектория луча в диэлектрическом волноводе при учете влияния сдвигов Гуса-Хэнхена.

Аналогично можно рассмотреть поведение луча и на границе раздела 1-3. Следовательно, как видно из рисунка 3, мы можем считать, что луч распространяется по зигзагообразной траектории в волноводе, эффективная толщина которого ω определяется выражением

$$w = 2a + x_{s12} + x_{s13} . \qquad (14)$$

С помощью формул (12) и (13) для глубины проникновения луча, поляризованного так, что вектор электрического поля перпендикулярен плоскости xz, и с помощью введенных выше определений для величин p и r выражение (14) можно представить в виде:

$$w_{\perp} = 2a + 1/p + 1/r . \qquad (15)$$

Аналогично для волн, поляризованных в плоскости xz:

$$w_{\parallel} = 2a + \frac{1}{p\left(\frac{\beta^2}{n_1^2 k^2} + \frac{\beta^2}{n_2^2 k^2} - 1\right)} + \frac{1}{r\left(\frac{\beta^2}{n_1^2 k^2} + \frac{\beta^2}{n_3^2 k^2} - 1\right)} =$$

$$= 2a + \frac{n_1^2 n_2^2 (p^2 + q^2)}{p\left(p^2 n_1^4 + q^2 n_2^4\right)} + \frac{n_1^2 n_3^2 (r^2 + q^2)}{r\left(r^2 n_1^4 + q^2 n_3^4\right)} . \qquad (16)$$

Интуитивно понятно, что эффективная толщина ω характеризует область локализации электромагнитного поля в волноводе. Как мы увидим в дальнейшем, эта величина еще понадобится нам при анализе плоских трехслойных диэлектрических волноводов с точки зрения теории электромагнитных мод.

Групповая скорость

Выведем выражения для величины групповой скорости. Для этого важно понимать, что сдвиг Гуса-Хэнхена приводит также к временной задержке, связанной с временем, которое затрачивает луч на распространение в среде с меньшим показателем преломления. При рассмотрении временного аспекта эффекта Гуса-Хэнхена используем рассуждения, применяемые при исследовании пространственного волнового пакета. Так, время задержки $2\tau_s$ можно найти, рассматривая две плоские волны со слегка различающимися угловыми частотами $\omega \pm \Delta\omega$. Тогда величина τ_s имеет вид:

$$\tau_s = -\frac{1}{2}\partial\delta/\partial\omega = -\frac{1}{2c}\partial\delta/\partial k, \qquad (17)$$

где δ - сдвиг фазы при полном внутреннем отражении, а c - скорость света в свободном пространстве. Время, затрачиваемое лучом, пересекающим волновод толщиной 2a, равно:

$$\tau_a = \frac{2am_1}{c\sin\phi}, \qquad (18)$$

где m_1 - групповой показатель преломления в диэлектрической среде 1:

$$m_1 = n_1 - \lambda\frac{dn_1}{d\lambda} \equiv n_1 + \omega\frac{dn_1}{d\omega} \equiv n_1 + k\frac{dn_1}{dk}, \qquad (19)$$

и при вычислении времени задержки мы использовали групповую скорость в среде 1, а не соответствующую фазовую скорость c/n_1. Поэтому полное время τ_ω, затрачиваемое лучом на пересечение волновода с кажущейся толщиной ω, равно

$$\tau_w = \tau_a + \tau_{s12} + \tau_{s13}, \qquad (20)$$

где τ_{s12} и τ_{s13} - времена задержки при отражении от границ раздела 1-2 и 1-3 соответственно. За это время луч прошел вдоль оси волновода расстояние, равное ω*ctg(φ), где ширина ω определена формулой (14). Наконец, полная групповая скорость V_g при зигзагообразном распространении луча в волноводе определяется как отношение полного смещения вдоль оси ω*ctg(φ) к затраченному времени τ_ω:

$$V_g = \frac{2a \operatorname{ctg} \phi + z_{s12} + z_{s13}}{\tau_a + \tau_{s12} + \tau_{s13}}. \qquad (21)$$

Покажем теперь, что этот результат совпадает с результатом, полученным путем непосредственного дифференцирования характеристического уравнения для волновода. Так, используя это уравнение в его первоначальном виде (5) и дифференцируя его по β (предполагая, что n_1 является функцией частоты ω), имеем:

$$4a \sin \phi \left(n_1 + k \frac{dn_1}{dk}\right) \frac{\partial k}{\partial \beta} + 4akn_1 \cos \phi \frac{d\phi}{d\beta} - \frac{\partial}{\partial \beta} (\delta_{12} + \delta_{13}) - \\ - \frac{\partial k}{\partial \beta} \frac{\partial}{\partial k} (\delta_{12} + \delta_{13}) = 0, \qquad (22)$$

или, вводя в это уравнение частоту ω и групповой показатель преломления m_1, определенный формулой (19), получаем:

$$\frac{4am_1 \sin\phi}{c}\frac{\partial\omega}{\partial\beta} + 4akn_1\cos\phi\frac{d\phi}{d\beta} - \frac{\partial}{\partial\beta}(\delta_{12}+\delta_{13}) - \\ -\frac{\partial\omega}{\partial\beta}\frac{\partial}{\partial\omega}(\delta_{12}+\delta_{13}) = 0. \qquad (23)$$

Далее, дифференцируя выражение (3):

$$\frac{d\phi}{d\beta} = \frac{(m_1\cos\phi/c)\,\partial\omega/\partial\beta - 1}{n_1 k\sin\phi}, \qquad (24)$$

и исключая производную $d\varphi/d\beta$ из уравнения (23), после простых преобразований находим выражение для величины $d\omega/d\beta$ (или, что эквивалентно, для групповой скорости V_g):

$$V_g = \frac{\partial\omega}{\partial\beta} = \frac{2a\,\mathrm{ctg}\,\phi + \frac{1}{2}\frac{\partial\delta_{12}}{\partial\beta} + \frac{1}{2}\frac{\partial\delta_{13}}{\partial\beta}}{\frac{2am_1}{c\sin\phi} - \frac{1}{2}\frac{\partial\delta_{12}}{\partial\omega} - \frac{1}{2}\frac{\partial\delta_{13}}{\partial\omega}}. \qquad (25)$$

Анализ выражения (25) позволяет заключить, что второй и третий члены в числителе оказываются равными величинам z_{s12} и z_{s13} , в то время как слагаемые в знаменателе есть не что иное, как величины τ_a, τ_{s12} и τ_{s13}. Таким образом, мы доказали, что выражение (21) для групповой скорости, выведенное исключительно на основе физических соображений из рассмотрения распространения луча в зигзагообразной модели, совпадает с выражением (25), полученным путем дифференцирования характеристического уравнения для волновода.

Двумерная среда с параболическим показателем преломления

Градиент показателя преломления возникает в процессе изготовления волноводов, например в планарных волноводах при диффузии или ионной имплантации, и позволяет создавать волноводы с характеристиками, уникальными по виду получаемого изменения показателя преломления. Поэтому

в данном разделе будет рассмотрен наиболее простой вид симметричного изменения показателя преломления n(x) - распределение (профиль) показателя преломления, подчиняющееся параболическому закону:

$$n^2(x) = n_1^2\left[1 - 2\Delta\left(\frac{x}{a}\right)^2\right], \qquad (26)$$

где x - направление, перпендикулярное оси распространения z; n_1 - значение показателя преломления на оси волновода, а Δ и a - два параметра, задающие изменения показателя преломления; их значения будут определены ниже. Следует отметить, что выражение (26) фактически описывает изменение по параболическому закону величины $n^2(x)$, а не n(x), т. е. изменение диэлектрической проницаемости, а не показателя преломления. Тем не менее при не слишком больших значениях величины x, для которых выполняется условие $2\Delta(x/a)^2 << 1$, мы приходим к параболическому закону изменения также и для показателя преломления:

$$n(x) \approx n_1\left[1 - \Delta\left(\frac{x}{a}\right)^2\right]. \qquad (27)$$

Выражения (26) и (27) можно выбрать в качестве первых приближений при описании любого симметричного распределения показателя преломления, допускающего разложение в ряд Тейлора около оси волновода (x = 0). В качестве конкретного примера часто рассматривается случай градиентного распределения показателя преломления в сердцевине |x| < a, окруженной оболочкой с однородным распределением показателя преломления n_2. Рассматриваемая ситуация иллюстрируется на рисунке 4 для случая, когда показатель преломления не испытывает скачка на границе |x| = a, т. е. когда $n(a) = n^2$.

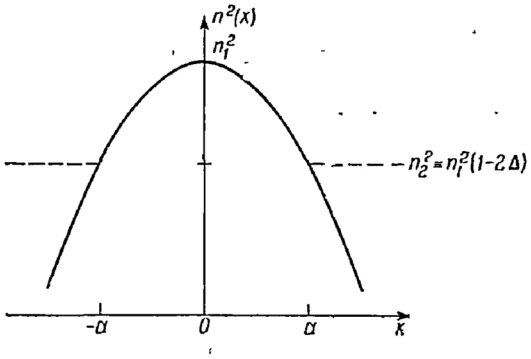


Рисунок 4 - Параболическое распределение диэлектрической проницаемости
Сплошная кривая соответствует неограниченному параболическому профилю; штриховая линия - ограниченному параболическому профилю.

Из этого рисунка становится понятным определение параметра Δ:

$$2\Delta = \frac{n_1^2 - n_2^2}{n_1^2}. \qquad (28)$$

В приближении слабонаправляющих волноводов, т. е. при выполнении условия n_1-n_2 << n_1, n_2, выражение (28) приводится к виду

$$\Delta \approx \frac{n_1 - n_2}{n_1}; \qquad (29)$$

величина Δ в этом приближении является относительным изменением показателя преломления при переходе от сердцевины к оболочке плоского волновода. Такие определения величин Δ и а оказываются полезными (хотя и несущественными) и при нахождении нормированной частоты υ, определяемой по аналогии со случаем плоского волновода следующим образом:

$$v^2 = a^2k^2\left(n_1^2 - n_2^2\right) = a^2k^2 2\Delta n_1^2, \qquad (30)$$

где $k = 2\pi/\lambda$. Выражение (30) облегчает сопоставление результатов для среды с параболическим распределением показателя преломления и результатов для плоских и других волноводов. Однако при использовании рассмотренного выше распределения показателя преломления следует соблюдать определенную осторожность: параболическое распределение (26) или (27) задает моды, характеристики которых имеют сходство с характеристиками мод для параболического распределения в сердцевине, ограниченной оболочкой (штриховая линия на рисунке 4), лишь в случае, когда поле компактно локализовано у центра сердцевины. По мере увеличения проникновения поля в слои оболочки приближение, задаваемое выражениями (26) и (27), будет постепенно становиться все хуже и хуже, и при отсечке это распределение поля уже не будет иметь никакого сходства с реальным распределением поля для ограниченного параболического профиля. Имея это в виду, мы будем рассматривать здесь только неограниченную среду с параболическим распределением показателя преломления, описываемую выражениями (26) и (27), сохраняя пока определения (28) и (29) для удобства сравнения с результатами, полученными для других профилей показателя преломления.

Описание среды с параболическим распределением показателя преломления в лучевом приближении

Начнем изложение с отчасти нового вывода уравнений распространения луча в среде с параболическим распределением показателя преломления; в основу этого вывода положен метод локально плоских волн. Решения уравнений задают траектории лучей и демонстрируют фокусирующее действие рассматриваемой среды; они позволяют определить время распространения вдоль траектории луча. При выводе характеристического уравнения для допустимых значений постоянной распространения привлекается условие фазового резонанса; в результате возникает необходимость рассмотреть изменение фазы, происходящее на каустике. Наконец, будет показано, что время распространения, найденное

непосредственно из характеристического уравнения, оказывается совпадающим со временем распространения, рассчитанным с помощью траекторий лучей.

Вывод лучевых уравнений с помощью метода локально-плоских волн

Лучевые уравнения можно вывести из скалярного волнового уравнения, решая его на поверхностях с постоянными фазами и используя тот факт, что все лучи перпендикулярны этим поверхностям. Их можно также вывести из принципа Ферма или из уравнении Гамильтона по аналогии с классической механикой. Поскольку нас в основном интересуют лучевые уравнения в среде с параболическим распределением показателя преломления, мы приведем менее формальный вывод этих уравнений, основанный на разложении локального волнового вектора. Такой подход, уже упоминавшийся нами кратко, хотя и уступает в строгости другим методам, обладает определенными преимуществами, поскольку допускает простую физическую интерпретацию и позволяет выразить интегралы движения через уже известные волновые параметры.

На рисунке 5, а представлено разложение локального волнового вектора kn(x) с началом в точке x, z в двумерной декартовой системе координат. Мы предполагаем здесь (поскольку используется лучевое описание), что все линейные размеры задачи существенно превышают длину волны λ и, в частности, показатель преломления n(x) изменяется достаточно медленно в направлении x, т. е. в масштабе длины волны λ его можно считать постоянным. Следовательно, компонентами z и x локального волнового вектора являются величины β и q соответственно, причем:

$$q^2 = k^2 n^2(x) - \beta^2. \qquad (31)$$

Локальное направление луча, исходящего из точки x, z, задается вектором kn(x). Элемент длины вдоль траектории луча определяется следующим образом:

$$ds^2 = dx^2 + dz^2, \qquad (32)$$

а его разложение по направлениям x и z представлено на рисунке 5, б. Сравнивая рисунки 5, а и б, нетрудно получить уравнения для лучей, распространяющихся в двумерной среде с показателем преломления n(x):

$$\frac{ds}{dx} = \frac{kn(x)}{[k^2n^2(x) - \beta^2]^{1/2}},$$
$$\frac{dz}{dx} = \frac{\beta}{[k^2n^2(x) - \beta^2]^{1/2}}, \quad (33)$$

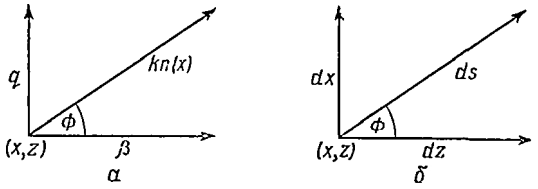


Рисунок 5 - разложение локального волнового вектора на компоненты в плоскости распространения xz (а), разложение элемента длины луча па компоненты в плоскости распространения xz (б).

Заметим, что интегралом движения в этих уравнениях является величина β, которая совпадает с обычной продольной постоянной распространения, определенной выше.

Как видно из уравнений (33), существуют некоторые значения переменной x, при которых dx = q = 0, т. е. β = kn(x). Эти значения отвечают точкам поворота, или каустикам, между которыми параметр q принимает действительные значения, соответствующие действительному лучу; для значений x, лежащих вне этой области, величина q мнимая (что отвечает затухающей волне). Следовательно, траектория луча заключена между значениями x, для которых выполняется условие β = kn(x). С помощью уравнений (33) можно также определить время x распространения луча вдоль отрезка такой среды длиной L:

$$\tau = \frac{L}{c} \frac{\oint n(x)\, ds}{\oint dz}, \quad (34)$$

где c - скорость света в вакууме, а интеграл $\oint$ означает, что интегрирование ведется по полному периоду луча (величина которого предполагается меньшей L). Величина n(x)ds в выражении (34) задает элемент длины оптического пути вдоль луча, а не длину геометрического пути ds. Здесь мы, безусловно, пренебрегаем эффектами, связанными с материальной дисперсией, поскольку при более строгом описании в качестве элемента длины оптического пути следует использовать выражение mds, где m - групповой показатель преломления (n - $\lambda dn/d\lambda$).

Траектория луча и время его распространения

Чтобы найти траекторию луча в системе координат x, z, подставим в уравнение (33) выражение (26) для распределения показателя преломления. Тогда это дифференциальное уравнение примет вид:

$$\frac{dz}{dx} = \frac{\beta}{\left[k^2 n_1^2 - k^2 n_1^2 2\Delta\,(x/a)^2 - \beta^2\right]^{1/2}} = \frac{\alpha\beta}{\left[u^2 - v^2\,(x/a)^2\right]^{1/2}}, \qquad (35)$$

где величина υ^2 определяется формулой (30).

Если граничные условия выбраны в виде $x = x_0$ при $z = 0$, то (интегрирование (35) приводит к выражению:

$$\frac{zv}{a^2\beta} = \arcsin\left(\frac{vx}{ua}\right) - \arcsin\left(\frac{vx_0}{ua}\right), \qquad (36)$$

из которого видно, что траектория луча имеет периодический характер, причем ее период равен $2\pi a^2\beta/\upsilon$, а максимальная амплитуда осцилляции равна ua/υ. Чтобы представить (36) в более удобном виде, введем величину $\mathrm{tg}(\varphi_0)$, определяющую начальный наклон луча dx/dz при $x = x_0$, $z = 0$:

$$a\beta \operatorname{tg} \phi_0 = \left[u^2 - v^2 \left(\frac{x_0}{a}\right)^2\right]^{1/2}. \qquad (37)$$

Тогда выражение (36) после преобразования примет вид

$$x = x_0 \cos\left(\frac{zv}{a^2\beta}\right) + \frac{a^2\beta}{v} \operatorname{tg} \phi_0 \sin\left(\frac{zv}{a^2\beta}\right); \qquad (38)$$

оно задает величину смещения x как функцию расстояния вдоль оси z для луча с начальной координатой x_0 и начальным наклоном $\operatorname{tg}(\varphi_0)$.

Поскольку нас в основном интересуют лучи, которые не сильно отклоняются от оси (чтобы среда с параболическим распределением показателя преломления могла служить моделью при описании слабонаправляющей параболической среды, ограниченной оболочкой), то практически мы будем иметь дело с параксиальным приближением (угол луча φ на рисунке 5 должен быть малым). В этом случае выполняется условие $\beta \approx kn_1$ и выражение (38) окончательно принимает вид:

$$x \approx x_0 \cos\left((2\Delta)^{1/2} \frac{z}{a}\right) + \frac{a\phi_0}{\sqrt{2\Delta}} \sin\left((2\Delta)^{1/2} \frac{z}{a}\right). \qquad (39)$$

Заметим, что из выражения (39) следует, что в параксиальном приближении все лучи имеют одинаковый период осцилляции. На рисунке 6 приведены некоторые траектории лучей, вычисленные с помощью данного выражения.

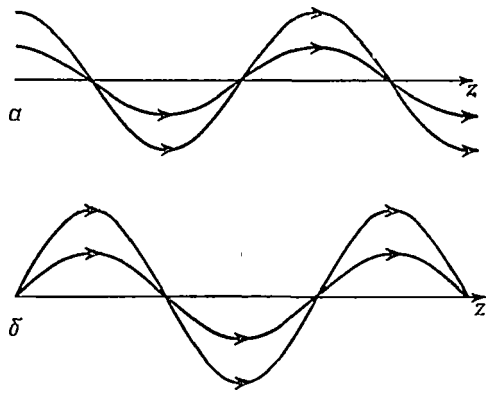


Рисунок 6 - Траектории лучей в среде с параболическим распределением показателя преломления, полученные с помощью выражения (39): а) лучн, параллельные на входе φ_0; б) лучи, пересекающие на входе ось среды ($x_0 = 0$).

Периодический характер траекторий лучей приводит к возникновению эффекта фокусировки в среде с параболическим распределением показателя преломления; в такой среде на расстояниях, кратных периоду осцилляции луча $2\pi a/(2\Delta)^{1/2}$ будут формироваться изображения. Следовательно, участок такой среды с определенной длиной L будет работать как линза. Это свойство среды иллюстрируется на рисунке 7, на котором изображена траектория луча, входящего в среду параллельно оси z ($\varphi_0 = 0$). После прохождения участка среды длиной L этот луч выходит из среды с поперечным смещением x, которое выражается следующей формулой:

$$x = x_0 \cos\left((2\Delta)^{1/2}\frac{L}{a}\right). \qquad (40)$$

Если среда вне области, занятой средой с параболическим распределением показателя преломления, имеет показатель преломления n_0, то угол выхода луча (рисунок 7) в параксиальном приближении выражается формулой:

$$\rho \approx -\frac{n_1 (2\Delta)^{1/2} x_0}{n_0 a}\sin\left((2\Delta)^{1/2}\frac{L}{a}\right). \qquad (41)$$

Фокусное расстояние f участка среды длиной L с параболическим распределением показателя преломления определяется как расстояние между точкой z = L и точкой пересечения луча с осью z:

$$f \approx \left|\frac{x}{\rho}\right| \approx \frac{n_0 a}{n_1 (2\Delta)^{1/2}} \operatorname{ctg}\left((2\Delta)^{1/2} \frac{L}{a}\right). \qquad (42)$$

Следует отметить, что в параксиальном приближении выражение (42) для фокусного расстояния не зависит от начального смещения x_0 входящего луча, и поэтому среда с параболическим распределением показателя преломления для всех лучей с небольшим смещением относительно оси работает как настоящая линза.



Рисунок 7 - Участок среды длиной L с параболическим распределением показателя преломления, действующий как линза с фокусным расстоянием f.

Время распространения луча в среде с параболическим профилем показателя преломления можно найти с помощью выражения (34). Знаменатель этого выражения вычисляется (в общем случае без использования параксиального приближения) с помощью интегрирования по полупериоду луча от x_0=-ua/υ до x=ua/υ:

$$\oint dz = \frac{a^2 \beta \pi}{v}. \qquad (43)$$

Аналогично числитель в (34) можно найти с помощью непосредственного интегрирования по полупериоду луча:

$$\oint n(x)\,ds = \int_{-ua/v}^{ua/v} \frac{kn^2(x)\,dx}{[k^2n^2(x) - \beta^2]^{1/2}} =$$
$$= \int_{-ua/v}^{ua/v} \frac{akn_1^2[1 - 2\Delta(x/a)^2]\,dx}{[u^2 - v^2(x/a)^2]^{1/2}} = \frac{a^2kn_1^2\pi}{v}\left(1 - \frac{\Delta u^2}{v^2}\right). \qquad (44)$$

Следовательно, время распространения луча принимает вид:

$$\tau = \frac{L}{c}\left(\frac{kn_1^2}{\beta}\right)\left(1 - \frac{\Delta u^2}{v^2}\right). \qquad (45)$$

Заметим, что в рамках параксиального приближения $\beta \approx kn_1$, и выражение (45) дает для времени распространения луча τ одинаковую для всех лучей величину Ln/c. Физический смысл этого результата можно понять из рассмотрения траекторий лучей на рисунке 6. Как видно из этого рисунка, геометрические длины траекторий различны для разных лучей, причем лучам с большим смещением относительно оси соответствует большая длина пути. Однако большую часть пути такие лучи распространяются в среде с меньшим показателем преломления; следовательно, они распространяются быстрее других лучей, не имеющих больших поперечных смещений. Таким образом, при соответствующем выборе профиля показателя преломления полное время распространения всех лучей может быть одинаковым. Как мы доказали, параболический профиль в параксиальном приближении обладает именно таким свойством.

Сдвиг фазы на каустике

Чтобы вывести на основе лучевого описания характеристическое уравнение для среды с параболическим распределением показателя преломления, нужно воспользоваться условием поперечного фазового резонанса, уже рассмотренным нами для волноводов в гл. 1 и 2. Однако в данном случае ситуация оказывается несколько более сложной, поскольку необходимо учесть фазовые сдвиги на каустиках $x = \pm au/v$, определяющих границы отклонения луча от оси z. Как легко убедиться с помощью уравнения (33), для любого профиля показателя

преломления n(x) все лучи с одинаковыми β касательны к каустикам. Чтобы вычислить величину изменения фазы на каустике, можно воспользоваться общими выражениями для фазового сдвига на границе раздела диэлектрика, аппроксимируя непрерывную кривую n(x) ступенчатой кривой, а затем переходя к пределу бесконечно малых ступенек показателя преломления и касательного падения лучей. Рассмотрим луч, распространяющийся в среде с показателем преломления $n(x_c-\Delta x)$, непосредственно до достижения им любой каустики (положение каустики определяется координатой x_c, а величина Δx предполагается малой); значение показателя преломления с противоположной стороны каустики будет $n(x_c+\Delta x)$. Тогда выражение для сдвигов фазы электрического поля световой волны, перпендикулярного и параллельного плоскости падения, находятся следующим образом:

$$\delta_{\perp} = 2\,\mathrm{arctg}\left\{\left[\frac{(\beta/k)^2 - n^2(x_c+\Delta x)}{n^2(x_c-\Delta x) - (\beta/k)^2}\right]^{1/2}\right\},$$
$$\delta_{\parallel} = 2\,\mathrm{arctg}\left\{\frac{n^2(x_c-\Delta x)}{n^2(x_c+\Delta x)}\left[\frac{(\beta/k)^2 - n^2(x_c+\Delta x)}{n^2(x_c-\Delta x) - (\beta/k)^2}\right]^{1/2}\right\}. \qquad (46)$$

где $\beta = n(x)*k*\sin(\Theta_1)$.

В пределе $\Delta x \to 0$, который требуется вычислить для перехода к касательному падению лучей и непрерывному распределению n(x), как числитель, так и знаменатель функции arctg в (46) стремятся к нулю. Для вычисления этого предела можно разложить функции в ряд Тейлора в окрестности точки x_c, либо использовать правило Лопиталя. В результате, можно показать, что сдвиг фазы на каустике для обеих ортогонально поляризованных волн стремится к $\pi/2$.

Лекция 9

Электромагнитная теория планарных волноводов. Волновые уравнения планарных волноводов. Моды тонкопленочного волновода. Свойства мод тонкопленочного волновода. Волновые уравнения для градиентных планарных волноводов.

Рассмотрим направляемые моды плоского диэлектрического волновода, определяемые непосредственно из уравнений Максвелла. Покажем также, что характеристическое уравнение, полученное при таком подходе, совпадает с соответствующим уравнением, выведенным ранее на основе лучевого описания. Для этого вернемся к рассмотрению асимметричного плоского диэлектрического волновода без потерь. Записывая систему уравнений Максвелла для каждого из трех слоев волновода с показателем преломления n_j , j=1,2,3 и предполагая, что величина магнитной проницаемости везде равна магнитной проницаемости свободного пространства μ_0, имеем

$$\nabla \times H = n_j^2 \varepsilon_0 dE/dt, \quad (1)$$

$$\nabla \times E = -\mu_0 dH/dt, \quad (2)$$

$$\nabla \cdot E = 0, \quad (3)$$

$$\nabla \cdot H = 0, \quad (4)$$

Применяя оператор дивергенции к уравнению (2) и исключая из уравнения величину H с помощью уравнения (1), получаем

$$\nabla \times (\nabla \times E) = -\mu_0 \nabla \times (dH/dt) = -\mu_0 n_j^2 \varepsilon_0 d^2 E/dt^2, \quad (5)$$

Используя известное тождество

$$\nabla \times (\nabla \times A) = \nabla(\nabla \cdot A) - \nabla^2 A,$$

справедливое для любого вектора A, и уравнение (3), находим для уравнения (5) следующее выражение:

$$\nabla^2 E = \mu_0 \varepsilon_0 n_j^2 d^2 E / dt^2$$

Наконец, предполагая, что $E \sim \exp(-i\omega t)$, получаем известное волновое уравнение для однородной диэлектрической среды с показателем преломления n_j:

$$\nabla^2 E + k n_j^2 E = 0, \qquad (6)$$

где $k^2 = \omega^2 \mu_0 \varepsilon_0$. Уравнение (6) можно упростить, если учесть, что в рассматриваемом двумерном случае все производные по y обращаются в нуль. Это следует из предположения, что в положительном и отрицательном направлениях оси y волновод неограничен, т. е. распределение полей мод в направлении оси y однородно. Далее, если предположить, что зависимость поля от координаты z задается множителем exp(iβz), где β - продольная постоянная распространения, то уравнение (6) можно записать для каждой из трех областей волновода следующим образом:

Область 1: $\frac{\partial^2 E_3}{dx^2} - r^2 E_3 = 0, \qquad (7а)$

Область 2: $\frac{\partial^2 E_1}{dx^2} + q^2 E_1 = 0, \qquad (7б)$

Область 3: $\frac{\partial^2 E_2}{dx^2} - p^2 E_2 = 0, \qquad (7в)$

здесь использованы определения:

$$q^2 = n_1^2 k \ - \beta^2, \quad p^2 = \beta^2 - n_2^2 k^2, \quad r^2 = \beta^2 - n_3^2 k^2$$

Из уравнений Максвелла (1) - (4) можно вывести также аналогичные уравнения для магнитного поля H в соответствующих трех областях.

Поперечные электрические (TE) направляемые моды

Из сделанного выше предположения $\partial / \partial y = 0$ следует (см. уравнения (1) - (4)), что отличны от нуля только компоненты E_y, H_x и H_z поля TE-мод Опуская для удобства множитель exp(iβz-iωt), задающий зависимость поля от времени и координаты z, из уравнения (2) получаем:

$$H_x = -\frac{\beta}{\omega \mu_0} E_y, \qquad (8)$$

$$H_z = -\frac{i}{\omega \mu_0} \frac{\partial E_y}{\partial x}. \qquad (9)$$

Соотношения (8) и (9) связывают две отличные от нуля компоненты магнитного поля с единственной отличной от нуля компонентой электрического поля Ey, которая находится из решения волновых уравнений (7а) - (7в) для соответствующей области волновода. Другое требование, которому должны удовлетворять эти компоненты поля, заключается в непрерывности тангенциальных компонент E_y и H_z на границах раздела 1-2 и 1-3 между соответствующими слоями диэлектрика. Выберем начало оси x на границе раздела 1-3; тогда граница раздела 1-2 будет иметь координату x = -2a.

Энергия направляемых мод в основном должна быть сосредоточена в центральном слое волновода, т. е. в области 1. Как видно из уравнений (7а) - (7в), этому требованию будет удовлетворять решение, осциллирующее в области 1 (q^2>=0) и обладающее затухающими хвостами в области обкладок 2 и 3 (p^2,

q^2>=0). Учитывая такие условия, для величины постоянной распространения p получаем неравенства, аналогичные полученным ранее:

$$n_1 k \geq \beta \geq n_2 k \geq n_3 k \ .$$

Принимая во внимание вышеизложенное, можно сразу записать решение для компоненты E_y направляемой моды в трех областях волновода (опуская зависимость от времени t и координаты):

$$E_y = \begin{cases} Ae^{-rx}, & x \geq 0, \\ A\cos(qx) + B\sin(qx), & 0 \geq x \geq -2a, \\ (A\cos(2aq) - B\sin(2aq))e^{p(x+2a)}, & -2a \geq x. \end{cases} \qquad \text{(10а) – (10в)}$$

Решения (10а) - (10в) выбраны в таком виде, чтобы для величины E_y выполнялось условие непрерывности при x = 0 и x = -2a. Наконец, необходимо обеспечить непрерывность величины H_z. Выражение для этой компоненты поля находится из (9) и имеет вид

$$H_z = -\frac{i}{\omega\mu_0}\begin{cases} -rAe^{-rx}, & x \geq 0, \\ q(-A\sin(qx) + B\cos(qx)), & 0 \geq x \geq -2a, \\ p(A\cos(2aq) - B\sin(2aq))e^{p(x+2a)}, & -2a \geq x. \end{cases} \qquad \text{(11а) – (11в)}$$

Условие непрерывности приводит к двум уравнениям

при x=0: $-rA = qB,$

при x=-2a: $q(A\sin(2aq) + B\cos(2aq)) = p(A\cos(2aq) - B\sin(2aq)).$

Исключая из этих уравнений отношение A/B, получаем характеристическое уравнение для TE-мод:

$$tg(2aq)=\frac{q(p+r)}{q^2-pr}. \qquad (12)$$

Это уравнение совпадает с уравнением ранее, полученным с помощью лучевого подхода.

Поперечные магнитные (TM) направляемые моды

Для такой поляризации отличны от нуля только компоненты H_y, E_x и E_z поля TM-мод. Снова опуская множитель exp(iβz-iωt), из уравнения (11) находим:

$$E_x=\frac{\beta}{\omega n_j^2\varepsilon_0}H_y \quad (j=1,2,3), \qquad (13)$$

$$E_z=\frac{\beta}{\omega n_j^2\varepsilon_0}\frac{\partial H_y}{\partial x} \quad (j=1,2,3), \qquad (14)$$

Эти уравнения связывают компоненты электрического поля E_x и E_z с единственной отличной от нуля компонентой магнитного поля H_v, которая является решением волновых уравнений для трех областей волновода. Аналогично можно получить и решение для компоненты H_y:

$$H_y=\begin{cases} Ce^{-rx}, & x\geq 0, \\ C\cos(qx)+D\sin(qx), & 0\geq x\geq -2a, \\ (C\cos(2aq)-D\sin(2aq))e^{p(x+2a)}, & -2a\geq x. \end{cases} \qquad (15а)-(15в)$$

Тогда, согласно (14), находим выражение для E_z в виде:

$$E_z = \frac{i}{\omega\varepsilon_0}\begin{cases}\left(-rC/n_3^2\right)e^{-rx}, & x \geq 0,\\ \left(q/n_1^2\right)\left(-C\sin(qx)+D\cos(qx)\right), & 0 \geq x \geq -2a,\\ \left(p/n_2^2\right)\left(C\cos(2aq)-D\sin(2aq)\right)e^{p(x+2a)}, & -2a \geq x.\end{cases} \qquad (16а) - (16в)$$

Требование непрерывности компоненты Ez при x = 0 и x = —2a дает

$$-rC/n_3^2 = qD/n_1^2,$$
$$\left(q/n_1^2\right)\left(C\sin(2aq)+D\cos(2aq)\right) = \left(p/n_2^2\right)\left(C\cos(2aq)-D\sin(2aq)\right).$$

Наконец, исключая из этих уравнений отношение C/D, получаем характеристическое уравнение для TM-мод:

$$tg(2aq) = \left(n_3^2 p + n_2^2 r\right)n_1^2 q/\left(n_2^2 n_3^2 q^2 - n_1^4 pr\right), \qquad (17)$$

которое совпадает с уравнением, найденным с помощью лучевого подхода.

Модовые числа и условие отсечки

Для обозначения моды, в распределении поля которой имеется N узлов, будем использовать обозначение TE_N или аналогичное обозначение TM_N. Чтобы найти величину N, представим аргумент тангенса в характеристическом уравнении (12) или (17) в виде 2aq-Nπ, как это имело место в случае уравнений (17) или (18). Условие отсечки задается выражением:

$$\beta = лm_2,$$

которое на языке лучевого описания соответствует нарушению условия полного внутреннего отражения, а на языке модового описания означает нарушение локализации поля излучения и расплывание этого поля в область 2.

С помощью двух приведенных выше определений мы можем найти выражения, задающие частоты отсечки для направляемых ТЕ- и ТМ-мод в плоском волноводе. Для случая ТЕ-мод, подставляя в уравнение (12) значения p, q и r при отсечке, получаем:

$$tg\left(2ak_c\left(n_1^2-n_2^2\right)^{1/2}-N\pi\right)=\left(\frac{n_2^2-n_3^2}{n_1^2-n_2^2}\right)^{1/2},$$

где k_c - значение волнового числа, соответствующее отсечке. Далее, если мы введем нормированную частоту υ:

$$\upsilon=ak\left(n_1^2-n_2^2\right)^{1/2},$$

то из приведенного выше выражения нетрудно найти значение υ_c при отсечке:

ТЕ-мода: $$\upsilon_c=\frac{1}{2}arctg\left(\left(\frac{n_2^2-n_3^2}{n_1^2-n_2^2}\right)^{1/2}\right)+N\pi/2, \qquad (19)$$

где функция arctg определена в интервале $0 - \pi/2$. Это соотношение можно также использовать для подсчета количества направляемых ТЕ-мод. Если предположить, что первая мода обозначена индексом N = 0, то из общего выражения (19) для нормированной частоты υ общее число направляемых мод волновода M находится следующим образом:

ТЕ-мода: $$M=\left\{\frac{1}{\pi}\left(2\upsilon-arctg\left(\left(\frac{n_2^2-n_3^2}{n_1^2-n_2^2}\right)^{1/2}\right)\right)\right\}_{\text{int}}, \qquad (20)$$

где индекс «int» указывает на то, что следует брать минимальное целое число, превышающее значение выражения в скобках. Нетрудно получить соответствующие результаты и для ТМ-мод. Так, для частоты отсечки, согласно (17), имеем:

ТМ-мода: $$\upsilon_c = \frac{1}{2} arctg\left((n_1/n_3)^2 \left(\frac{n_2^2 - n_3^2}{n_1^2 - n_3^2} \right)^{1/2} \right) + N\pi/2, \qquad (21)$$

а для числа M направляемых мод волновода находим:

ТМ-мода: $$M = \left\{ \frac{1}{\pi} \left(2\upsilon - arctg\left((n_1/n_3)^2 \left(\frac{n_2^2 - n_3^2}{n_1^2 - n_2^2} \right)^{1/2} \right) \right) \right\}_{\text{int}}, \qquad (22)$$

Нормировка на величину потока энергии

Усредненное по времени значение потока энергии P в волноводе равно интегралу по поперечному сечению волновода от компоненты вектора Пойтинга S_Z в направлении оси z:

$$P = \int_{-\infty}^{\infty} S_z \, dx = \frac{1}{2} \int_{-\infty}^{\infty} \mathrm{Re}\,(\mathbf{E} \times \mathbf{H}^*)_z \, dx. \qquad (23)$$

Для ТЕ-мод компонента S_Z задается выражением

$$S_z = -\frac{1}{2} E_y H_x^* = \frac{\beta}{2\omega\mu_0} |E_y|^2, \qquad (24)$$

где использовано уравнение, комплексно-сопряженное (18). Интеграл в выражении (23) следует разложить на три слагаемых, соответствующих трем

слоям волновода. Обозначая мощность, распространяющуюся в каждой области, величиной P_j (j=1, 2, 3) и используя (20а) - (20в), для ТЕ-мод находим:

$$P_3 = \left(\frac{\beta}{2\omega\mu_0}\right) A^2 / 2r, \quad (25а)$$

$$P_1 = \left(\frac{\beta}{2\omega\mu_0}\right) \frac{A^2}{2} \left(\frac{q^2 + r^2}{q^2}\right) \left(2a + \frac{p}{q^2 + p^2} + \frac{r}{q^2 + r^2}\right), \quad (25б)$$

$$P_2 = \left(\frac{\beta}{2\omega\mu_0}\right) \frac{A^2}{2p} \left(\frac{q^2 + r^2}{p^2 + q^2}\right). \quad (25в)$$

Для получения выражений (256) и (25в) в таком простом виде были использованы граничные условия, связывающие величины А и В, и характеристическое уравнение (22). Таким образом, полная мощность Р, распространяющаяся в волноводе, задается выражением:

$$P = P_1 + P_2 + P_3 = \left(\frac{\beta}{2\omega\mu_0}\right) \frac{A^2}{2} \left(\frac{q^2 + r^2}{q^2}\right) \left[2a + \frac{1}{p} + \frac{1}{r}\right]. \quad (26)$$

В последнем множителе выражения (26) мы узнаем величину $\omega_\perp$, введенную при лучевом описании и названную «эффективной толщиной волновода». На языке приведенного здесь модового описания эта величина, как видно из выражений (20а) - (20в), представляет собой расстояние между точками в областях 2 и 3, в которых интенсивность поля падает в е раз по отношению к значению интенсивности на соответствующих границах раздела, т. е. величина $\omega_\perp$ действительно оказывается мерой проникновения поля в слои обкладок.

Выражение (26) можно использовать для нормировки полей мод (20а) - (20в) на величину потока энергии Р:

$$A^2 \equiv \left(\frac{qB}{r}\right)^2 = \frac{P}{\left(\frac{\beta}{2\omega\mu_0}\right)\left(\frac{q^2 + r^2}{2q^2}\right)\left(2a + \frac{1}{p} + \frac{1}{r}\right)}. \quad (27)$$

Кроме того, выражения (25) и (26) можно использовать для получения относительно простых выражений, задающих отношение мощности в сердцевине P_{core} к величине полной мощности P и отношение мощности излучения в области обкладок P_{clad} к полной мощности P:

$$P_{core}/P = \frac{2a + \dfrac{p}{q^2+p^2} + \dfrac{r}{q^2+r^2}}{2a + 1/p + 1/r}, \quad (28)$$

$$P_{clad}/P = \frac{\dfrac{q^2}{p\left(q^2+p^2\right)} + \dfrac{q^2}{r\left(q^2+r^2\right)}}{2a + 1/p + 1/r}. \quad (29)$$

Из найденных выражений нетрудно видеть, что величины отношений P_j/P (/j = 1,2,3) совпадают с коэффициентами поглощения a, для луча в слабонаправляющем волноводе. Иными словами, выражение можно переписать в виде

$$\alpha = (P_1\alpha_1 + P_2\alpha_2 + P_3\alpha_3)/P.$$

Этот результат можно также получить непосредственно из характеристического уравнения с помощью теории возмущений.

Для TM-мод составляющая вектора Пойтинга вдоль направления z имеет вид:

$$S_z = \frac{1}{2} H_y^* E_x = \frac{\beta}{2\omega n_j^2 \varepsilon_0} |H_y|^2 \quad (j = 1,\ 2,\ 3), \quad (30)$$

где было использовано выражение (13). Мощность, заключенная в каждой области P_j, в этом случае находится с помощью выражений (15а) - (15в):

$$P_3 = \left(\frac{\beta}{2\omega n_3^2 \varepsilon_0}\right)\frac{C^2}{2r}, \quad (31а)$$

$$P_1 = \left(\frac{\beta}{2\omega n_1^2 \varepsilon_0}\right)\frac{C^2}{2}\left(\frac{q^2 n_3^4 + r^2 n_1^4}{q^2 n_3^4}\right)\left(2a + \frac{p n_1^2 n_2^2}{q^2 n_2^4 + p^2 n_1^4} + \frac{r n_1^2 n_3^2}{q^2 n_3^4 + r^2 n_1^4}\right), \quad (31б)$$

$$P_2 = \left(\frac{\beta}{2\omega n_2^2 \varepsilon_0}\right)\frac{C^2}{2p}\left(\frac{n_2^4}{n_3^4}\right)\left(\frac{q^2 n_3^4 + r^2 n_1^4}{q^2 n_2^4 + p^2 n_1^4}\right). \quad (31в)$$

здесь были использованы характеристическое уравнение (17) и значение величины C/D. Таким образом, выражение для полной мощности P принимает вид

$$P = P_1 + P_2 + P_3 = \left(\frac{\beta}{2\omega\varepsilon_0}\right)\frac{C^2}{2}\left(\frac{q^2 n_3^4 + r^2 n_1^4}{q^2 n_1^2 n_3^4}\right)\times$$
$$\times\left[2a + \frac{n_1^2 n_2^2 (q^2 + p^2)}{p\left(q^2 n_2^4 + p^2 n_1^4\right)} + \frac{n_1^2 n_3^2 (q^2 + r^2)}{r\left(q^2 n_3^4 + r^2 n_1^4\right)}\right]. \quad (32)$$

Последний сомножитель в выражении (2.102), так же как и в случае TE-мод, оказывается эффективной шириной $\omega_{\|}$ для TM-мод, определенной ранее. Выражение (32) можно использовать для нормировки полей мод (15а) - (15в) на поток энергии P:

$$C^2 \equiv (q n_3^2 D / r n_1^2)^2 =$$
$$= \frac{P}{\left(\frac{\beta}{2\omega\varepsilon_0}\right)\left(\frac{q^2 n_3^4 + r^2 n_1^4}{2q^2 n_1^2 n_3^4}\right)\left[2a + \frac{n_1^2 n_2^2 (p^2 + q^2)}{p\left(q^2 n_2^4 + p^2 n_1^4\right)} + \frac{n_1^2 n_3^2 (q^2 + r^2)}{r\left(q^2 n_3^4 + r^2 n_1^4\right)}\right]}. \quad (33)$$

Соответствующие отношения мощности в сердцевине мощности в обкладках Pciad к полной мощности P имеют вид:

$$P_{core} / P = \frac{2a + \frac{p n_1^2 n_2^2}{q^2 n_2^4 + p^2 n_1^4} + \frac{r n_1^2 n_3^2}{q^2 n_3^4 + r^2 n_1^4}}{2a + \frac{n_1^2 n_2^2\left(q^2 + p^2\right)}{p\left(q^2 n_2^4 + p^2 n_1^4\right)} + \frac{n_1^2 n_3^2\left(q^2 + r^2\right)}{r\left(q^2 n_3^4 + r^2 n_1^4\right)}}, \quad (34)$$

$$P_{clad}/P = \frac{\dfrac{q^2 n_1^2 n_2^2}{q^2 n_2^4 + p^2 n_1^4} + \dfrac{q n_1^2 n_3^2}{q^2 n_3^4 + r^2 n_1^4}}{2a + \dfrac{n_1^2 n_2^2 \left(q^2 + p^2\right)}{p\left(q^2 n_2^4 + p^2 n_1^4\right)} + \dfrac{n_1^2 n_3^2 \left(q^2 + r^2\right)}{r\left(q^2 n_3^4 + r^2 n_1^4\right)}}. \qquad (35)$$

В заключение следует отметить, что выражения (26) и (32) для полного потока энергии P можно представить в чрезвычайно простом виде, используя эффективную ширину волновода ω и максимальное значение в распределении поля. Нетрудно показать, что в случае TE-мод компонента поля E_y как функция координаты x достигает своего максимального значения E_{ym} при:

$$x = (1/q)\,\mathrm{arctg}\,(-r/q),$$

и это максимальное значение равно

$$\frac{E_{ym}}{A} = \frac{(q^2 + r^2)^{1/2}}{q}. \qquad (36)$$

Следовательно, полный поток энергии для TE-мод, можно представить в простом виде:

$$P = (\beta/(4\omega\mu_0))\, E_{ym}^2 w_\perp \equiv (1/4)\, E_{ym} H_{xm} w_\perp, \qquad (37)$$

где было использовано выражение (8), а H_{xm} — максимальное значение компоненты поля H_x.

Аналогично в случае TM-мод максимальное значение H_{ym} компоненты поля H_y достигается при

$$x = (1/q)\,\mathrm{arctg}\,[-r n_1^2/(q n_3^2)]$$

и равно

$$H_{ym}/C = (q^2 n_3^4 + r^2 n_1^4)^{1/4}/(q n_3^2). \qquad (38)$$

Следовательно, полный поток энергии P для TM-мод задается простым выражением:

$$P = \left(\beta/\left(4\omega\varepsilon_0 n_1^2\right)\right) H_{ym}^2 w_{\parallel} \equiv (1/4) E_{xm} H_{ym} w_{\parallel}; \qquad (39)$$

здесь использовано выражение (13), а E_{xm} — максимальное значение компоненты поля E_x.

Относительная простота выражений (37) и (39) для потоков энергии и их очевидная физическая интерпретация являются элегантной демонстрацией больших возможностей понятия эффективной толщины.

Описание среды с параболическим распределением показателя преломления на языке электромагнитных мод

В основу лучевого описания, приведенного в предыдущем параграфе, было положено скалярное волновое уравнение

$$\nabla^2 \psi + k^2 n^2 (x) \psi = 0, \qquad (40)$$

где функция ψ описывает некоторую компоненту поля. Будет проведен более детальный анализ рассматриваемой задачи, основанном на решении уравнений Максвелла, учитывающий поляризационные свойства поперечных мод. Поэтому нас будет интересовать вывод волнового уравнения в гораздо более точном виде. Однако, как будет показано ниже, в случае слабонаправляющих сред приближение, обладает разумной точностью, и в предельном случае результаты, полученные при точном модовом описании, переходят в соответствующие результаты, полученные при лучевом анализе в предыдущем параграфе.

Векторное волновое уравнение для среды с градиентным распределением показателя преломления общего вида

Для среды с градиентным распределением показателя преломления общего вида уравнения Максвелла можно записать в следующем виде (Маркузе [409]):

$$\nabla \times H = n^2(x)\varepsilon_0 \frac{dE}{dt}, \qquad (41)$$

$$\nabla \times E = -\mu_0 \frac{dH}{dt}, \qquad (42)$$

$$\nabla \cdot \left[n^2(x)\varepsilon_0 E\right] = 0, \qquad (43)$$

$$\nabla \cdot H = 0. \qquad (44)$$

Подействуем оператором дивергенции на уравнение (42), а затем исключим из него H с помощью уравнения (41), тогда получим.

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \nabla \times \frac{d\mathbf{H}}{dt} = -\mu_0 n^2(x)\,\varepsilon_0 \frac{d^2\mathbf{E}}{dt^2} \qquad (45)$$

Далее, используя известное векторное тождество:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E}$$

и уравнение (43), уравнение (45) можно представить в виде

$$\nabla^2\mathbf{E} + \nabla\left(\frac{\mathbf{E} \cdot \nabla n^2(x)}{n^2(x)}\right) + k^2 n^2(x)\,\mathbf{E} = 0, \qquad (46)$$

где предполагается, что поле E зависит от времени как $\exp(-i\omega t)$, и используется обычное обозначение $k^2=\omega^2\mu_0\varepsilon_0$. Уравнение (46) является векторным волновым уравнением для электрического поля E.

Чтобы вывести соответствующее волновое уравнение для магнитного поля, подействуем сначала оператором дивергенции на уравнение (41), а затем исключим из него E с помощью уравнений (41) и (42), тогда получим:

$$\nabla \times (\nabla \times \mathbf{H}) = n^2(x)\,\varepsilon_0 \left(\nabla \times \frac{d\mathbf{E}}{dt}\right) + \varepsilon_0 \left(\nabla n^2(x) \times \frac{d\mathbf{E}}{dt}\right) =$$
$$= -\,n^2(x)\,\varepsilon_0\mu_0 \frac{d^2\mathbf{H}}{dt^2} + \frac{1}{n^2(x)} \left(\nabla n^2(x)\right) \times (\nabla \times \mathbf{H}). \qquad (47)$$

Поскольку в данном случае, согласно (44), V-H = 0, то, используя то же самое векторное тождество, как и выше, придем к векторному волновому уравнению для магнитного поля H:

$$\nabla^2 \mathbf{H} + \frac{1}{n^2(x)} [\nabla n^2(x)] \times (\nabla \times \mathbf{H}) + k^2 n^2(x)\,\mathbf{H} = 0. \qquad (48)$$

Далее, исходя из того факта, что показатель преломления n(x) - функция только поперечной координаты x, представим уравнения (46) и (48) в виде скалярных волновых уравнений для компонент поля E_x и H_x соответственно. Предполагая, что поле зависит от координаты z как exp(iβz), и считая d/dy=0, из (48) находим:

$$\frac{d^2 H_x}{dx^2} + [k^2 n^2(x) - \beta^2]\, H_x = 0. \qquad (49)$$

С другой стороны, из (46) следует:

$$\frac{d^2 E_x}{dx^2} + \frac{d}{dx}\left[\frac{E_x}{n^2(x)} \frac{dn^2(x)}{dx^2}\right] + [k^2 n^2(x) - \beta^2]\, E_x = 0. \qquad (50)$$

Чтобы исключить в (50) член dE_x/dx, сделаем замену:

$$E_x = \psi / n(x). \qquad (51)$$

Тогда, подставив (51) в (50), придем к скалярному волновому уравнению для величины ψ:

$$\frac{d^2\psi}{dx^2} + \left[\frac{1}{2n^2(x)}\frac{d^2(n^2(x))}{dx^2} - \frac{3}{4n^4(x)}\left(\frac{d(n^2(x))}{dx}\right)^2 + \right.$$
$$\left. + k^2n^2(x) - \beta^2\right]\psi = 0. \qquad (52)$$

Итак, мы вывели два скалярных волновых уравнения - (49) и (52) - для компонент поля H_x и E_x (через ψ) в среде, показатель преломления которой изменяется только в поперечном направлении x.

Решения для ТЕ-мод в среде с параболическим распределением показателя преломления

Пусть в этом случае $E_x = 0$; предположим, что компонента поля H_x известна как решение уравнения (49). Тогда, согласно (41) и (42), для остальных компонент поля имеем:

$$E_z = H_y = 0, \qquad (53)$$

$$E_y = -\frac{\omega\mu_0}{\beta}H_x, \qquad (54)$$

$$H_z = \frac{i}{\beta}\frac{dH_x}{dx}. \qquad (55)$$

В частном случае среды с параболическим распределением показателя преломления, когда $n^2(x)$, уравнение (49) можно записать следующим образом:

$$\frac{d^2H_x}{dx^2} + [k^2n_1^2(1 - 2\Delta(x/a)^2) - \beta^2]H_x = 0. \qquad (56)$$

Будем искать решение этого уравнения в виде:

$$H_x(x) = X(x)\exp(-x^2/w_0^2). \qquad (57)$$

Тогда после подстановки его в уравнение (56) получаем:

$$\frac{d^2X}{dx^2} - \frac{4x}{w_0^2}\frac{dX}{dx} - \frac{2}{w_0^2}(1 - 2x^2/w_0^2)X + \\ + [k^2 n_1^2 (1 - 2\Delta(x/a)^2) - \beta^2]X = 0. \qquad (58)$$

Определяя теперь «ширину» гауссова пучка ω_0, входящую в выражение (57), следующим образом:

$$w_0^2 = \frac{2a}{kn_1(2\Delta)^{1/2}} = \frac{2a^2}{v}, \qquad (59)$$

исключим в уравнении (58) член, содержащий x^2. Далее, вводя переменную:

$$x' = x\sqrt{2}/w_0$$

и учитывая (59), уравнение (58) для функции X можно представить в более известном виде

$$\frac{d^2X}{dx'^2} - 2x'\frac{dX}{dx'} + \left[(k^2 n_1^2 - \beta^2)\frac{w_0^2}{2} - 1\right]X = 0. \qquad (60)$$

Решения этого уравнения выражаются через полиномы Эрмита $H_N(x')$, индекс N которых связан с постоянной распространения β:

$$2N = (k^2 n_1^2 - \beta^2)w_0^2/2 - 1 \quad (N = 0,\ 1,\ 2,\ \ldots). \qquad (61)$$

Таким образом, согласно (57), решения уравнения (56) для ТЕ-мод в среде с параболическим распределением показателя преломления выражаются через функции Гаусса - Эрмита:

$$H_x = \frac{2^{1/4}}{\pi^{1/4}(2^N N! w_0)^{1/2}} H_N(x\sqrt{2}/w_0) \times \\ \times \exp(-x^2/w_0^2) \quad (N = 0, 1, 2, \ldots), \tag{62}$$

где константа находится из следующего условия нормировки:

$$\int_{-\infty}^{\infty} |H_x|^2 dx = 1. \tag{63}$$

Решение (62) можно выразить также через переменную υ. Для этого следует воспользоваться соотношением (59), связывающим υ и ω_0, что дает:

$$H_x = \frac{v^{1/4}}{\pi^{1/4}(2^N N! a)^{1/2}} H_N(v^{1/2}x/a) \times \\ \times \exp(-vx^2/(2a^2)) \quad (N = 0, 1, 2, \ldots). \tag{64}$$

На рисунке 1 представлены распределения полей для нескольких мод Гаусса — Эрмита низшего порядка. Характеристическое уравнение для этих мод задается выражением (61), которое в нормированных переменных можно записать следующим образом:

$$u^2 = v(2N + 1) \qquad (N = 0, 1, 2, \ldots). \tag{65}$$

Это уравнение совпадает с уравнением, полученным в рамках лучевого описания. Наконец, выражая это уравнение через нормированную постоянную распространения b, определяемую формулой, имеем

$$b = 1 - [(2N + 1)/v] \qquad (N = 0,\ 1,\ 2,\ \ldots). \tag{66}$$

На рисунке 2 представлена зависимость нормированной постоянной распространения b от переменной υ для нескольких мод низшего порядка, полученная согласно (66).

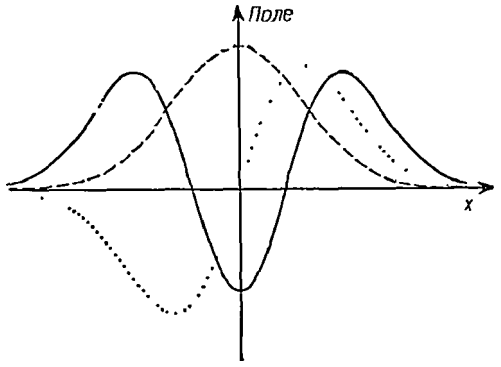


Рисунок 1 - Распределения полей для нескольких мод Гаусса-Эрмита низшего порядка. Штриховая кривая соответствует моде с N = 0, пунктирная кривая - моде с N = 1, а сплошная кривая - моде с N = 2.

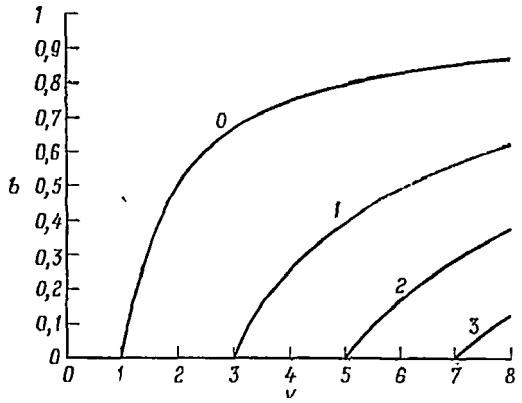


Рисунок 2 - Зависимость нормированной постоянной распространения b от нормированной частоты υ для среды с параболическим распределением показателя преломления, полученная согласно (66).

Обратим внимание на то, что при определенных значениях υ, а именно:

$$\upsilon_c = 2N + 1 \qquad (N = 0,\ 1,\ 2,\ \ldots), \tag{67}$$

нормированная постоянная распространения b = 0, т. е. $\beta = kn_2$. Однако это не является определением отсечки в волноводе в обычном смысле. При значениях υ, задаваемых выражением (67), поля мод не проявляют свойств, характерных для мод отсечки. Например, распределение полей в слоях оболочки трехслойного плоского волновода, рассмотренного ранее, при отсечке становится постоянным, так что энергия излучения во всем пространстве распределяется равномерно. Поля же мод Гаусса-Эрмита обладают свойствами, характерными для направляемых мод при всех значениях величины υ, поскольку распределение показателя преломления не ограничено в направлении x. Поэтому мы снова обратим внимание на то, что результаты для среды с параболическим распределением показателя преломления являются только полезным приближением при описании реальных волноводов, ограниченных оболочкой, и справедливы только для мод, поле которых компактно локализовано вблизи оси x = 0.

Решения для ТМ-мод в среде с параболическим распределением показателя преломления

Пусть в этом случае $H_x = 0$; предположим, что компоненту поля E_x можно найти как решение уравнения (52) с помощью преобразования (51). Тогда, согласно (41) и (42), для остальных компонент поля имеем:

$$E_y = H_z = 0, \tag{68}$$

$$E_z = \frac{i}{\beta n^2(x)} \frac{d}{dx}\left(n^2(x) E_x\right), \tag{69}$$

$$H_y = \frac{\omega n^2(x) \varepsilon_0}{\beta} E_x. \tag{70}$$

В случае среды с параболическим распределением показателя преломления уравнение (52) можно решить, разлагая члены, входящие в $n^2(x)$, в ряд и пренебрегая затем членами порядка $\Delta^2(x/a)^4$ и выше. При этом уравнение (52) принимает вид:

$$\frac{d^2\psi}{dx^2} + [(k^2n_1^2 - \beta^2 - 2\Delta/a^2) - (k^2n_1^2 2\Delta/a^2 + 16\Delta^2/a^4)\, x^2]\,\psi = 0. \qquad (71)$$

Уравнение (71) по форме имеет сходство с уравнением (56). Следовательно, решение уравнения (71) можно сразу записать по аналогии с решением уравнения (56):

$$E_x = \frac{\psi}{n(x)} = \frac{2^{1/4}}{n(x)\pi^{1/4}(2^N N!\, w_0)^{1/2}} H_N(x\sqrt{2}/w_0) \times$$
$$\times \exp(-x^2/w_0^2) \quad (N = 0,\ 1,\ 2,\ \ldots), \qquad (72)$$

где ширина пучка ω_0 в данном случае задается выражением:

$$w_0^2 = 2(k^2n_1^2 2\Delta/a^2 + 16\Delta^2/a^4)^{-1/2}. \qquad (73)$$

Наконец, соответствующее характеристическое уравнение имеет вид:

$$\beta^2 = k^2n_1^2 - 2\Delta/a^2 - (2N+1)(k^2n_1^2 2\Delta/a^2 +$$
$$+ 16\Delta^2/a^4)^{1/2} \quad (N = 0,\ 1,\ 2,\ \ldots). \qquad (74)$$

Обратим внимание на то, что решения для ТЕ- и ТМ-мод при $2\Delta << k^2 n_1^2 a^2$ совпадают, т. е в этом пределе выражения (72) - (74) сводятся к выражениям (64), (59) и (61) соответственно.

Групповая задержка и коэффициент локализации излучения

Чтобы сравнить выражения для постоянной распространения и, следовательно, для групповой задержки, которые получаются для ТЕ- и ТМ-мод из выведенных выше выражении, мы можем разложить выражения (61) и (74) соответственно в ряд по возрастающим степеням $\Delta^{1/2}$. Сохраняя члены разложения вплоть до членов порядка Δ, из выражения (61) получаем:

$$\text{ТЕ-}мода: \quad \beta \approx kn_1 - (2\Delta)^{1/2}(N + 1/2)/a - \Delta(N + 1/2)^2/(a^2 kn_1), \qquad (75)$$

а из выражения (74) имеем:

$$\text{ТМ-}мода: \quad \beta \approx kn_1 - (2\Delta)^{1/2}(N + 1/2)/a - \\ - \Delta[(N + 1/2)^2 + 1]/(a^2 kn_1). \qquad (76)$$

Дифференцируя каждое из этих выражений, находим величину групповой задержки τ:

$$\text{ТЕ-мода:} \quad \tau \approx \frac{Ln_1}{c}\left[1 + \Delta\left(N + \frac{1}{2}\right)^2 / \left(a^2 k^2 n_1^2\right)\right], \qquad (77)$$

$$\text{ТМ-мода:} \quad \tau \approx \frac{Ln_1}{c}\left[1 + \Delta\left(\left(N + \frac{1}{2}\right)^2 + 1\right) / \left(a^2 k^2 n_1^2\right)\right]. \qquad (78)$$

Из выражений (77) и (78) следует:

1) групповые задержки для всех ТЕ- и ТМ-мод, кроме нескольких мод низшего порядка, почти одинаковы;

2) для малых значений $\Delta / \left(a^2 k^2 n_1^2\right)$ ТЕ- и ТМ-моды квазивырождеиы;

3) для малых значений $\Delta / \left(a^2 k^2 n_1^2\right)$ групповая задержка, по существу, не зависит от номера моды, т. е. одинакова для всех мод.

Последний результат в случае слабонаправляющей среды (ТЕ-и ТМ-моды квазивырождены) можно сформулировать и в ином виде с помощью

нормированного уравнения (66). Согласно (66), нормированный дисперсионный параметр d(υb)/dυ, удовлетворяет условию:

$$d(vb)/dv = 1 \quad \text{для всех } N, \tag{79}$$

т. е. в этом приближении групповая задержка для всех мод одинакова.

Обратимся к вычислению коэффициента локализации излучения Г, определяемого как отношение энергии излучения, локализованного в сердцевине, к полной энергии.

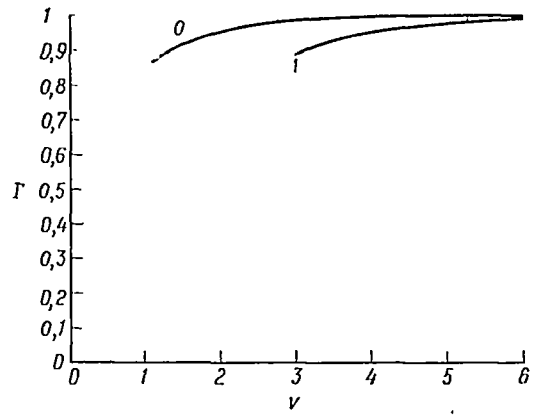


Рисунок 3 - Зависимость коэффициента локализации излучения Γ от υ для мод с N = 0 н N = 1 в среде с параболическим распределением показателя преломления, рассчитанная по (81) и (82)

В пределе малых Δ, т. е. в случае вырожденных ТЕ- и ТМ-мод, для этого коэффициента имеем:

$$\Gamma = \frac{v^{1/2}}{\pi^{1/2} 2^N N! a} \int_{-a}^{a} [H_N(v^{1/2}x/a)]^2 \exp(-vx^2/a^2)\, dx$$
$$(N = 0, 1, 2, \ldots), \qquad (80)$$

где для нормированных полей было использовано выражение (64). Несколько первых полиномов Эрмита задаются следующими выражениями:

$$H_0(y) = 1, \quad H_1(y) = 2y, \quad H_2(y) = 4y^2 - 2 \quad \text{и т. д.}$$

Оказывается, что величину Γ всегда можно выразить через функцию ошибок erf ($\upsilon^{1/2}$). Так, для двух первых мод имеем:

$$N = 0: \quad \Gamma = erf\left(\upsilon^{1/2}\right), \qquad (81)$$

$$N = 1: \quad \Gamma = erf\left(\upsilon^{1/2}\right) - \frac{2\upsilon^{1/2}}{\pi^{1/2}} e^{-\upsilon}. \qquad (82)$$

Зависимость этих функций от v дана на рисунке 3. Отметим, что значения Г приведены только для таких υ, которые превышают величину отсечки υ_c, задаваемую выражением (67). Поскольку условие отсечки (67) является искусственным, коэффициент локализации при этих значениях величин отсечки не обращается в нуль, и при использовании полученных результатов для описания реальных волноводов с параболическим изменением показателя преломления следует соблюдать большую осторожность.

Лекция 10

Элементы интегрально-оптических схем. Связь между волноводами. Разветвители. Рупорные переходы. Фокусирующие элементы. Поляризационные элементы. Спектральные фильтры и резонаторы.

Каждый тип элементов связи обладает целым рядом достоинств и недостатков; не один из них не является наилучшим для всех случаев использования интегральной оптики. Поэтому как потребителям оптических интегральных схем, так и их разработчикам необходимо знать особенности каждого вида элементов связи.

Для изготовления элементов волноводной связи в общем случае используют такие методы, как фоторезистивное маскирование, распыление, эпитаксиальное выращивание слоев, которые были рассмотрены ранее. Однако в данной лекции рассматриваются некоторые специальные методы изготовления элементов связи, например экспонирование фоторезиста голографическими методами для изготовления решеточных элементов связи.

Основы оптического согласования

Основными характеристиками любого элемента связи являются его эффективность и медовая избирательность. Эффективность сопряжения, как правило, определяется частью общей энергии оптического пучка, введенной в волновод или выведенной из него. Кроме того, ее можно определять, исходя из потерь сопряжения (в дБ). Для элементов связи с модовой избирательностью эффективность сопряжения можно выразить независимо для каждой моды, тогда как в случае многомодовых элементов связи они, как правило, характеризуются общей эффективностью. Однако имеются случаи, когда можно определить относительные модовые эффективности и для многомодовых элементов связи.

Поэтому самое общее определение эффективности сопряжения дается следующим отношением:

$$\eta_{Cm} = \frac{энергия,\ \ введенная\ \ в\ \ моду\ \ m-го\ \ порядка}{полная\ \ энергия\ \ оптического\ \ пучка}, \qquad (1)$$

а потери сопряжения равны:

$$L_{Cm} = -10\log\left(\frac{энергия,\ \ введенная\ \ в\ \ моду\ \ m-го\ \ порядка}{полная\ \ энергия\ \ оптического\ \ пучка}\right). \qquad (2)$$

Если нет возможности определить энергию в каждой моде отдельно, то используют усредненное значении η_{Cm} и L_{Cm}. Эффективность сопряжения очень сильно зависит от согласования между полями оптического пучка и волновой моды. Этот принцип можно хорошо проиллюстрировать при рассмотрении случая поперечных элементов связи.

Поперечные элементы связи

Поперечные элементы связи представляют собой такие элементы, в которых пучок фокусируется непосредственно на специально подготовленное поперечное сечение волновода. Для случая пучка в свободном пространстве (воздухе) такую фокусировку можно выполнить с помощью линзы. Поперечное сопряжение двух твердотельных волноводов можно осуществить, соединяя встык отполированные или сколотые торцы волноводов.

Прямое фокусирование

Самым простым методом поперечного сопряжения лазерного пучка с волноводом является прямое фокусирование, или метод светящегося конца, представленный на рисунке 1.

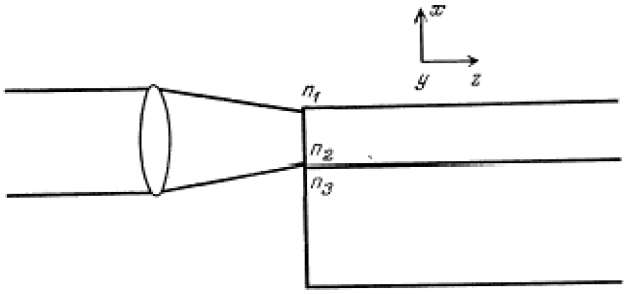


Рисунок 1 – Поперечное сопряжение методом светящегося конца.

Можно использовать как плоские, так и канальные волноводы, однако в данном случае допустим, что мы имеем дело с плоским волноводом. Передача оптической энергии лазерного пучка достигается за счет сопряжения полей пучка и волноводной моды. Эффективность сопряжения можно рассчитать из интеграла перекрытия картин полей падающего пучка и волноводной моды, который задается следующим выражением:

$$\eta_{Cm} = \left(\int A(x)B_m(x)dx\right)^2 / \int A(x)A^*(x)dx \int B_m(x)B_m^*(x)dx, \qquad (3)$$

где A(x) - амплитудное распределение входного лазерного пучка, а B(x) - амплитудное распределение моды m-го порядка.

Метод светящегося конца особенно полезен для сопряжения пучков от газовых лазеров и основной моды в волноводе, что -связано с относительно хорошим согласованием между гауссовым профилем лазерного пучка и формой TE_0-моды волновода. Конечно, для оптимального сопряжения диаметр пучка должен быть достаточно близким к толщине волновода. В принципе если контуры распределения полей тщательно согласованы, то можно достичь 100%-ной эффективности сопряжения. Однако на практике обычно достигают эффективности 60%. Это связано с тем, что толщины пленок волноводов имеют порядок 1 мкм, поэтому юстировка весьма критична. Такое сопряжение сравнительно часто используется в лабораторных условиях благодаря. Однако

трудности, связанные с монтажом оптической аппаратуры для столь точной юстировки без использования оптических скамей, существенно ограничивают его практическое использование.

Торцевое сопряжение встык

Поперечное сопряжение используется на практике в приложениях интегральной оптики в тех случаях, когда волновод сопрягается с полупроводниковыми лазерными источниками или с другим волноводом. При этом можно использовать стыковое соединение, как показано на рисунке 2. В таком случае можно достичь очень высокой эффективности сопряжения, что определяется тем, что ширину волновода можно приблизительно подогнать под ширину излучающего слоя лазера, и тем, что распределение поля в основной моде лазера достаточно хорошо согласуется с полем TE_0-моды волновода. Этот метод особенно полезен для сочленения лазерных диодов с плоскими волноводами, так как достичь эффективного сочленения инжекционного лазера с тонкопленочными волноводами трудно при использовании таких элементов связи, как призменные, решеточные элементы или сужающийся край пленки. Причина этого в том, что инжекционные лазеры дают сравнительно плохо коллимированные пучки с угловой полушириной расходимости обычно 10-20°. Как будет показано дальше в данной главе, призменные, решеточные элементы связи и элементы связи на утончающихся пленках крайне чувствительны к углу падения пучка излучения и требуют для эффективного сопряжения коллимации пучков менее 1°.

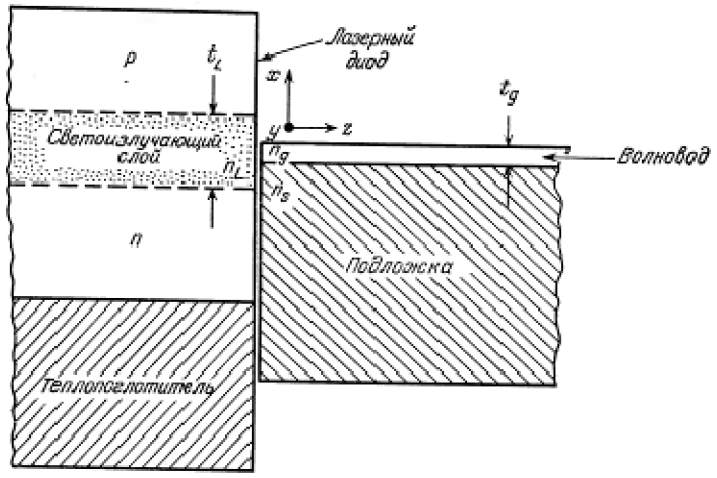


Рисунок 2 – Параллельное соединение лазерного диода с тонкопленочным волноводом

Для случая сочленения лазерного диода, работающего на основной моде TE_0, и плоского волновода, как это показано на рисунке 2, эффективность сопряжения на TE_0-моде дается выражением:

$$\eta Cm = \frac{64}{(m+1)^2\pi^2}\frac{n_L n_g}{(n_L+n_g)^2}\cos^2\left(\frac{\pi t_g}{2t_L}\right)\frac{1}{\left(1-\frac{t_g}{(m+1)t_L}\right)^4}\frac{t_g}{t_L}\cos^2\left(\frac{m\pi}{2}\right), \quad m = 0,1,2,\ldots \; . \qquad (4)$$

Это выражение основано на предположении, что все волноводные моды хорошо ограничены и $t_g \le t_L$. Интересно отметить отсутствие сопряжения для волноводных мод нечетного порядка, что определяется последним сомножителем в выражении (4). Это связано с тем, что распределения полей имеют компенсирующие лепестки, когда берутся интегралы перекрывания с четной (m = 0) модой лазерного излучения. Первый сомножитель в выражении (4) является нормировочным членом, второй обусловлен отражением, которое имеет место на границе лазер - волновод. Остальные сомножители связаны с рассогласованием распределений полей в лазере и волноводе.

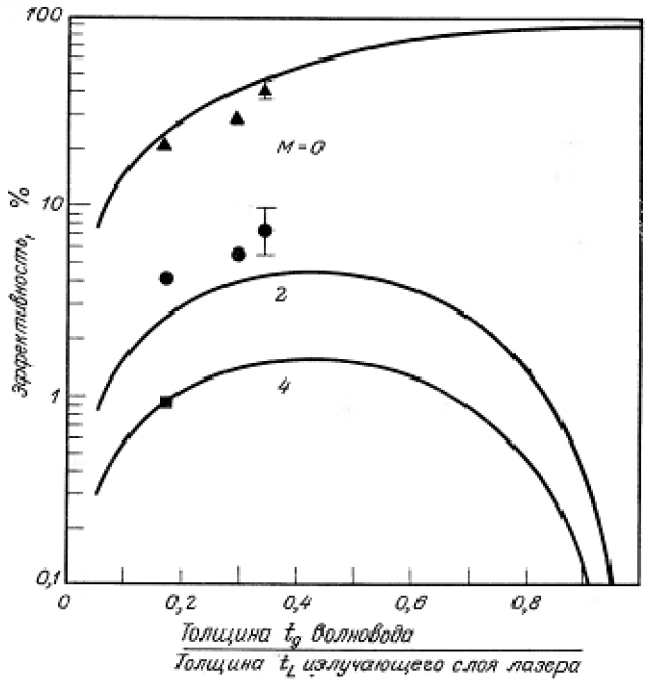


Рисунок 3 - Сравнение экспериментальных и теоретических значений эффективности связи в зависимости от толщины волновода: n_L=3,6, n_g=2,0, t_L=5,8 мкм.

Рассчитанные кривые зависимости η_{Cm} от относительной толщины волноводов представлены на рисунке 3 наряду с экспериментальными данными, полученными для сочленения лазерного диода на GaAs с волноводом из Ta_2O_5 на стеклянной подложке. Если $t_g \approx t_L$, эффективность сопряжения для моды наинизшего порядка теоретически приближается к 100%. В таком случае сопряжение для волноводных мод более высокого порядка почти равно нулю. Данные, касающиеся эффективности сопряжения на рисунке 3, являются оптимальными значениями, соответствующими идеальной настройке лазера и волновода. Эффективность сопряжения очень чувствительна к поперечным боковым рассогласованиям в направлении x. Смещение x волновода относительно лазера, как показано на рисунке 4, уменьшает эффективность сопряжения в соответствии с выражением:

$$P / P_0 = \cos^2(\pi X / t_L), \qquad (5)$$

где P_0 - энергия сопряжения при X=0. В этом выражении предполагают, что $t_g<t_L$ и $X \le (t_L - t_g)/2$. Штриховая линия на рисунке 4 соответствует теоретически вычисленным значениям P/P_0 для t_L=5,8 мкм и t_g=2 мкм, в то время как сплошные кривые представляют собой экспериментально полученные данные. Для определения относительной энергии, которая вводится при стыковом соединении для каждой из трех наблюдаемых мод, использовался призменный элемент связи на выходе волновода.

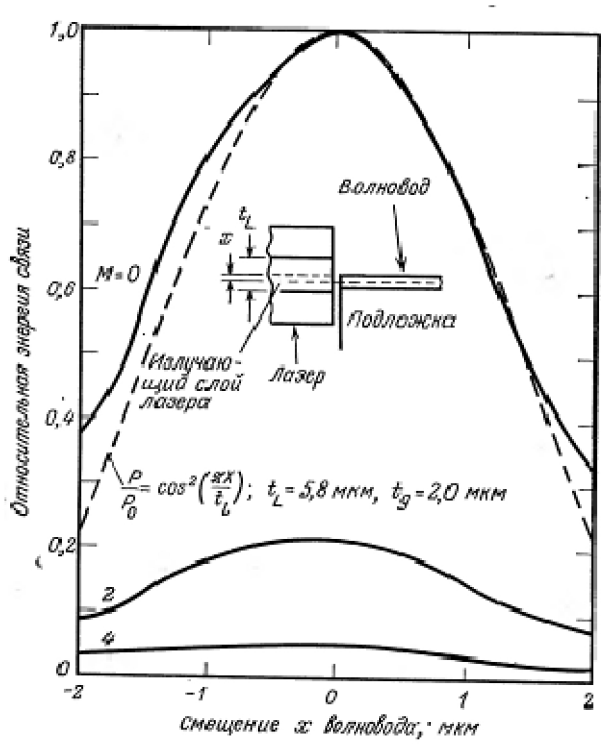


Рисунок 4 – Сравнение экспериментальных данных по эффективности связи (сплошная линия) с теоретическими (штриховая линия) в зависимости от поперечного смещения лазера и волновода.

Величина зазора между лазером и волноводом в направлении оси z также является очень критичной, и для достижения оптимального сопряжения она должна контролироваться с точностью порядка длины волны излучения. На рисунке 5 представлена зависимость экспериментально полученных изменений

введенной мощности от смещения в направлении оси z. Осциллирующая форма данной кривой связана с модуляцией эффективной отражательной способности выходной поверхности лазера за счет резонансных явлений в эталоне Фабри-Перо, образованном параллельными плоскостями границ раздела лазер—волновод. В принципе от этого эффекта можно избавиться, если использовать жидкость с промежуточным значением показателя преломления между лазером и волноводом. Если сделать это, то введенная мощность изменяется плавно в зависимости от смещения до оси z, как показано на рисунке 5 штриховой линией.

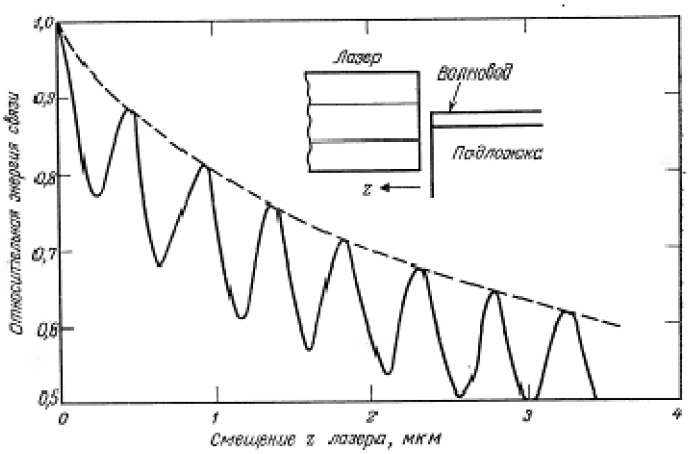


Рисунок 5 – Экспериментальная зависимость эффективности связи от величины зазора между лазером и волноводом

Представленные до сих пор результаты показывают, что стыковое соединение может быть достаточно эффективным для сопряжения лазерных диодов с волноводами. Кэмпбелл получил подобные результаты для стыкового соединения лазера и оптического волокна. Однако очевидно, что для получения оптимального сопряжения требуется субмикрометровая настройка. Настройку с такими допусками можно реализовать, если использовать микрометрические головки с пьезоэлектрическим приводом, которые отличаются тем, что имеют маленький пьезоэлектрический кристалл-держатель, закрепленный на конце обычного микрометрического винта. Грубая юстировка проводится за счет вращения микрометрического винта, а окончательная подстройка производится

при подаче напряжения на пьезоэлектрическую насадку для передвижения лазера (или волновода), в то время как вводимая оптическая мощность сопряжения измеряется с помощью соответствующего элемента связи и фотодетектора. Имеются в продаже пьезоэлектрические микрометры чувствительностью выше 40 А на вольт, работающие в диапазоне до 2000 В. Таким образом, настройку на уровне лучше 0,1 мкм можно сравнительно легко реализовать. После того как один раз -проведена подобная настройка, радиатор лазера можно прочно скрепить с держателем волновода с помощью эпоксидной смолы или металлического крепежа. Так как размер и масса лазера и ОИС сравнительно малы, чувствительность к вибрациям не является серьезной проблемой и настройка не нарушается. В самом деле, планируется, что стыковое сочленение будет использоваться для соединения лазерных источников с волноводами в интегрально-оптических анализаторах спектров высокой частоты.

Призменные элементы связи

Поперечное сочленение можно использовать лишь тогда, когда обнажена плоскость поперечного сечения торцевого конца волновода. Во многих случаях задача состоит в том, чтобы ввести излучение в волновод, находящийся внутри ОИС, когда доступна лишь одна поверхность. Если пытаться проводить фокусирование излучения на поверхность волновода при косом падении, как показано на рисунке 6, то это связано с принципиальными трудностями

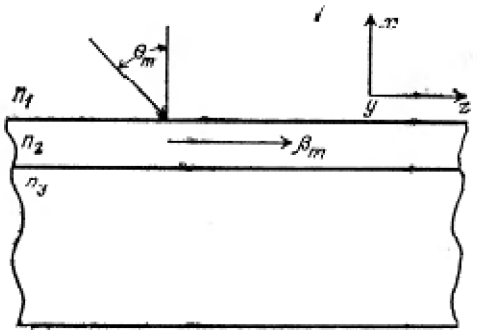


Рисунок 6 - Схема пробного эксперимента по наклонному введению света в волновод через его поверхность.

Чтобы сопряжение имело место, необходимо равенство составляющих фазовых скоростей волн в направлении оси z как для волновода, так и для пучка. Таким образом, надо удовлетворить условие согласования фаз, которое требует:

$$\beta_m = kn_1 \sin\Theta_m = (2\pi / \lambda_0) n_1 \sin\Theta_m . \qquad (6)$$

Однако в ранее было показано, что для мод в волноводе

$$\beta_m > kn_1 . \qquad (7)$$

Объединяя выражения (6) и (7), получаем $\sin\Theta_m > 1$, что, конечно, невозможно.

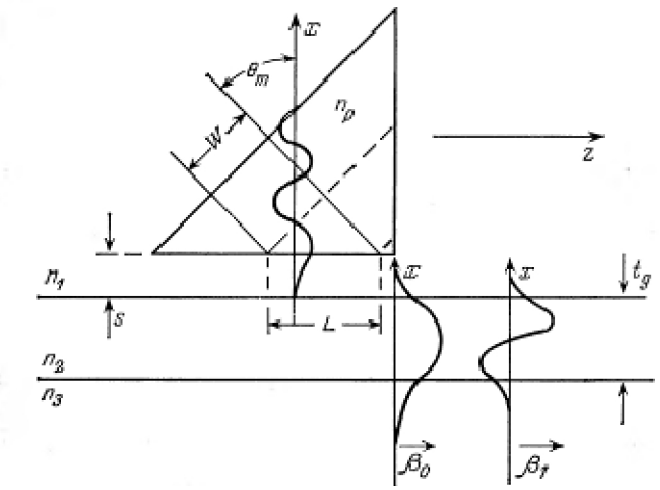


Рисунок 7 - Схема призменного элемента связи.

Одним из решений этой проблемы согласования фаз является использование призмы, как показано на рисунке 7. Пучок излучения шириной W направляется на основание призмы, которая имеет $n_p > n_1$. Пучок испытывает полное внутреннее отражение на поверхности раздела $n_p - n_1$, что приводит к установлению моды стоячей волны в призме (рисунок 7). Эта мода является стационарной в направлении оси x, но движется в направлении оси z с фазовой

постоянной β_m. В волноводе может существовать несколько мод, движущихся в направлении оси z с фазовыми постоянными β_m. Все такие моды имеют экспоненциальные «хвосты», немного заходящие за поверхность раздела $n_1 - n_2$. Если зазор s между призмой и волноводом настолько мал, что «хвосты» волноводных мод перекрываются хвостом моды призмы, то при выборе Θ_m таким образом, что $\beta_p = \beta_m$ имеет место когерентная перекачка энергии моды призмы в волноводную моду m-го порядка. Условие для согласования значений β определяется следующим соотношением:

$$\left(2\pi n_p / \lambda_0\right)\sin\Theta_m = \beta_m . \qquad (8)$$

Хотя возбуждение определенной моды будет иметь место при определенном значении угла Θ_m, с помощью одной-единственной призмы можно осуществить возбуждение нескольких различных мод просто за счет изменения угла падения оптического пучка. Падающий пучок ,не обязательно должен быть перпендикулярным к поверхности призмы. Но если направление падения пучка не перпендикулярно поверхности призмы, то условия преломления на поверхности раздела требуют изменения выражения (8) для угла Θ_m.

Процесс перекачки оптической энергии через перекрывающиеся «хвосты» мод, в то время как падающий пучок имеет тенденцию к полному внутреннему отражению в призме, иногда называют оптическим туннелированием по аналогии с туннельным эффектом в квантовой механике при прохождении частиц через энергетические барьеры. Моды волновода очень слабо связаны с модой призмы. Поэтому имеет место сравнительно небольшая деформация форм мод. Разумеется, условие:

$$\Theta_m > \Theta_C = \arcsin\left(n_1 / n_p\right), \qquad (9)$$

также должно выполняться для достижения полного внутреннего отражения, где Θ_C — критический угол полного внутреннего отражения.

Длина L, на которой происходит взаимодействие между модами призмы и волновода, определяется размерами призмы. Из теории слабосвязанных мод следует, что полный обмен энергией между фазово-согласованными модами имеет место, если длина взаимодействия в направлении оси z удовлетворяет соотношению

$$\chi L = \pi / 2, \qquad (10)$$

где χ — коэффициент связи. Коэффициент χ зависит от n_p, n_1 и n_2, определяющих форму «хвостов» мод, а также от зазора s между призмой и волноводом. Из выражения (10) можно определить длину, необходимую для полной перекачки:

$$L = W / \cos\Theta_m = \pi / 2\chi . \qquad (11)$$

Таким образом, для заданной длины L значение коэффициента связи, требуемое для полной перекачки, имеет вид

$$\chi = \pi \cos\Theta_m / 2W . \qquad (12)$$

При выводе этого условия для полной перекачки предполагается, что амплитуда электрического поля равномерно распределена по всей ширине W пучка излучения. На практике это никогда не реализуется. Для пучка гауссовой формы это приводит к максимальной эффективности связи, равной примерно 80%. Кроме того, видно, что для достижения 100%-ной связи, даже в случае равномерного распределения энергии по ширине пучка, правая граница пучка должна точно проходить через вершину прямого угла призмы. Если она проходит справа от этой точки, то часть падающей оптической энергии будет либо отражаться, либо непосредственно проходить в волновод и будет исключаться из моды призмы. Если же она будет проходить левее этой точки, то часть энергии, введенной в волновод при сопряжении, будет выводиться за счет сопряжения обратно в призму.

Призменные элементы связи довольно часто используются в различных областях применения интегральной оптики благодаря их универсальности. Их можно использовать в качестве как входных, так и выходных элементов. При использовании их в качестве выходных элементов призма располагается точно так, как изображено на рисунке 7, за исключением того, что направление распространения каналированного в волноводе излучения совпадает с отрицательным направлением оси z. Если в волноводе распространяется несколько мод, то вывод света происходит под своим определенным углом для каждой определенной моды. В связи с данной особенностью призменные элементы связи можно использовать в качестве приборов для анализа относительной мощности в каждой из волноводных мод. Кроме того, можно использовать призму для определения волноводных потерь, передвигая ее относительно волновода. Однако при этом следует обратить особое внимание на то, чтобы механическое давление, оказываемое на призму, было одинаково во всех измерениях, чтобы зазор, а следовательно, и коэффициент связи были бы постоянными.

Одним из недостатков призменных элементов связи является то, что величина n_p должна быть больше не только величины n_1 но и величины n_2. Это так, потому что показатель преломления n_2 волновода вообще близок к показателю преломления n_3 подложки, что приводит к выводу

$$\beta_m \approx kn_2 = (2\pi / \lambda_0)n_2 . \qquad (13)$$

Так как $\sin\Theta_m \leq 1$, то из выражения (13), связанного с (8), следует, что $n_p > n_2$. В случае стеклянных волноводов с показателями преломления ~1,5 сравнительно легко подобрать подходящий материал для призмы с $n_p > n_2$. Однако для полупроводниковых волноводов, материалы которых обычно имеют значения показателей преломления ~3 или 4, значительно труднее подобрать подходящий материал для призмы. При под боре материала для призм следует учитывать как показатель преломления, так и его .пропускание излучения на рабочей длине

волны. В таблице 1 перечислены материалы, пригодные по своему оптическому качеству для изготовления призменных элементов связи, а также их показатели преломления в различных волновых диапазонах.

Таблица 1 - Материалы, пригодные для изготовления призменных элементов связи

Материал	Приблизительное значение пока-	Диапазон длин волн
Титанат стронция	2,3	Видимая — ближняя ИК
Рутил	2,5	Видимая — ближняя ИК

Другим недостатком призменных элементов связи является то, что падающий пучок должен быть очень хорошо коллимирован из-за острой зависимости от угла эффективности связи для определенной моды. По этой причине призменные элементы связи нельзя эффективно использовать в сочетании с полупроводниковыми лазерами, которые дают пучки с угловой полушириной расходимости излучения 10 - 20°.

Призменные элементы связи достаточно хорошо подходят для проведения исследований в лабораторных условиях, где требуется большая маневренность относительно положения падающего пучка излучения. Однако необходимость постоянного механического давления для удержания призмы на месте делает их менее полезными для прикладных областей интегральной оптики при наличии вибраций и изменений температуры. Для преодоления этих трудностей без потери преимуществ модовой избирательности можно использовать решеточные элементы связи.

Решеточные элементы связи

Решеточные элементы, подобно призменным, используются для обеспечения фазового согласования между определенной волноводной модой и

оптическим пучком, падающим наклонно к поверхности волновода, как показано на рисунке 8. Необходимо помнить, что без решетки условия фазового согласования задаются выражением (6) и их нельзя удовлетворить при любом угле Θ_m.

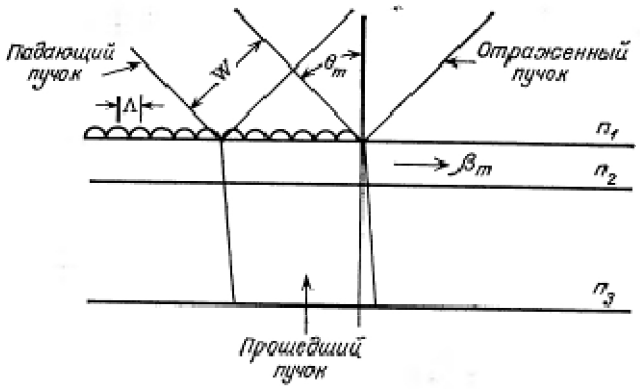


Рисунок 8 - Схема решеточного элемента связи.

Вследствие своей периодической структуры решеточный элемент связи возмущает волноводные моды в области, находящейся под решеткой, вызывая тем самым разложение каждой моды в ряд пространственных гармоник с постоянными распространениями в направлении оси z, определяемыми выражением

$$\beta_v \approx \beta_0 + (2\pi v / \Lambda), \qquad (14)$$

где v=0, ±1, ±2,..., а Λ — период решетки. Слагаемое β_0 приблизительно равняется β_m для определенной моды в области данного волновода, не закрытой решеткой. Так как величина v может быть отрицательной, то в таком случае условие согласования фаз (6) может выполняться так, чтобы:

$$\beta_v \approx kn_1 \sin \Theta_m, \qquad (15)$$

даже при $\beta_m > kn_1$. Так как все пространственные гармоники каждой моды согласованы для образования законченного поверхностного волнового поля в области решетки, энергия, введенная из пучка в любую пространственную гармонику, очевидно, переходит в основную (v=0) гармонику по мере ее распространения вправо и дальше за пределы области решетки. Эта основная гармоника очень близка к β_m-моде (и в конце концов становится ею) вне области решетки. Поэтому решеточный элемент связи можно использовать для избирательной передачи энергии от оптического пучка определенной моде волновода за счет соответствующего подбора угла падения пучка. Решеточный элемент связи можно использовать и для выведения энергии из волновода, так как в силу принципа взаимности энергия волноводной моды будет согласованно выводиться при определенных углах Θ_m, соответствующих конкретным модам. В предыдущем разделе описывалась работа решеточного элемента связи в обычных и относительно простых терминах. Однако детали явления такого сопряжения достаточно сложны и очень сильно зависят от формы поперечного сечения штрихов решетки, а также расстояния между ними. Как и в случае призменных элементов связи, теоретически можно достичь оптимальной эффективности связи около 80%, когда с помощью решетки вводят пучок света с гауссовой формой распределения. Однако типичные решетки «без блеска» (с симметричными профилями) обычно имеют эффективность 10 - 30%. Основной причиной такого положения является то, что большая часть падающей оптической энергии, как правило, проходит через волновод и теряется в подложке, так как в отличие от призменных элементов связи решетка работает в условиях, далеких от условий полного внутреннего отражения. Кроме того, энергия может уходить в дифракционные пучки более высокого порядка, производимые решеткой, несмотря на то, что отношение периода решетки к длине волны в волноводе составляет приблизительно 1.

Эффективность решеточного элемента связи можно значительно повысить, если придать асимметричный профиль его штрихам, чтобы иметь «блеск» решетки для получения оптимальных условий ввода света на рабочей длине

волны. Например, теоретический максимум эффективности связи моды TE_0 или TM_0 с пучком, выходящим из воздуха, для решетки с симметричным профилем составляет приблизительно 50%, тогда как для решетки с асимметричными пилообразными профилями штрихов - более 95%. Такие теоретические предсказания подтверждаются экспериментальными результатами, полученными на решетках «с блеском», и дальнейшими теоретическими работами.

Основное преимущество решеточных элементов связи заключается в том, что после изготовления они являются составной частью волноводной структуры. Следовательно, их эффективность сопряжения остается постоянной и заметно не изменяется при воздействии вибраций или окружающих условий. Кроме того, решеточные элементы можно использовать на волноводах из полупроводников с высокими показателями преломления, к которым трудно подобрать подходящий материал для изготовления призменных элементов. Однако из-за сильной зависимости эффективности решеточных элементов от угла их нельзя успешно использовать с относительно расходящимися пучками полупроводниковых лазеров. Может быть, самый большой недостаток решеточных элементов состоит в том, что они очень трудны в изготовлении, требующем использования сложных методов маскирования и травления.

Решеточную структуру можно сформировать либо маскированием с последующим травлением поверхности волновода, либо маскированием поверхности и нанесением на нее тонкопленочного рельефа решетки. Как при том, так и при другом методе самой трудной частью технологии является процесс получения четкой картины близко расположенных штрихов решетки. Расстояние между штрихами должно составлять порядка длины волны излучения в материале волновода. Поэтому для материалов с показателем преломления 1,4 - 4, используемых в видимой и ближней инфракрасной областях излучения, расстояния между штрихами решетки обычно находятся в пределах 100 - 300 нм. Обычные фоторезисты, используемые в микроэлектронной промышленности, имеют адекватное разрешение, однако практический предел обычных фотомасок составляет около 1 мкм. Поэтому обычно решеточные элементы связи

изготовляют с использованием оптического интерференционного процесса, который иногда называют голографическим (рисунок 9).

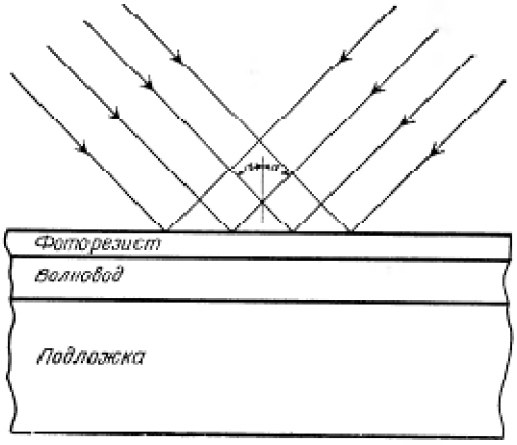


Рисунок 9 - Голографическое экспонирование фоторезиста,

При проведении такого процесса подложка с волноводом сначала покрывается фоторезистом. Затем данный фоторезист экспонируют под интерференционной картиной, которая образуется на поверхности от когерентных лучков лазера, как показано на рисунке 9 (Оба пучка обычно получают от одного лазера за счет использования расщепителя пучка). Простые геометрические соображения показывают, что соотношение между периодом решетки Λ и углом падения пучка α имеет вид:

$$\Lambda = \lambda_0 / 2\sin\alpha , \qquad (16)$$

Из выражения (16) очевидно, что Λ ограничено значением большим $\lambda_0/2$. Однако если использовать прямоугольную призму, как показано на рисунке 10, то при экспонировании можно получить решетки с меньшим расстоянием между штрихами, чем, при обычном использовании данного лазерного источника.

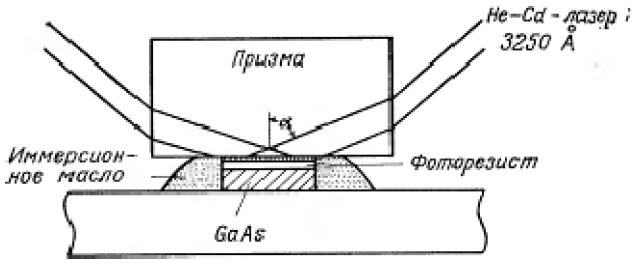


Рисунок 10 - Схема использования призмы в голографическом методе изготовления дифракционных решеток первого порядка в волноводе из GaAs.

В этом, случае соотношение (16) принимает вид

$$\Lambda = \lambda_0 / 2n \sin\alpha , \qquad (17)$$

где n - показатель преломления материала призмы. Используя призму из кварца и He-Cd лазер, получают решетки с расстояниями между штрихами до 115 нм. После облучения фоторезист обрабатывается с помощью обычных методов до получения необходимой маски на поверхности волновода. Далее для формирования структуры решетки используют либо химическое, либо ионное травление. Обычно считают, что при ионном травлении получаются решетки с более однородной структурой, а при химическом травлении материал волновода меньше повреждается. Кроме того, химическое травление при тщательном подборе травителя дает возможность получать более глубокие канавки, без побочной обработки материала. Например, при использовании раствора NH_4OH можно стравливать GaAs до глубины более 100 мкм. Другим преимуществом химического травления является то, что соответствующим подбором травителя и ориентации материала подложки можно достичь условий анизотропного травления, которое приводит к асимметричным решеткам «с блеском» на нужной длине волны.

В качестве варианта метода получения дифракционных решеток травлением служит метод получения решеток нанесением на волноводную поверхность тонкопленочных штрихов с использованием фоторезистивной маски для

ограничения формы штрихов. Фактически сам фоторезист можно использовать в качестве материала, наносимого в виде штрихов решетки. Обычно такие методы изготовления приводят к решеточным элементам связи с большими потерями (за счет рассеяния и поглощения излучения) по сравнению с элементами связи, изготовленными с использованием травления, так как часть пучка при выходе из волновода задерживается.

Сужающиеся элементы связи

Принцип действия сужающихся элементов связи основан на том факте, что волноводы толщиной, меньшей требуемой условием отсечки, передают энергию в излучательные моды.

Толщина волновода уменьшается в области связи, тем самым образуя волновод клинообразной формы с уменьшающейся критической длиной волны, как показано на рисунке 11. Механизм связи удобнее представить себе, если использовать оптико-лучевое приближение. Направляемая волна, подходя к данному элементу связи, испытывает зигзагообразные отражения с постоянно уменьшающимся углом падения на плоскость раздела волновод — подложка, который отсчитывается от нормали к этой плоскости. Когда значение этого угла падения станет меньше критического для полного внутреннего отражения, оптическая энергия будет переходить в подложку волновода. Энергия от последующих лучей также выходит из волновода аналогичным образом, поэтому можно достичь эффективности связи вплоть до 70%. Остальные 30% приходятся на потери, связанные в основном с излучением в воздух, когда толщина волновода достигает значения, соответствующего отсечке. После достижения точки отсечки мода волновода полностью выводится на расстоянии, составляющем около восьми длин волн излучения в вакууме. Главное преимущество сужающихся элементов связи состоит в том, что они просты в изготовлении и достаточно хорошо действуют в качестве выходных элементов связи, но формируют расходящийся луч (рисунок 11), который имеет угловые размеры 1-20° в зависимости от

величины сужения. Подобный расходящийся луч иногда неудобно использовать, однако такое положение допустимо во многих случаях использования, когда форма выходного луча не является критическим параметром.

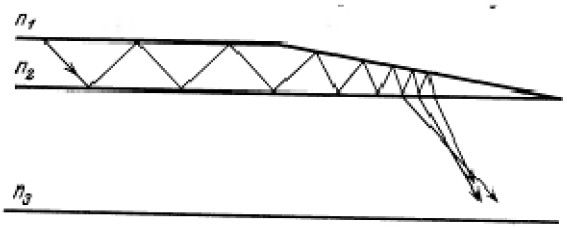


Рисунок 11 – Сужающийся элемент связи

В принципе сужающиеся элементы связи можно также использовать в качестве входных элементов. Однако, чтобы достичь высокой эффективности связи, необходимо сформировать сходящийся входной пучок излучения, который был бы обратным по отношению к расходящемуся пучку, представленному на рисунке 11. Так как это практически неосуществимо, то эффективность использования таких элементов связи на входе волновода крайне низка. Практическое использование сужающихся элементов связи возможно при сочленении тонкопленочных волноводов с оптическим волокном, так как конец волокна можно расположить очень близко к волноводу, и этой торцевой поверхности волокон можно придать форму, необходимую для увеличения эффективности связи. Такой тип использования подробно рассматривается в следующем разделе, посвященном связи волноводов с волокном.

Элементы связи волокно - волновод

Ключевым элементом интегральных оптических систем является элемент связи, передающий оптические волны между волоконным световодом, используемым для передачи сигналов на большие расстояния, и волноводом, входящим в состав ОИС, используемых для обработки сигналов. Исследования последних лет привели к разработке ряда различных видов элементов связи

волокно - волновод. Хотя часть из них обладает достаточной эффективностью, нельзя сказать, что найдено окончательное решение проблемы, в частности более трудной задачи сопряжения с одномодовым оптическим волокном, имеющим малый диаметр светопроводящей жилы. Однако работы в данном направлении продолжаются на обнадеживающем уровне, и не видно серьезных теоретических ограничений для достижения эффективной связи; необходимо преодолеть лишь чисто практические трудности.

Волокно можно непосредственно состыковать с волноводом без какого-либо промежуточного устройства. Если поперечные сечения сердцевины волокна и волновода достаточно близки по размерам, то можно достичь высокой эффективности соединения, как и в случае стыкового соединения канальных волноводов. Для уменьшения потерь на границе раздела можно использовать жидкость с промежуточным показателем преломления. Главная проблема при использовании подхода, основанного на применении стыкового соединения, состоит в том, что крайне трудно достичь и поддерживать правильную юстировку, поскольку сердцевина волокна и волновод, как правило, имеют размеры порядка 1 мкм.

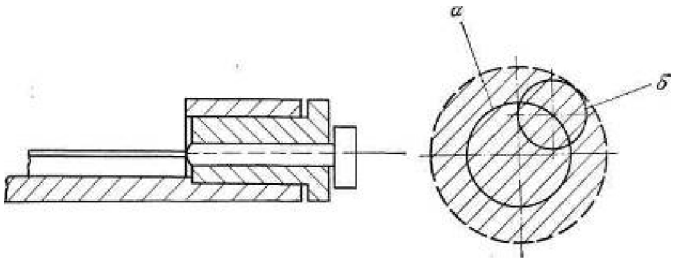


Рисунок 12 - Устройство связи волокно – волновод. Вращением внутренней втулки достигается перемещение волокна по окружности (б), а вращением большой втулки — по окружности (а). Таким образом, волокно можно перемещать по всей заштрихованной площади.

Был предложен механический соединитель, который облегчает юстировку волокна с волноводом. Волокно эксцентрично закрепляется в соединительной втулке, которая в свою очередь монтируется эксцентрично во второй большей

втулке, как показано на рисунке 12. Вращательное движение обеих втулок дает возможность перемещать волокно по значительной площади в пределах соединителя, обеспечивая тем самым его юстировку с волноводом.
Совсем недавно для решения проблемы юстировки канального волновода с одномодовым волокном были использованы методы V-образных канавок в кремнии и монтажа с переворотом пластины кремния, которые уже ряд лет применяются в технологии промышленного изготовления электрических интегральных схем. Шим с сотр. использовали подобный элемент связи для соединения диффузионного волновода на основе $LiNbO_3$ шириной 3 мкм с жилой одномодового волокна диаметром 4,5 мкм. Они достигли эффективности связи 75% как на TE-, так и на TM-модах.

Другой метод стыкового соединения, позволяет избежать трудностей, связанных с юстировкой, так как при этом используется стыковое соединение конца волокна с верхней поверхностью, а не с торцом волновода. Капля жидкости с большим показателем преломления используется для получения получить утечки оптической энергии. Величина этой утечки зависит от угла, однако при соответствующей ориентации волокна была достигнута эффективность связи порядка 50%. Этот метод хорошо себя зарекомендовал в лабораторных условиях, но он мало полезен в прикладных областях интегральной оптики из-за трудностей, связанных с сохранением юстировки.

Другим вариантом является создание канавки к волноводу с помощью установки дисковой резки. Волокно укладывается в канаву и заклеивается в ней. При этом глубина канавки выбирается таким образом, чтобы совместить сердцевину волокна и волновод.

Элементы связи сужающейся пленки с волокном

Можно использовать сужающуюся часть волновода для связи с волоконным световодом, как показано на рисунке 13, а.

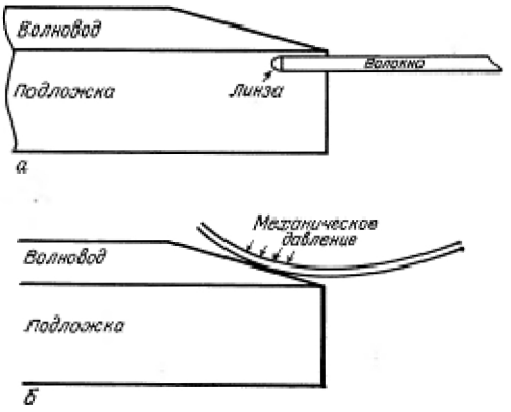


Рисунок 13 - Элементы связи сужающейся пленки с волокном.

а - связь через подложку; б - связь через поверхность.

Волокно вставляется в цилиндрическое отверстие, просверленное в подложке чуть ниже волновода. Полусферическое дно этого отверстия заполняется материалом с высоким показателем преломления, чтобы сформировалась линза. Световые волны, выведенные из тонкопленочного волновода в подложку, собираются волокном. В случае многомодового волоконного световода с достаточно большим значением углов ввода достигнута приемлемая эффективность связи. Например, в одном эксперименте, в котором показатели преломления пленки, подложки, волокна и линзы равнялись 1,53, 1,49, 1,51, 1,55 соответственно, свет длиной волны 632,8 нм от He-Ne лазера был введен с эффективностью 60%.

Другой подход к сопряжению волокно - волновод с использованием сужающейся пленки был продемонстрирован Техом и Стегманом. Вместо того чтобы вставлять волокно в подложку под сужающейся частью волновода, они сняли оболочку волоконного световода, а затем прижали небольшой участок обнаженной жилы к поверхности сужающейся части волновода, как показано на рисунке 13, 6. Так как при этом использовалось многомодовое волокно диаметром 60 мкм, то оно собирало около 90% излучаемого из волновода света.

Решеточные элементы связи для соединения волокно – волновод

Хаммер с сотр. показали, что решеточные элементы связи можно использовать для связи волокна малого показателя преломления с волноводом из материала высокого показателя преломления. Вдавливанием неэкранированного многомодового волокна из пластика (ПММА) ($n \approx 1,5$) в решетку, полученную травлением диффузионного волновода на основе $LiNb_xTa_{1-x}O_3$ ($n=2,195$), они достигли эффективности связи 6%. Однако их расчеты показывают, что при оптимальном «блеске» дифракционной решетки можно достичь эффективности 86%.

Балмер и Уильсон также использовали решетку для связи одномодового волоконного световода со стеклянным волноводом, полученным методом распыления. В этом случае защитная оболочка не удалялась, а волокно нагревалось и вытягивалось для уменьшения поперечного сечения волокна диаметром 18 мкм на одну треть (диаметр светопроводящей жилы 3 мкм). Такое уменьшение толщины оболочки допускает туннелирование оптической энергии, если решетка обеспечивает фазовое согласование. Эффективность составляла лишь 0,4%, однако такая малая величина обусловливалась в основном крайне малой площадью согласования между волокном и используемым широким плоским волноводом. Если провести расчет поправки на рассогласование за счет такой разницы в размерах, то с ее учетом теория предсказывает эффективность связи в 30% для канальных волноводов тех же размеров, что и сердцевины волоконного световода.

Связь между волноводами

Явление оптического туннелирования можно использовать не только для связи волоконного световода или пучка света с волноводом, но и для связи между волноводами. Элементы связи подобного типа, как правило, называются направленными ответвителями, поскольку в данном случае энергия передается когерентно с сохранением направления -распространения. Направленные ответвители изготовляются в виде двух основных геометрических структур: многослойной панарной структуры и двух параллельных канальных волноводов, расположенных достаточно близко друг от друга.

Многослойные плоские волноводные элементы связи

Хотя стыковое соединение можно использовать для двух плоских волноводов, как упоминалось в ранее, более обычным методом является тот, при котором волноводы располагают в непосредственной близости друг от друга и

создают условия, необходимые для связи за счет фазово-когерентного перехода энергии (оптического туннелирования), как показано на рисунке 14.

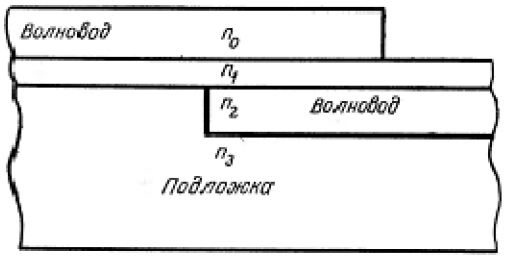


Рисунок 14 - Связи между двумя плоскими волноводами за счет оптического туннелирования.

Переход энергии осуществляется при наличии фазового синхронизма через изолирующий слой с показателем преломления n_1.

При этом показатели преломления n_0 и n_2 волноводных слоев должны быть больше n_1 и n_3, а толщина промежуточного слоя *l* должна быть настолько мала, чтобы перекрывались экспоненциальные «хвосты» мод, распространяющихся в волноводах. Чтобы имел место переход энергии из одного волновода в другой, они должны иметь одинаковые постоянные распространения. Таким образом, показатели преломления и толщины волноводных слоев должны тщательно контролироваться, чтобы обеспечить согласование постоянных распространения. Использование такого труднодоступного метода дает прекрасные результаты. Для достижения оптимального согласования, как и в случае устройств, использующих синхронизм связанных волн, например призменный элемент связи, должна тщательно выбираться длина взаимодействия. Условия для полной передачи энергии снова даются выражением (10), однако значение к для пары перекрывающихся плоских волноводов, очевидно, будет отличаться от значения для призменного элемента связи.

Данный метод связи трудно использовать для волноводов, полученных нанесением тонких пленок, так как такая технология не позволяет столь жестко контролировать получаемые величины показателей преломления и толщины

волноводов. Однако в случае эпитаксиально выращенных волноводов при наличии очень точного контроля толщин и показателей преломления (за счет контроля состава) проблему согласования постоянных распространения мод решить гораздо легче. Очевидно, этот метод особенно хорошо подходит для связи интегрально-оптических лазерных диодов с волноводами. Например, с использованием многослойной планарной структуры для связи CdS_xSe_{1-x} лазера с волноводом, изготовленным из того же материала. В этом случае промежуточный изолирующий слой был образован из GdS. Данный метод также был использован авторами работы для связи лазеров и волноводов на основе GalnAsP, разделенных тонким слоем из InP. Такую структуру сдвоенных волноводов можно использовать для изготовления эффективных одномодовых лазеров с распределенной обратной связью, обеспечиваемой решетками. Можно изготовить лазеры с очень большой эффективностью, если использовать два различных, но связанных друг с другом волновода, чтобы изолировать активную область, в которой идет генерация фотонов от области распределенной обратной связи.

Двухканальные направленные ответвители

Двухканальные направленные ответвители, являющиеся аналогами двухволновых дырочных микроволновых ответвителей, в принципе состоят из параллельных канальных оптических волноводов, расположенных настолько близко друг от друга, что энергия излучения переходит из одного волновода в другой путем оптического туннелирования, как показано на рисунке 15. Такой переход энергии обеспечивается за счет связи в условиях когерентности и синхронизма между перекрывающимися экспоненциальными «хвостами» мод, каналированных в каждом из волноводов. Фотоны моды, распространяющейся в одном из волноводов, например в волноводе 0, переходят в моду, которая распространяется в волноводе 1, сохраняя фазовую когерентность во время такого перехода. Этот процесс происходит кумулятивно на значительной длине;

следовательно, для достижения такой синхронной связи свет должен распространяться с одинаковой фазовой скоростью в каждом канале.

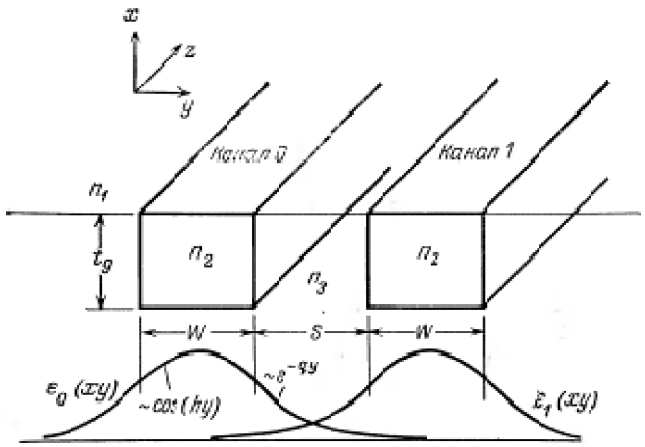


Рисунок 15 - Схема двухканального направленного ответвителя. Внизу представлено характерное для них распределение амплитуд электрической составляющей.

Часть переданной энергии на расстоянии в единицу длины определяется перекрытием мод в отдельных каналах. Поэтому она зависит от расстояния между каналами s, длины взаимодействия L, а также проникновения моды в пространство между каналами, которое можно охарактеризовать коэффициентами ослабления p и q, характерные дл волноводных мод.

В двухканальных ответвителях энергия попеременно переходит из одного волновода в другой и обратно, если длина взаимодействия между ними является достаточно большой. Если бы было проведено измерение плотности распространяющейся оптической энергии в направлении оси z (вдоль одного из каналов направленного ответвителя), то модно было бы наблюдать зависимость синусоидального распределения от расстояния.

Модовые линзы

Ранее было показано, что эффективный показатель преломления волновода увеличивается с увеличением толщины. Изменение эффективного показателя преломления с изменением толщины волновода зависит от порядка моды m.

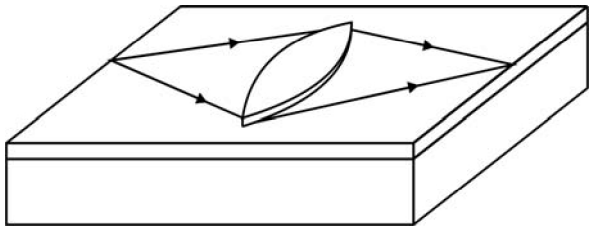


Рисунок 16 - Модовая линза.

Линзы, работа которых основана на изменении n_e с изменением толщины волновода, обладают аберрациями, если в волноводе распространяется более одной моды. В дальнейшем мы будем рассматривать только одномодовые волноводы. Если разность показателей преломления подложки и волновода составляет 1%, толщина одномодового волокна должна быть порядка пяти длин волн. Если длина волны приблизительно равна 1 мкм, изменение эффективного показателя преломления составляет около 0,002 на микрометр изменения толщины. Допустим, мы хотим создать линзу в волноводе, увеличивая толщину волновода, как показано на рисунке 16. Согласно уравнению линзы, необходимый радиус кривизны R равен:

$$R = f(\Delta n / n), \quad (18)$$

где f - фокальное расстояние. Для излучения с длиной волны 1мкм радиус приблизительно равен f*Δd/1000, где Δd измеряется в микрометрах. Это выражение показывает возможность получить линзу с фокусом несколько миллиметров, создавая изменение толщины волновода порядка нескольких микрометров.

Линзы, изготовленные путем размещения одного и более слоев с различным показателем преломления на выпуклый остов, называются модовыми линзами.

По-видимому, они являются простейшими для производства и анализа типом волноводных линз, но имеют два недостатка. Первый недостаток уже был отмечен - наличие аберраций в многомодовых волноводах. К тому же на этих линзах происходят некоторые потери по той причине, что концы волноводов не являются идеально гладкими и, следовательно, имеют некоторые неровности на уровне микрометров. Эти неровности приводят к рассеянию и как следствие к ухудшению контрастности. Тем не менее, из-за простоты изготовления таких линз они нашли применение и в планарной оптике, и в оптических интегральных схемах.

При рассмотрении лазеров с обратной связью отмечалось, что часто предпочтительнее протравливать волновод, а не располагать его на поверхности в виде наружного слоя. Это объясняется высоким показателем преломления большинства материалов, из которых производятся волноводы. Травление волновода снижает показатель преломления. Поэтому положительная линза, изготовленная травлением, должна быть вогнутой, а не выпуклой. Такие линзы более практичны, чем выпуклые линзы, изготовленные нанесением наружного слоя, так как при травлении можно получить большее изменение показателя преломления, чем при нанесении наружного слоя.

Лунебергские линзы

В планарных волноводах можно изготавливать и градиентные линзы. На рисунке 17 приведен простой пример такой линзы. Выпуклая полоска размещается на поверхности волновода. Показатель преломления полоски должен быть больше или равен показателю преломления волновода, поэтому полного внутреннего отражения на границе волновод - полоска нет. Следовательно, верхняя поверхность полоски становится верхней поверхностью волновода. Как уже отмечалось, эффективный показатель преломления волновода увеличивается при увеличении оптической толщины. Поэтому в центре волновода эффективный показатель преломления имеет самое большое значение и постепенно

уменьшается при удалении от центра. Следовательно, полоска фокусирует свет точно так же, как и градиентные волноводы. Участок плоского волновода с нанесенной на поверхность полоской, профиль показателя преломления которой точно рассчитан, становится градиентной линзой. Рассмотренное устройство является разновидностью модовых линз. Они могут быть изготовлены напылением полоски на подложку с использованием масок для контроля пространственного распределения количества напыляемого вещества.

Используемые в интегральной оптике линзы могут обладать радиальной симметрией. Такие линзы называются лунебергскими линзами. Принцип их действия похож на принцип действия одномерных линз. Для расчета точного распределения профиля показателя преломления лунебергских линз используется принцип Ферма. Из-за математической сложности сами расчеты не приводятся, однако необходимо сделать одно важное замечание относительно таких линз.

На рисунке 18 показан вид сверху на лунебергскую линзу. Допустим, линза без искажений формирует изображение некоторой точки A в другой точке B. Линза имеет радиальную симметрию, следовательно, нет выделенных оптических осей.

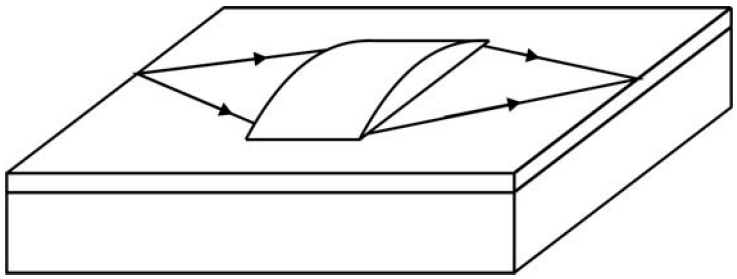


Рисунок 17 - Градиентная линза.

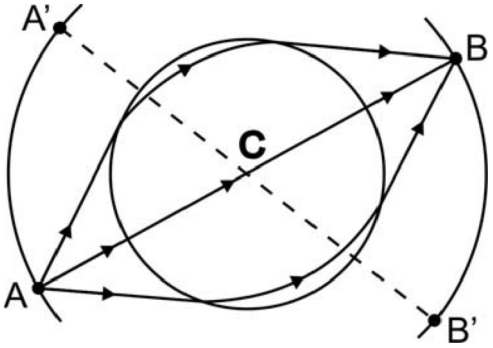


Рисунок 18 – Лунебергская линза (вид сверху)

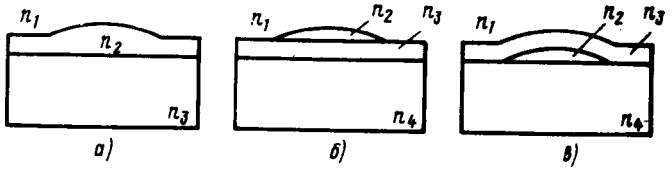


Рисунок 19 – Варианты реализации лунебергских линз.

Любая другая точка A’, лежащая на окружности, проходящей через A и с центром в точке C, эквивалентная точке A. Ее изображение B’ не содержит искажений и эквивалентно B. Поскольку A и A’ лежат на одной окружности, то B и B’ также находятся на одной окружности. Лунбургская линза является примером совершенной оптической системы. За исключением того, что объект и его изображение лежат на окружности, правильно сконструированная линза не вносит искажений. Варианты конструкций подобных линз приведены на рисунке 19.

Геодезическая линза

Лунебергские линзы относятся к классу модовых линз. Они лучше всего работают в одномодовых волноводах или в волноводах с несколькими модами низшего порядка. Линзы другого вида называются геодезическими, т.к. лучи в них распространяются по кратчайшему пути между двумя точками. На рисунке 20 показаны вид сверху и вид сбоку на геодезическую линзу в волноводе. Линза

образована сферическим углублением в подложке. Волновод искривляется, повторяя форму углубления, сохраняя постоянную толщину. Луч света, распространяющийся по диаметру углубления, проходит больший путь, чем луч, распространяющийся касательно к линзе. Соответственно, длины путей лучей, пересекающих линзу по другим направлениям, принимают промежуточные значения. Т.к. фокусировка определяется равенством проходимых оптических путей, то геодезическая линзы принципиально могут фокусировать свет.

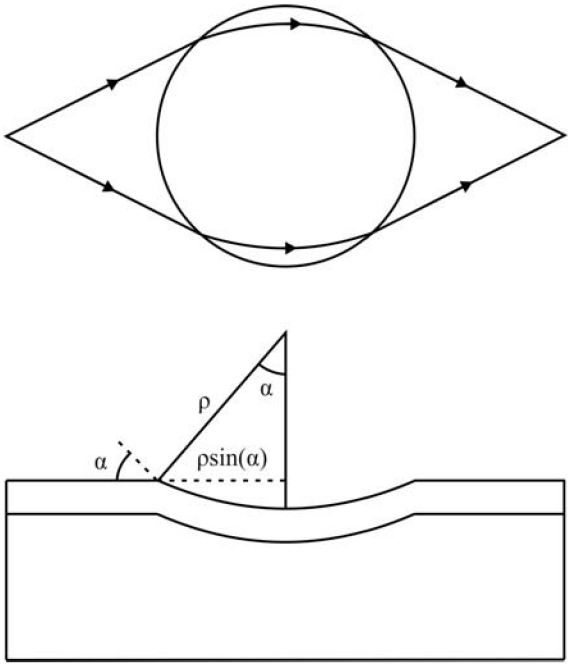


Рисунок 20 - Геодезическая линза.

Фокусное расстояние сферической геодезической линзы в параксиальном приближении можно определить следующим образом. Обозначим радиус кривизны углубления ρ, тогда радиус углубления в плоскости волновода равен ρ*sinα. где α - угол, показанный на рисунке 20. Падающий на геодезическую линзу луч, проходя вдоль диаметра, распространяется под углом α к оптической оси. Таким образом, в параксиальном приближении геодезическая линза может рассматриваться как толстая двояковыпуклая линза с радиусом кривизны ρ*sinα, поверхности которой разделены расстоянием, вдвое превышающим это значение, а относительный показатель преломления равен $n_e'/n_e=1/\cos(\alpha)$.

Фокусное расстояние такой линзы можно вычистить. В рассматриваемом случае оба радиуса кривизны равны R расстояние между поверхностями линзы равно 2R. следовательно:

$$f = R/2(1-1/n). \quad (19)$$

Полученный результат тот же, что и в случае сферической линзы с показателем преломления n. В случае геодезической линзы n следует заменить на 1/cosα, R равно ρsina. Фокус сферической геодезической линзы равен:

$$f = \rho \sin(\alpha) / 2(1 - \cos(\alpha)). \quad (20)$$

На практике линзы с большими апертурами обладают сферической аберрацией. К тому же при больших α на краях линзы имеют место значительные потери. Реальные геодезические линзы должны иметь клиновидную форму для уменьшения потерь на краях, а следовательно, отличный от сферического профиль углубления.

Геодезические линзы могут быть изготовлены путем формирования углубления на поверхности подложки и создания поверх этого углубления волновода. Оптический путь вдоль диаметра углубления больше, чем вдоль касательной, и само углубление играет роль фокусирующей линзы. Более легко контролируемый способ производства геодезической линзы заключается в автоматизированном изготовлении углубления при помощи токарного станка с алмазным резцом, управляемого компьютером. Иногда этот процесс называют алмазным профилированием, и он может быть использован для прецизионного изготовления углублений любой формы.

Еще одно преимущество геодезической линзы состоит в том. что ее фокус не зависит от порядка моды, так как от порядка моды не зависит отношение n_e'/n_e. Следовательно, такая линза может использоваться в многомодовых волноводах, не внося искажений, связанных с распространением мод с разными индексами.

Дифракционные решетки

Дифракционные решетки для устройств различного назначения создаются травлением поверхности пленарного волновода или нанесением на нее слоев материала. Дифракционная линза эквивалентна одномерной зонной пластинке

Френеля, нанесенной на поверхность волновода или вытравленной внутри него.. В отличие от зонных пластинок, состоящих из чередующихся прозрачных и поглощающих участков, соседние полупериодные зоны дифракционной линзы вносят сдвиг фазы падающей волны, отличающийся на π. Это изменение фазы точно компенсирует разность хода лучей от различных зон, поэтому все зоны вносят вклад в общую интенсивность изображения. В обычных зонных пластинках вклад вносит только половина зон. В результате интенсивность изображения, формируемого фазовыми пластинками, в четыре раза больше, чем интенсивность изображения, формируемого обычными полупериодными пластинками.

На рисунке 21 схематично показан вид сверху на тонкую дифракционную линзу. Зоны могут быть получены, например, созданием в волноводе слоев с меньшим эффективным показателем преломления в каждой зоне путем травления. Длина волны в волноводе равна длине волны в вакууме, деленной на эффективный показатель преломления, поэтому для расчета точных значений параметров линзы необходимо знание эффективного показателя преломления волновода. Для повышения эффективности дифракционные линзы можно выполнять не только дискретными, но и аналоговыми (рисунок 22).

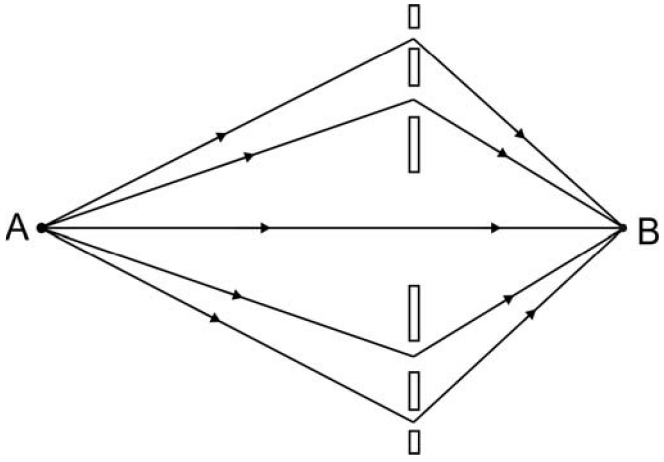
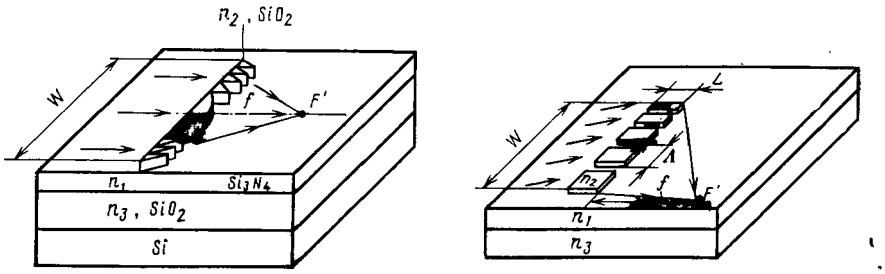


Рисунок 21 – Волноводная дифракционная решетка (вид сверху).



а) б)

Рисунок 22 – Тонкопленочная аналоговая линза Френеля (а) и апериодическая дискретная линза Френеля (б)

Рассмотрим схему волноводной решетки, приведенную на рисунке 21. Решетка может быть получена травлением или напылением на поверхность волновода, либо решетка может быть объемной, полученной периодическим легированием материала волноводной пленки. Сначала рассмотрим моду подложки, или волну, дифрагировавшую в подложку:

$$m\lambda = n_s d \sin(i_s) - n_g d \sin(i_g), \qquad (21)$$

где индекс s относится к подложке, g - к волноводному слою. Символ m здесь означает порядок дифракции. Предположим, что волновод поддерживает распространение только одной моды. Тогда будет справедливо равенство:

$$m\lambda = n_s d \sin(i_s) - n_e d\,. \qquad (22)$$

Из сравнения (21) и (22) следует, что волна в волноводе может быть описана одним из двух эквивалентных способов: либо как волна, распространяющаяся под углом i_g к плоскости волновода в материале с показателем преломления n_g, либо как волна, распространяющаяся параллельно поверхности волновода, но в материале с показателем преломления, равным эффективному показателю преломления n_e.

Так как n_e в (22) всегда меньше n_s, величина m должна быть меньше нуля. Это означает, что могут существовать только отрицательные порядки дифракции. Следовательно, для наименьшего порядка дифракции (m = -1):

$$n_s \sin(i_s) = n_e - \lambda / d . \qquad (23)$$

Аналогично для воздуха:

$$\sin(i_a) = n_e - \lambda / d , \qquad (24)$$

где индекс а относится к воздуху и $n_a = 1$.

Определив n_e из (23) и подставив в (24), получим:

$$\sin(i_s) = \sin(i_a) / n_s . \quad (25)$$

Так как n_s всегда больше 1, $\sin(i_s)$ меньше 1 тогда, когда $\sin(i_a)$ меньше 1, хотя обратное не всегда верно. Следовательно, если существует волна, дифрагировавшая в воздух, также существует волна, дифрагировавшая в подложку.

Для увеличения эффективности дифракции следует ограничить дифракционную картину только одним порядком дифракции. Можно, например, подавить волну, дифрагировавшую в воздух, или воздушную моду, потребовав, чтобы $\sin(i_a)$ в (24) был меньше -1. Тогда, согласно уравнению (25), $i_s < 0$, дифрагировавшая волна распространяется в обратном направлении. В зависимости от угла дифракции эта волна либо может оказаться локализованной в подложке из-за полного внутреннего отражения, либо может выйти из подложки и распространяться практически параллельно поверхности. Это наиболее часто наблюдается в материалах с высоким показателем преломления, например в арсениде галлия. Поэтому может возникнуть необходимость в использовании призмы, контактирующей с нижней поверхностью подложки, для вывода света.

Поляризационные устройства

Устройства для модуляции, разделения или контроля поляризации могут найти широкое применение в перспективных системах оптической связи когерентного типа. Все эти устройства были разработаны на основе волноводов Ti:$LiNbO_3$. Для определения поляризации излучения могут быть использованы различные базисы параметров, однако для наших целей удобно использовать две характеристики угол поляризации Θ и фазовый сдвиг φ. Тогда нормированные комплексные амплитуды для TE и TM мод можно записать в виде:

$$\begin{pmatrix} A_{TE} \\ A_{TM} \end{pmatrix} = \begin{pmatrix} \cos(\Theta) \\ \sin(\Theta) e^{j\phi} \end{pmatrix}, \quad (26)$$

где Θ определяет отношение амплитуд TE и TM волн, а φ есть фазовый сдвиг между TE и TM компонентами поля. Свет линейно поляризован под углом Θ, если φ = 0. В случае Θ = 0 присутствует только TE поляризация, а при Θ = π/2 - только TM. Свет с правой круговой поляризацией будет, например, определяться следующими параметрами Θ = π/4 и φ = π/2. В пассивных волноводах Ti:$LiNbO_3$ свет линейно-поляризованный вдоль главных осей, распространяется так, что его поляризация вдоль главных осей сохраняется. Например, излучение TE (TM) поляризации в волноводе, изготовленном на z (x или y) срезе ниобата лития, сохраняет исходную поляризацию. Кроме того, при вводе эллиптически поляризованного излучения, соотношение амплитуд TE и TM мод на выходе сохранится, хотя, если волноводы обладают двулучепреломлением, как в случае Ti:$LiNbO_3$, относительный сдвиг фазы может измениться

TE ⇔ TM преобразование

Устройство, использующее наибольший электрооптический коэффициент для преобразования ТЕ волн в ТМ, изображено на рисунке 23.

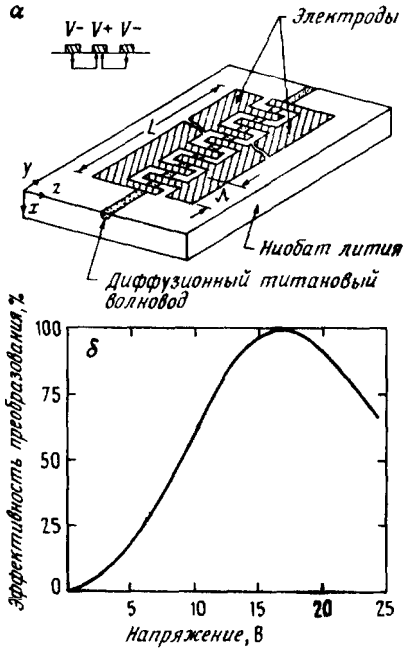


Рисунок 23 – волноводный электрооптический преобразователь поляризации ТЕ ⇔ ТМ с синхронизированными периодическими электродами (а), экспериментальная зависимость эффективности преобразования поляризации ТЕ ⇔ ТМ в синхронном режиме от приложенного напряжения, L = 6 мм и λ = 1,32 мкм (б).

Для этого устройства основным является недиагональный элемент электро-оптического тензора r_{51}, позволяющий преобразовывать TE (TM) волны в TM (TE) Периодическая электродная структура необходима ввиду двулучепреломления ниобата лития и различия в эффективных показателях преломления для мод различной поляризации. Эффективное преобразование может быть реализовано для периода Λ, определенного выражением:

$$\frac{2\pi}{\lambda}\left|N_{TM} \quad N_{TE}\right| = \frac{2\pi}{\Lambda}. \qquad (27)$$

В случае ниобата лития требуемые значения периода Λ составляют 7 и 18 мкм соответственно при λ = 0,6 и 1,3 мкм. Выше мы показали, что этот процесс преобразования описывается уравнениями связанных мод при Δβ=0. Таким образом:

$$\eta = \sin^2 kL, \qquad (28)$$

где

$$k = \frac{\Gamma\pi}{\lambda} n^3 r_{51} \frac{V}{G}. \qquad (29)$$

Как следует из структуры электрооптического тензора для $LiNbO_3$, направленное вдоль оси x электрическое поле может приводить к связи поляризованных по осям x и z компонент оптического поля через коэффициент r_{51}. Для достижения 100% модуляции поляризации необходимо, чтобы kL = π/2, что является обязательным условием для любого типа модулятора. Экспериментальные характеристики TE ⇔TM преобразователя мод, в котором обеспечен синхронизм для длины волны λ = 1,32 мкм, приведены на 23.

Вращатель плоскости поляризации

Для ряда практических применений требуется устройство, способное преобразовать данную поляризацию входного сигнала с поляризацией (Θ_1, φ_1) в определенную поляризацию сигнала на выходе. Устройства для преобразования поляризации играет важную роль, поскольку обычное одномодовое волокно не сохраняет поляризацию, в то время как когерентные системы часто требуют известного состояния поляризации. Во многих случаях достаточно, чтобы выходной сигнал представлял собой ТМ или ТЕ моду. При этом на вход устройства из стандартного (без двулучепреломления) одномодового волокна может поступать произвольный, в общем случае эллиптически поляризованный сигнал. Для обеспечения произвольного преобразования поляризации требуется более сложное устройство, чем ТЕ <-> ТМ преобразователь. Наиболее простой вариант такого устройства изображен на рисунке 24.

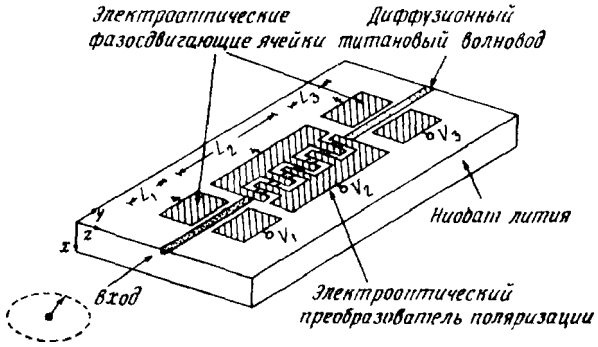


Рисунок 24 – Волноводный электрооптический преобразователь поляризации ТЕ <-> ТМ с синхронизованными периодическими электродами (а) и экспериментальная зависимость эффективности преобразования поляризации ТЕ <-> ТМ в синхронном режиме от приложенного напряжения, L=6 мм, λ=1,32 мкм.

Оно состоит из синхронизованного ТМ <-> ТЕ преобразователя, расположенного между двумя фазосдвигающими участками, которые при подаче

электрического напряжения изменяют сдвиг фазы между TE и TM модами. Для любого входного сигнала (Θ_1, φ_1) путем перемножения трех характеристических матриц участков устройства можно найти амплитуды TE и TM мод на выходе:

$$\begin{pmatrix} A_{TE} \\ A_{TM} \end{pmatrix} = \begin{pmatrix} \cos(kL_2)\cos(\Theta_1) - je^{j(\phi_1+\Delta\phi_1)}\sin(kL_2)\sin(\Theta_1) \\ e^{j\Delta\phi_3}\left(e^{j(\phi_1+\Delta\phi_1)}\cos(kL_2)\sin(\Theta_1) - j\sin(kL_2)\cos(\Theta_1)\right) \end{pmatrix}. \quad (30)$$

Лекция 11

Интегрально-оптические модуляторы и затворы. Основные технические параметры и характеристики. Электрооптические модуляторы Маха-Цандера, бегущей волны, на связанных волноводах, дифракционные, поглощающие. Акустооптические модуляторы и дефлекторы.

Электрооптические модуляторы

Во многих случаях один и тот же прибор действует как модулятор или как переключатель в зависимости от силы взаимодействия между оптическими волнами и управляющего электрического сигнала, а также от структуры входа и выхода. Считают, что прибор является модулятором, если его основная функция состоит в вводе информации в несущую волну света за счет временного изменения одной из его характеристик. С другой стороны, переключатель изменяет пространственное положение света, иначе говоря, включает и выключает его. При проектировании или оценке как модуляторов, так и переключателей необходимо учитывать те же самые факторы. Поэтому логично рассматривать их совместно.

Основные рабочие характеристики модуляторов и переключателей

Глубина модуляции

Одной из самых важных характеристик модуляторов и переключателей является глубина или степень модуляции η. В случае модуляции по интенсивности, когда приложенный электрический сигнал действует с целью уменьшения интенсивности проходящего света, величина η дается выражением:

$$\eta = (I_0 - I)/I_0, \qquad (1)$$

где I - интенсивность прошедшего света, I_0 - значение интенсивности в отсутствие электрического сигнала. Если приложенный электрический сигнал действует с целью увеличения интенсивности проходящего света, то величина η дается выражжением:

$$\eta = (I - I_0)/I_m, \qquad (2)$$

где I_m - интенсивность проходящего света в случае приложения максимального электрического сигнала. Максимальная глубина модуляции, или коэффициент ослабления, определяется формулой:

$$\eta_{макс} = (I_0 - I_m)/I_0, \quad при \quad I_m \le I_0, \qquad (3)$$

$$\eta_{макс} = (I_m - I_0)/I_m, \quad при \quad I_m \ge I_0,$$

Можно также определить глубину модуляции для фазовых модуляторов, поскольку изменение фазы можно связать с изменением интенсивности в виде функции. Для случая интерференционных модуляторов можно показать, что глубина модулялии описывается:

$$\eta = \sin^2(\Delta\varphi/2), \qquad (4)$$

где Δφ - изменение фазы. Глубину модуляции можно определить для модуляторов интенсивности (и косвенно для фазовых модуляторов); однако аналогичный критерий качества, или максимальная девиация, для частотных модуляторов определяется выражением:

$$D_{макс} = |f_m - f_0|/f_0, \qquad (5)$$

где f_0 - оптическая несущая частота, а f_m - смещенная оптическая частота, соответствующая приложению максимального электрического сигнала.

Полоса пропускания

Другой важной характеристикой модуляторов и переключателей является полоса пропускания или диапазон частот модуляции, в котором может работать прибор. Обычно считают, что полоса пропускания модулятора определяется разностью между верхней и нижней частотами, при которых глубина модуляции уменьшается на 50% ее максимальной величины. Для переключателей частотную характеристику обычно определяют в виде скорости или времени переключения. Время T переключения связано с полосой Δf пропускания следующим образом:

$$T = 2\pi / \Delta f, \qquad (6)$$

Минимизация времени переключения модуляторов приобретает очень важное значение при использовании матриц с большим числом переключателей для распределения света по нужным направлениям. Аналогично ширина полосы модуляторов является определяющим параметром в тех случаях, когда несколько информационных каналов вводятся параллельно в один и тот же оптический луч. Поэтому-то высокая скорость переключения и широкая полоса волноводных переключателей и модуляторов, о чем будет сказано ниже, делают их особенно перспективными для использования в больших телекоммуникационных системах.

Потери, вносимые модуляторами и переключателями

Другой важной характеристикой оптических модуляторов и переключателей, которую необходимо знать при проектировании систем, являются потери, вносимые этими приборами. Такие потери, как правило,

выражают в децибелах, и в случае, когда модулирующий сигнал приводит к уменьшению интенсивности, они задаются выражением:

$$L_1 = 10\lg(I_1 / I_0), \qquad (7)$$

где I_1 - интенсивность света, которая передавалась бы волноводом в отсутствие модулятора, а I_0 - передаваемая интенсивность при наличии модулятора, но без приложенного электрического сигнала. Для модулятора, в котором при приложении электрического сигнала увеличивается интенсивность проходящего света, вводимые потери определяются формулой:

$$L_2 = 10\lg(I_1 / I_m), \qquad (8)$$

где I_m - передаваемая интенсивность, соответствующая приложению максимального электрического сигнала. Вносимые потери, конечно, являются оптическими. Однако в конце концов они приводят к увеличению электрической энергии, подводимой к системе, так как требуется использовать оптические источники большей мощности.

Потребляемая мощность

Для обеспечения работы модуляторов и переключателей необходимо, конечно, затрачивать электрическую энергию. В случае модуляторов потребляемая электрическая мощность увеличивается с увеличением частоты модуляции. Поэтому характерным показателем является потребляемая мощность на единицу ширины полосы модулятора $P/\Delta f$, которая обычно выражается в милливаттах на мегагерц. Например, основным преимуществом модуляторов канальных волноводов перед объемными модуляторами является: гораздо меньшее значение отношения $P/\Delta f$.

Мощность, потребляемую при работе оптических переключателей, работающих при больших скоростях переключения сигналов, например в случае временного уплотнения и разделения различных сигналов, можно оценить во многом сходным способом, как это делают в случае модуляторов. Поэтому отношение $P/\Delta f$ является полезным критерием. Однако, если переключение осуществляется при сравнительно низких скоростях, более существенной величиной является мощность,, затрачиваемая для удержания переключателя в определенном рабочем положении. В идеале переключатель может потреблять, мощность только во время смены состояния; мощность, затрачиваемая в нем на поддержание состояния, должна быть незначительной. Так как для электрооптических переключателей необходимо присутствие электрического поля для поддержания хотя бы одного положения, то с этой точки зрения их нельзя считать идеальными. Однако, исключая ток утечки, для поддержания электрического поля в малом объеме волноводного переключателя требуется малая мощность.

Изоляция

Большое внимание при проектировании систем уделяется изоляции между различными входами и выходами переключателя или модулятора. В модуляторе изоляция между входом и выходом является просто максимальной глубиной модуляции, как было определено выше. Однако ее обычно выражают в децибелах. В случае переключателя изоляция между двумя каналами (либо входом, либо выходом) задается выражением

$$Изоляция\ [дБ] = 10\lg(I_2 / I_1), \qquad (9)$$

где I_1 — интенсивность света в действующем канале, а I_2 — интенсивность света в отключенном канале. Таким образом, переключатель с утечкой или

перекрестным действием сигнала величиной 1% по отношению к этим двум каналам будет иметь изоляцию -20 дБ.

Электрооптический эффект

Явление, на котором основано действие большинства электрооптических модуляторов и переключателей, связано с изменением показателя преломления вещества под действием приложенного электрического поля. В общем случае этот эффект является анизотропным и имеет как линейную (эффект Поккельса), так и нелинейную (эффект Керра) составляющие. В твердых кристаллических веществах изменение показателя преломления, обусловленное линейным электрооптическим эффектом, можно очень удобно характеризовать изменением компонент оптической индикатрисной матрицы. Уравнение эллипсоида показателей преломления в присутствии электрического поля имеет вид:

$$\left(1/n^2\right)_1 x^2 + \left(1/n^2\right)_2 y^2 + \left(1/n^2\right)_3 z^2 + 2yz\left(1/n^2\right)_4 + 2xz\left(1/n^2\right)_5 + 2yx\left(1/n^2\right)_6 = 1, \qquad (10)$$

Если оси x, y и z выбрать параллельно главным осям кристалла, линейное изменение коэффициентов, обусловленное введением электрического поля E, задается следующим образом:

$$\Delta\left(1/n^2\right)_i = \sum_{i=1}^{3} r_{ij} E_j , \qquad (11)$$

где i=l, 2, …, 6, j= 1, 2, 3 связаны с x, y и z соответственно. Если выражение (11) записать в матричной форме, то такая 6x3 матрица называется электрооптическим тензором. Можно показать, что линейный электрооптический эффект существует лишь в кристаллах, не обладающих инверсной симметрией. Даже в случае нецентросимметричных кристаллов для большинства классов симметрии лишь несколько элементов электрооптического тензора отлично от нуля. Поэтому при

проектировании электрооптических модуляторов и переключателей необходимо тщательно выбирать как тип материала для волноводов, так и его ориентацию относительно приложенного электрического поля. Тем не менее большинство материалов, на основе которых можно изготовлять волноводы с малыми оптическими потерями, например GaAs, GaP, $LiNbO_3$, $LiTaO_3$ и кварц, при определенной ориентации также имеет достаточно большие коэффициенты Поккельса. Таким образом, в различных областях применения интегральной оптики широко используется линейный электрооптический эффект. Нелинейный (квадратичный) электрооптический эффект Керра выражен сравнительно слабо у большинства материалов, используемых для изготовления волноводов. Нелинейная зависимость от электрического поля также вводит нежелательное искажение в модулированный сигнал. Следовательно, этот факт ограничивает использование данного эффекта во многих случаях применения интегральной оптики.

Одноволноводные электрооптические модуляторы

На основе одноволноводных структур можно изготовить несколько различных типов электрооптических модуляторов и переключателей. При этом могут использоваться как плоские, так и канальные волноводы. Примером может служить относительно простой плоский волновод, представленный на рисунке 1, который может функционировать как фазовый модулятор, амплитудный (по интенсивности) модулятор или как оптический переключатель. Волновод, показанный на рисунке 1, сформирован на основе $Ga_{1-x}Al_xAs$ с использованием гетероэпитаксиальной технологии. Однако можно использовать любой электрооптический полупроводник, например $GaAs_xP_{1-x}$, GaAs или GaP; с другой стороны, волновод можно изготовить на основе метода, связанного с уменьшением концентрации носителей.

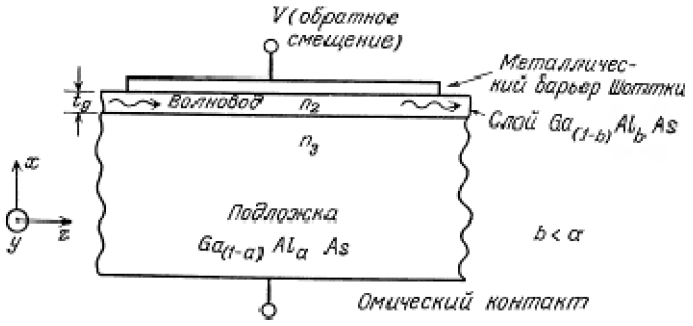


Рисунок 1 - Основная структура электрооптического модулятора (вид сбоку поперечного сечения пластины <100>).

Фазовая модуляция

Волновод, изображенный на рисунке 1, асимметричен, так как показатели преломления воздуха и металла, находящегося на верхней границе, много меньше n_2, в то время как n_3 относительно ближе к n_2. Полное изменение величины показателя преломления между волноводом и подложкой дается выражением:

$$\Delta n_{23} = n_2 - n_3 = \Delta n_{хим} + \Delta n_{укн} + \Delta n_{эо}, \qquad (12)$$

где $\Delta n_{хим}$ - изменение показателя преломления, обусловленное различными концентрациями а и б алюминия; $\Delta n_{укн}$ - изменение, связанное с уменьшением концентрации носителей (если таковое вообще имеет место) в волноводном слое; $\Delta n_{эо}$ - изменение показателя преломления, вызванное электрооптическим эффектом. Чтобы реализовать фазовый модулятор, размеры и легирование волновода выбирают так, чтобы он находился выше отсечки для моды m=0 на желаемой длине волны, но ниже отсечки для моды m=1. Таким образом, имеем:

$$\Delta n_{хим} + \Delta n_{укн} > 1/32 n_2 (\lambda_0 / t_g)^2 < 9/32 n_2 (\lambda_0 / t_g)^2. \qquad (13)$$

При подаче напряжения V обратной полярности на диод на барьере Шоттки, как показано на рисунке 1, волновод становится частью обедненного слоя этого диода и электрическое поле вызывает изменение фазы световой волны, проходящей по волноводу, которое пропорционально величине V. Для указанной ориентации кристалла изменение показателя преломления, вызванное электрическим полем, для ТЕ-волны (поляризованной в направлении оси y) описывается выражением:

$$\Delta n_{эо} = n_2^3 r_{41} V / 2t_g , \qquad (14)$$

тогда как аналогичного изменения показателя преломления для ТМ-волн не происходит (ТМ-волны поляризованы в направлении оси x). По определению имеем:

$$\Delta n = \Delta\beta / k = \Delta\beta\lambda_0 / 2\pi . \qquad (15)$$

Поэтому если выражение (15) подставить в выражение (14), то изменение фазы, вызываемое электрическим полем, можно определить из выражения:

$$\Delta\varphi_{эо} = \Delta\beta L = (\pi / \lambda_0) n_2^3 r_{41} VL / t_g, \qquad (16)$$

где L - длина модулятора в направлении z. Было продемонстрировано несколько вариантов базовой структуры одноволноводных фазовых модуляторов. Например, Холл с сотр. изготовили плоский модулятор, подобный показанному на рисунке 1, на основе GaAs, у которого $\Delta n_{хим} = 0$, но изменение показателя преломления за счет легирования в слоях $\Delta n_{укн}$ было достаточным для обеспечения светопроводности. Каминов с сотр. изготовили модулятор упомянутого типа на основе автодиффузионного плоского волновода из $LiNbO_3$ с потребляемой мощностью 0,4 мВт/МГц для достижения изменения фазы в 1 рад при длине волны излучения 632,8 нм. При использовании ионно-лучевого травления для получения гребенчатых канальных волноводов шириной 19 мкм в модуляторе на основе $LiNbO_3$ потребление мощности удалось снизить до 20 мкВт/МГц/рад. В этом случае электрическое поле вводилось в волновод с помощью двух электродов в форме металлических полос с каждой стороны волновода. Напряжение величиной 1,2 В вызывало изменение фазы в 1 рад. Дальнейшее уменьшение потребляемой мощности до 1,7 мкВт/МГц/рад было достигнуто при использовании одноканального волноводного модулятора на основе диффузии Ti в $LiNbO_3$ шириной 5 мкм.

Поляризационные модуляторы

Использование фазовых модуляторов ограничено тем фактом, что совместно с ними должны использоваться фазово-когерентные системы детектирования. Чтобы избежать такого усложнения, модулятор, схема которого представлена на рисунке 1, можно использовать в несколько измененном виде, вводя оптический пучок, плоскость поляризации которого составляет угол 45° с

осями x и y. Так как изменение фазы имеет место только для волн поляризованных в направлении оси y, а не для волн, поляризованных в направлении оси x то при распространении волн в направлении оси z будет иметь место вращение вектора поляризации. Такое изменение поляризации можно обнаружить, используя поляризационно чувствительный детектор или помещая перед детектором селективный поляризационный фильтр (обычно называемый анализатором). В случае дискретного волноводного модулятора, который используется для пучков света в воздухе, в качестве анализатора можно воспользоваться обычным решеточным поляризатором или абсорбционным поляризационным фильтром. Аналогичную систему можно использовать для интегральной оптической схемы. Например, как призменные, так и решеточные элементы связи чувствительны к поляризации, и поэтому их можно применять в качестве анализаторов. Однако использование поляризационных модуляторов в оптических интегральных схемах ограничивается трудностями изготовления эффективных монолитных анализаторов что привело к предпочтительному использованию модуляции интенсивности.

Модуляция интенсивности

Так как поляризационную и фазовую модуляции обнаружить 'труднее, чем модуляцию интенсивности, то устройство, схема которого представлена на рисунке 1, очень часто используется в режиме модуляции интенсивности. Для осуществления модуляции интенсивности необходимо тщательно подобрать разность показателей преломления на границе раздела волновод - подложка, чтобы в отсутствие электрического поля волновод находился бы точно в припороговых условиях для пропускания моды наинизшего порядка. Затем при прикладывании электрического поля за счет подачи напряжения на электроды оно вызывает небольшое дополнительное изменение показателя преломления, которое приводит к тому, что волновод становится проводящим. Полное изменение показателя преломления дается выражением (8.3.1). Следовательно, пороговые

условия при нулевом электрическом поле для модуляторов интенсивности такого типа описываются выражением

$$\Delta n_{23} = \Delta n_{хим} + \Delta n_{укн} = 1/32 n_2 \left(\lambda_0 / t_g\right)^2, \qquad (16)$$

где было использовано выражение для условий отсечки в случае асимметричных волноводов. Модуляторы интенсивности такого типа были впервые разработаны Холлом с сотр. на основе плоских волноводов из GaAs, полученных методом уменьшения концентрации носителей для излучения с длиной волны 1,15 мкм. Для перевода волновода в состояние выше отсечки на TE_0-моде требовалось напряжение 130 В. Кэмпбелл с сотр. изготовили модулятор такого типа на основе канальных волноводов из GaAs, максимальная глубина модуляции которого составляла 95%, полоса пропускания 150 МГц, а потребляемая мощность менее 300 мкВт/МГц. Совсем недавно Кавабе с сотр. изготовили модулятор интенсивности для света с длиной волны 632,8 нм на основе канальных волноводов шириной 2,4 мкм, полученных диффузией Ti в $LiNbO_3$. При подаче напряжения -10 В (15 кВ/см) мода E^z_{11} хорошо проходила, а при напряжении +10 В - отсекалась. Коэффициент ослабления между двумя экстремумами составил -19 дБ. Очевидно, что модуляторы интенсивности, подобные написанным выше, кроме того, могут использоваться в качестве эффективных оптических переключателей, поскольку величина ($\Delta n_{хим} + \Delta n_{укн}$) выбирается такой, чтобы поддерживать волновод в припороговом состоянии, а приложенное поле V/t_g является достаточно большим, чтобы перевести его в проходящее состояние на нужной длине волны.

Электроабсорбционная модуляция

Работа модуляторов, рассмотренных до сих пор, основывалась на использовании линейного электрооптического эффекта. Однако имеются и модуляторы другого типа, электроабсорбционные, которые должны

классифицироваться как электрооптические, так как для появления модуляции интенсивности излучения в них используется электрическое поле, но не применяется эффект Поккельса. Вместо этого в электроабсорбционных модуляторах используется эффект Франца-Келдыша. При подаче сильного электрического поля граница полосы собственного поглощения в полупроводниках смещается в сторону больших длин волн, как показано на рисунке 2 для GaAs при подаче поля напряженностью $1,3*10^5$ В/см.

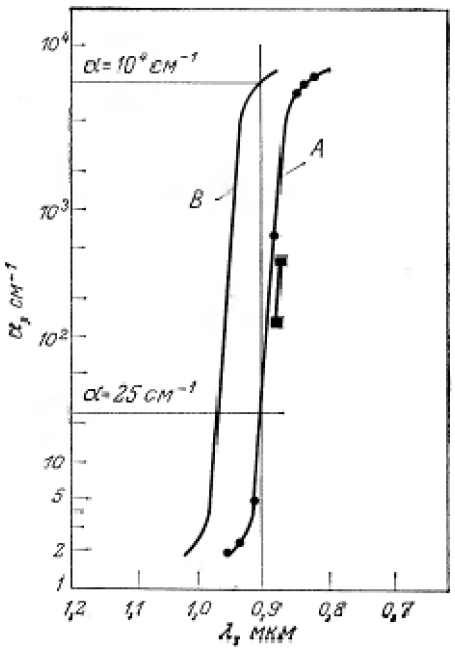


Рисунок 2 – Сдвиг границы поглощения для GaAs за счет эффекта Франца-Келдыша: А – кривая поглощения

Благодаря наличию резкой границы полосы поглощения у материалов с прямой запрещенной зоной, например GaAs, приложение электрического поля может привести к сильному изменению поглощения длин волн вблизи границы полосы. В частном случае, приведенном на рисунке 2 для излучения с длиной волны 900 нм, при подаче электрического напряжения коэффициент поглощения а увеличивается в пределах 25 - 10^4 $см^{-1}$.

Механизм, ответственный за появление эффекта Франца-Келдыша, можно описать непосредственно, обращаясь к картине энергетических полос полупроводника, представленной на рисунке 3.

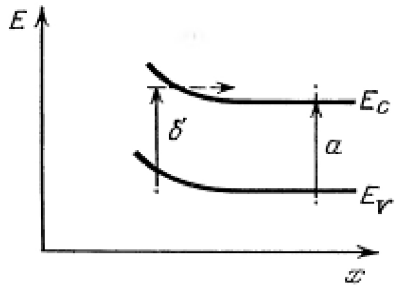


Рисунок 3 - Картина энергетических полос (зон) полупроводника, иллюстрирующая возникновение эффекта Франца-Келдыша в присутствии сильного электрического поля.

x - расстояние, отсчитываемое от поверхности полупроводника; E - энергия электрона:, Ec и Ev - границы зоны проводимости и валентной зоны соответственно.

При подаче сильного электрического поля границы зон изгибаются. Левая часть рисунке 3 соответствует поверхности полупроводника, в которой формируется контакт типа барьера Шоттки либо неглубокий p-n переход. Приложение напряжения обратного смещения на этот выпрямляющий переход вызывает образование обедненного носителями слоя, распространяющегося вглубь проводника на расстояние x. Внутри этого слоя; существует неоднородное электрическое поле, причем его амплитуда достигает наибольшего значения на поверхности полупроводника. Поэтому наибольший изгиб зон также имеет место на поверхности. Вне обедненного слоя, в котором существует электрическое поле, зоны плоские, чему соответствует правая часть рисунке 3. В этой области фотон может быть поглощен только в том-случае, если он обладает энергией, достаточной для переброса электрона через запрещенную зону, как в переходе а. Ближе к поверхности, где зоны изгибаются под действием поля, может иметь место переход б, когда энергия фотона достаточна лишь для компенсации части энергии запрещенной зоны для электрона. В обычных условиях подобный

переход может быть запрещен, так как в запрещенной зоне нет разрешенных состояний. Однако приложенное электрическое поле существенна уширяет разрешенные состояния зоны проводимости, поэтому имеется определенная вероятность нахождения электрона в за-лрещенной зоне. Этот факт, конечно, снижает эффективную ширину запрещенной зоны и тем самым приводит к сдвигу границы полосы в сторону больших длин волн. Можно показать, что эффективное изменение ширины запрещенной зоны ΔE дается выражением:

$$\Delta E = 3/2(m^*)^{-1/3}\left(\frac{qhE}{2\pi}\right)^{2/3}, \qquad (17)$$

где m* — эффективная масса носителей, q - величина электрического заряда носителей, h - постоянная Планка, деленная на 2π, E - напряженность электрического поля.

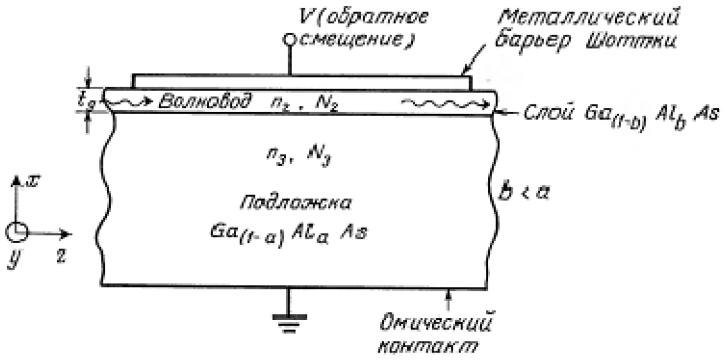


Рисунок 4 - Основная структура электроабсорбционного модулятора.

Концентрацию алюминия в волноводе следует выбирать такой, чтобы длина волны, соответствующая границе собственного поглощения, была немного меньше длины волны излучения, на которой работает волновод, т. е. волновод должен быть «прозрачен» при V=0. Концентрации носителей следует выбирать такими, чтобы $N_3 >> N_2$ тогда при подаче напряжения V в волноводном слое возникает сравнительно большое электрическое поле.

Так как ΔЕ зависит от напряженности электрического поля, -а величина а зависит от ΔЕ (как видно из рисунка 2), то можно изготовить очень эффективный электроабсорбционный модулятор для излучения с длиной волны, немного меньшей длины волны, соответствующей полосе поглощения полупроводника. Основная схема электроабсорбционного модулятора представлена на рисунке 4. В качестве поверхностного электрода можно «использовать контакт типа барьера Шоттки или неглубокий p-n переход. В любом случае в обедненном слое создается электрическое поле. В идеале концентрация примесей в волноводе N_2 должна быть настолько малой, чтобы обедненный слой распространялся через весь волноводный слой в направлении x. Длина модулятора и приложенное напряжение выбираются с использованием кривых поглощения, подобных приведенным на рисунке 2, для достижения минимальных внутренних потерь и максимальной глубины модуляции на данной длине волны. Хотя можно использовать любой тип волновода, структуру модулятора можно усовершенствовать, применяя волновод на гетероструктуре GaAlAs, как показано на рисунке 4. В таком случае можно подобрать концентрацию алюминия в волноводном слое для достижения оптимальных характеристик на данной длине волны. Например, Рейнхарт изготовил подобные структуры для работы на длине волны 900 нм, которые характеризовались изменением пропускания в 100 раз при приложенном напряжении -8 В. Мощность, необходимая для достижения 90%-ной модуляции, составляла 0,1 мВт/МГц. Используя многослойные волноводные структуры на основе четырехкомтюнентных соединений $Al_yGa_{1-yAs1-x}Sb_x$, Кэмпбелл с сотр. изготовили электроабсорбционные модуляторы, которые функционировали в диапазоне длин волн 0,9 - 1,2 мкм. Приведем достигнутые типичные рабочие характеристики при 1,06 мкм: вводимые модулятором потери - 3 дБ, коэффициент ослабления -16 дБ, полоса модуляции более 500 МГц.

Двухканальные волноводные электрооптические модуляторы

Два расположенных рядом канальных волновода могут функционировать как направленный ответвитель, в котором оптическая энергия при синхронизме переходит от одного волновода к другому. Подобный направленный ответвитель можно превратить в электрооптический модулятор добавлением двух электродов, как показано на рисунке 5. Если напряжение модулирующего сигнала подано на электроды, то оно приводит к небольшой разнице в показателях преломления в волноводах, в результате чего возникает разница в постоянных распространения $\Delta\beta$. Если вновь прибегнуть к подходу, основанному на теории связанных мод

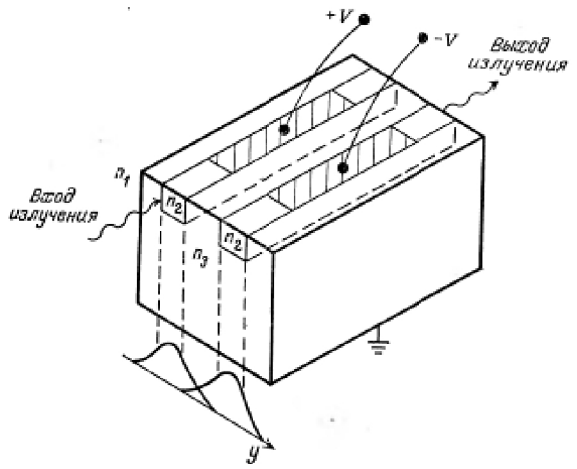


Рисунок 5 - Основная структура двухканального модулятора.

можно показать, что изменение эффективного показателя преломления для 100% модуляции определяется выражением:

$$\Delta n_g = \sqrt{3}\pi / kL, \qquad (18)$$

Обычно величина Δn_g, необходимая для 100%-ной модуляции, крайне мала. Например, в случае двухканального волноводного модулятора на основе GaAlAs длиной 1 см и площадью 3 мкм х 3 мкм, подобного изображенному на рисунке 5, излучение длиной волны в вакууме 900 нм можно полностью выключить, если

$\Delta n_g \approx 10\text{-}^4$. Можно определить, что величина требуемого электрического поля составляет около $3*10^4$ В/см, что соответствует напряжению 10 В для каналов толщиной 3 мкм. Геометрия модулятора, представленная на рисунке 5, является достаточно хорошим примером иллюстрации принципа действия двухканальных модуляторов. Однако возможны и многие другие конфигурации электродов, некоторые из которых имеют определенные преимущества в данных случаях применения.

Электрооптические модуляторы типа Маха-Цендера

Не во всех двухканальных модуляторах используется синхронизованный переход оптической энергии за счет перекрытия «хвостов» мод. Другой класс модуляторов основывается на волноводном варианте интерферометра Маха - Цендера, в котором интерференция происходит между фазово-когерентными волнами света, проходящими разные пути. Основная схема модулятора представлена на рисунке 6.

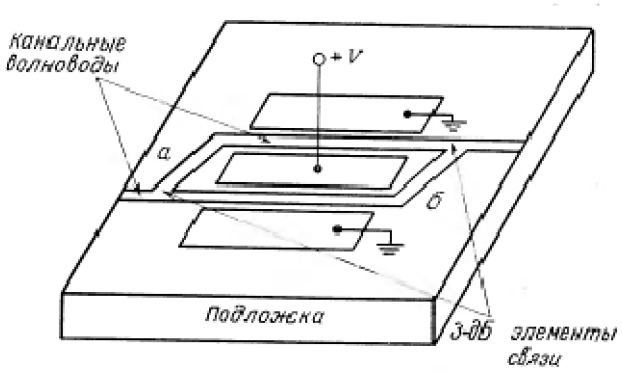


Рисунок 6 - Модулятор типа Маха - Цендера.

Свет входит в модулятор через одномодовый волновод. Разделитель пучка делит свет на два эквивалентных пучка, которые распространяются по волноводам а и б соответственно. При подаче напряжения на электроды можно изменять эффективные оптические пути пучков. В идеально разработанном

модуляторе такого типа оптические пути и характеристики волноводов идентичны, поэтому в отсутствие приложенного напряжения разделенные пучки объединяются в выходном волноводе и еще раз образуют моду наинизшего порядка. При приложении напряжения для достижения разности фаз π радиан между двумя плечами модулятора результирующее оптическое поле при объединении пучков будет иметь нулевое значение в центре выходного волновода, что соответствует моде первого порядка (m=1). Если в качестве выходного волновода используется одномодовый волновод, идентичный входному, то мода первого порядка отсекается и на коротком участке быстро затухает за счет излучения в подложку. Таким образом, модулятор можно переводить из пропускающего в непропускающее состояние за счет подачи электрического напряжения.

Эффективность использования подхода, основанного на интерферометре Маха — Цендера, была продемонстрирована на нескольких разновидностях модуляторов такого типа. Зернике первым предложил модулятор такого типа в интегрально-оптическом исполнении, используя 3-дБ направленные ответвители для разделения и последующего объединения пучка излучения вместо призм, используемых в обычном интерферометре с дискретными компонентами. Мартин изготовил модулятор Маха — Цендера на основе ZnSe, используя диффузионные волноводы, которые образовывали одномодовые Y-образные разветвления для получения нужного разделения и объединения пучков. Работая на волне 0,63 мкм, это устройство позволяло при подаче электрического напряжения 25 В переходить из состояния 60%-ного пропускания в состояние ~1%-ного пропускания. Омачи и Нода изготовили волноводный разветвляющийся интерферометр Маха-Цендера на основе $LiNbO_3$, используя диффузию TiO_2 в сочетании с высокочастотным травлением для изготовления канальных гребенчатых волноводов. Их устройство имело коэффициент ослабления 34% на волне 0,63 мкм при напряжении 19 В. Одна из главных проблем всех модуляторов на основе интерферометра Маха-Цендера, рассмотренных выше, состоит в том, что даже незначительные изменения параметров при изготовлении приводят к созданию

устройства, которое не находится во включенном состоянии при нулевом напряжении. Поэтому требуется тщательный контроль величины приложенного напряжения как для выключенного, так и для включенного состояний. Рамасвами с сотр. продвинулись еще на один шаг, используя электрически включаемые двухканальные направленные ответвители для разделения и последующего соединения пучка. Такая модификация обеспечивает электрическую подстройку как на стадии разделения, так и на стадии изменения разности фаз в плечах интерферометра. Это позволяет подобрать электрические характеристики с целью достижения оптимального значения коэффициента ослабления. После подбора напряжения на направленных ответвителях для управления переключением необходимо изменять лишь напряжение на плечах интерферометра. В устройстве такого типа, образованного диффузией Ti в $LiNbO_3$, длиной 38 мм было получено отношение включен - выключен, равное 22 дБ.

Электрооптические модуляторы, основанные на использовании явлений отражения или дифракции

Разработан целый ряд разнообразных модуляторов и переключателей, основанных на использовании электрооптического управления отражением или дифракцией каналируемого света. Работа дифракционных модуляторов базируется на эффекте Брэгга, который заключается в распределенных взаимодействиях света с множеством отражающих элементов, обычно в виде оптических решеток.

Электрооптические модуляторы на основе эффекта Брэгга

Типичный модулятор на основе эффекта Брэгга представлен на рисунке 7 и состоит из пары электродов, переплетенных друг с другом подобно двум гребенкам. Напряжение, подаваемое на электроды встречно-штыревого типа,

изменяет показатель преломления в материале под электродами, формируя в волноводе эффективную оптическую решеточную картину. Эта решетка вызывает изменение в направлении распространения светового пучка.

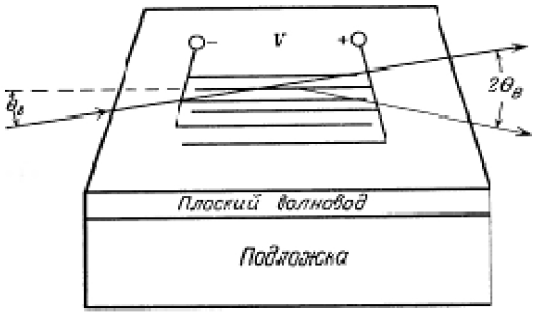


Рисунок 7 - Электрооптический модулятор на основе эффекта Брэгга.

Если направление пучка света в волноводе выбрано так, что он падает на штрихи этой решетки под углом $\Theta_Б$ эффекта Брэгга, то свет дифрагирует с максимальной эффективностью при угле $2\Theta_Б$ по отношению к входному пучку. Можно показать, что угол $\Theta_Б$ задается соотношением:

$$\sin\Theta_Б = \lambda_0 / 2\Lambda n_g , \qquad (19)$$

где Λ – период решетки, а n_g – эффективный показатель преломления волновода (β/k). Вывод выражения (19) основан на предположении толстой решетки:

$$2\pi\lambda_0 L >> \Lambda^2 . \qquad (20)$$

Если входной пучок падает на решетку под углом, отличным от угла Брэгга, то излучение все равно испытывает дифракцию в пределах угла А6Б, но с меньшей эффективностью. Диапазон изменения угла А0б при 50%-ном уменьшении эффективности определяется соотношением:

$$\Delta\Theta_Б = 2\Lambda / L, \qquad (21)$$

для малых углов $\Theta_Б$, так как в этом случае $\sin\Theta_Б \approx \Theta_Б$.

Интенсивность дифрагированного излучения зависит от приложенного электрического напряжения и обычно определяется следующим образом:

$$I / I_0 = \sin^2(VB), \qquad (22)$$

где I - интенсивность дифрагирующего света при наличии приложенного напряжения V, I_0 — интенсивность прошедшего света при V=0, B — постоянная, зависящая от показателя преломления волновода и выбранного элемента электрооптического тензора.

Электрооптические модуляторы, использующие отражение излучения

Линейный электрооптический эффект можно использовать для уменьшения показателя преломления волноводного слоя, в результате чего возникают условия полного внутреннего отражения пучка света. Модулятор такого типа представлен на рисунке 8. Устройство состоит из четырех расширяющихся канальных волноводов, которые образуют входы и выходы модулятора на основе плоского волновода. Модулятор содержит область, показатель преломления которой можно изменять за счет подачи электрического напряжения. В том случае, когда нет приложенного электрического напряжения, пучок света, входящий, например, через вход 1 или 2 не испытывая преломления, будет свободно проходить через выход 4. Если расширители тщательно спроектированы и изготовлены для сведения к минимуму рассеяния света и преобразования мод, то на выходе 3 будет наблюдаться крайне слабое перекрестное действие. Однако при подаче напряжения надлежащей полярности, уменьшающей показатель преломления в области между двумя электродами, внутри прибора возникают границы между областями с разными показателями преломления.

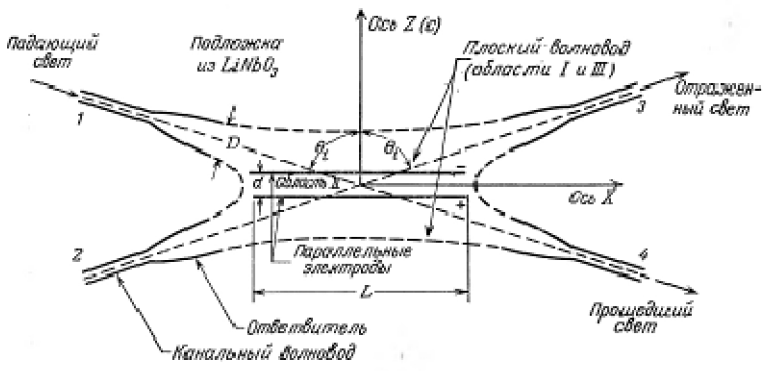


Рисунок 8 - Схема электрооптического модулятора и переключателя на основе явления полного внутреннего отражения

В таком случае излучение может испытывать полное внутреннее отражение на первой границе раздела, если угол падения больше критического угла, тем самым вызывая частичное (или, возможно, полное) переключение света на выход 3. Можно показать, что для устройства критический угол определяется формулой:

$$\Theta_c = \arcsin\left(1-\frac{1}{2}n_1^2 r_{33} V/d\right), \qquad (23)$$

где n_1 - эффективный показатель преломления материала вне области электрического поля, а d расстояние между электродами. В зависимости от напряжения, которое требуется для переключения пучка, падающего под углом Θ_i, выражение (23) принимает вид:

$$V/d = 2(1-\sin\Theta_i)/n_1^2 r_{33} \approx 1/\left(n_1^2 r_{33}\right)\left((\pi/2)-\Theta_i\right)^2, \qquad (23)$$

Переключатели описанного выше типа были изготовлены посредством диффузии Ti в $LiNbO_3$ Y-среза. Входные и выходные расширения имели длину 4,7 мм, а ширину перехода в пределах 4-40 мкм. Длина пары электродов L составляла 3,4 мм, а расстояние между ними d равно 4 мкм. Полное переключение пучка

излучения с длиной волны 632,8 нм наблюдалось при подаче напряжения, приблизительно равного 50 В. Было измерено, что перекрестное действие излучения на выходе 3 без подачи электрического напряжения составляет -15 дБ. По оценкам, быстродействие переключения превышало 6 ГГц, так как такие приборы являются сравнительно низкоемкостными.

Акустооптические модуляторы

В предыдущей главе было показано, что модуляторы и переключатели можно выполнить на основе использования электрооптического эффекта для получения периодических изменений показателя преломления материала внутри волновода. Такая решеточная структура приводит к дифракции оптических волн, распространяющихся по волноводу, что используется для модуляции и переключения оптических пучков.

Акустические волны можно также использовать для получения желаемого периодического распределения показателя преломления. Акустооптический эффект заключается в изменении показателя преломления, которое вызывается механическими напряжениями, возникающими при прохождении упругих акустических волн по материалу. В результате этого имеет место периодическое изменение показателя преломления, причем длина волны равна длине волны распространяющихся акустических волн. Рассмотрим два основных типа акустооптических модуляторов: конфигурации Брэгга и Рамана-Ната, которые отличаются друг от друга длиной взаимодействия между акустическими и оптическими волнами.

Основные принципы акустооптического эффекта

Механическое напряжение в твердом теле приводит к изменению его показателя преломления, который воздействует на фазу световой волны,

распространяющейся по этой механически напряженной среде. Этот фотоупругий эффект, как его называют, можно описать тензором четвертого ранга (тензором фотоупругости), который связывает тензор напряжений с оптической индикатрисой так же, как электрооптический тензор характеризует изменения в индикатрисе, производимые электрическим полем.

В случае акустооптического эффекта в материале за счет прохождения акустических волн возникает механическое напряжение. Поэтому такое напряжение вызывает изменение в показателя преломления за счет фотоупругого эффекта. Изменение показателя преломления Δn связано с акустической мощностью P_a следующим выражением:

$$\Delta n = \sqrt{n^6 p^2 10^7 P_a / 2\rho \upsilon_a^3 A}\,, \qquad (24)$$

где n - показатель преломления среды в отсутствие напряжения, p - соответствующий элемент тензора фотоупругости, P_a - полная акустическая мощность (в ваттах), ρ - плотность материала, υ_a - скорость звуковых волн, A - площадь поперечного сечения образца, через которую проходит волна. Все величины, входящие в выражение (24), приведены в единицах системы СГС, за исключением P_a. Если применить обычно используемый акустооптический критерий качества M_2, то выражение (24) можно переписать следующим образом:

$$\Delta n = \sqrt{M_2 P_a / 2A}\,, \qquad (25)$$

В кристаллических твердых телах, которые в основном используются в качестве подложек для ОИС, акустооптический эффект обладает ярко выраженной зависимостью от ориентации, т. е. от p. Однако этот эффект сравнительно слаб даже для оптимальных случаев подбора материала и ориентации. Например, при длине волны излучения 632,8 нм для плавленого кварца величина M_2 составляет $1{,}51*10^{-18}$ $с^3/см$, а для $LiNbO_3$ она составляет

$6{,}9*10^{-18}$ $с^3$/см. Таким образом, из выражения (25) следует, что величина Δn для этих материалов имеет порядок 10^{-4} даже в случае использования акустического излучения с плотностью мощности 100 Вт/$см^2$. Несмотря на столь малую величину Δn, которую можно достичь с помощью акустических волн, полное их воздействие на пучок света может быть весьма существенным, поскольку каждая пусть небольшая разница Δn, создаваемая пиком напряжения акустической волны, приводит к оптическим взаимодействиям, которые могут накапливаться (или не накапливаться), если обеспечено подходящее согласование фаз. Таким образом, можно получить значительные дифракционные эффекты.

Акустооптические модуляторы и переключатели, которые применяются в интегральных оптических схемах, основаны главным образом на использовании бегущих акустических волн. Поэтому индуцируемая решеточная структура профиля показателя преломления материала фактически движется относительно оптического пучка. Однако такое движение оказывает незначительное воздействие на работу большинства устройств. Среднее воздействие, связанное с таким движением решетки, идентично воздействию решетки, находящейся в стационарном состоянии, за исключением того, что спектр m-го порядка дифракции сдвинут по частоте на величину $\pm mf_0$, где f - частота акустических волн. Так как звуковые частоты, как правило, на 10 порядков по величине меньше оптических, то данный эффект вообще ничтожен.

Дифракции оптических волн можно достичь за счет взаимодействия с объемными акустическими волнами, которые проходят через объем среды, либо с поверхностными акустическими волнами (ПАВ), которые распространяются, грубо говоря, в приповерхностном слое, равном длине волны звука. Так как оптические волноводы обычно имеют толщину всего несколько микрометров, то ПАВ-модуляторы и переключатели совместимы с большинством приложений оптических интегральных схем.

Независимо от типа используемых акустических волн (объемных или поверхностных) возможны два основных типа модуляции. В модуляторах типа Рамана-Ната луч света падает перпендикулярно акустическому пучку и длина

взаимодействия между ними (т. е. ширина акустического пучка) относительно мала, поэтому оптические волны испытывают лишь простую дифракцию на фазовой решетке, что приводит к получению ряда интерференционных пиков в картине дальнего поля. Если акустический пучок настолько широк, что оптические волны испытывают многократную дифракцию, прежде чем покинут акустический пучок, то в этом случае получается существенно иная дифракционная картина. В таком случае дифракция во многом подобна объемной дифракции рентгеновского излучения на большом количестве атомных плоскостей в кристалле, которую впервые наблюдал Брэгг. В акустических модуляторах типа Брэгга оптический пучок направляют под определенным углом (углом Брэгга) к штрихам получаемой с помощью звука решеточной структуры, и при этом наблюдается лишь один дифракционный максимум в картине излучения в дальнем поле.

Модуляторы на основе дифракции Рамана-Ната

Основная схема модулятора Рамана-Ната представлена на рисунке 9. Свет проходит через это устройство в направлении оси z и испытывает сдвиг фазы, определяемый выражением:

$$\Delta\varphi = (2\pi l \Delta m / \lambda_0)\sin(2\pi y / \Lambda), \qquad (26)$$

где Δn - изменение показателя преломления, вызываемое акустическими волнами; l - длина акустического взаимодействия, а Λ - длина волны звука. Нулевая точка оси y соответствует центру падающего луча. Комбинируя выражения (25) и (26), приходим к уравнению

$$\Delta\varphi = (2\pi / \lambda_0)\sin(2\pi y / \Lambda)\sqrt{M_2 P_a l / 2a}\,, \qquad (27)$$

в котором используется тот факт, что площадь A равна l, умноженному на ширину акустического пучка a.

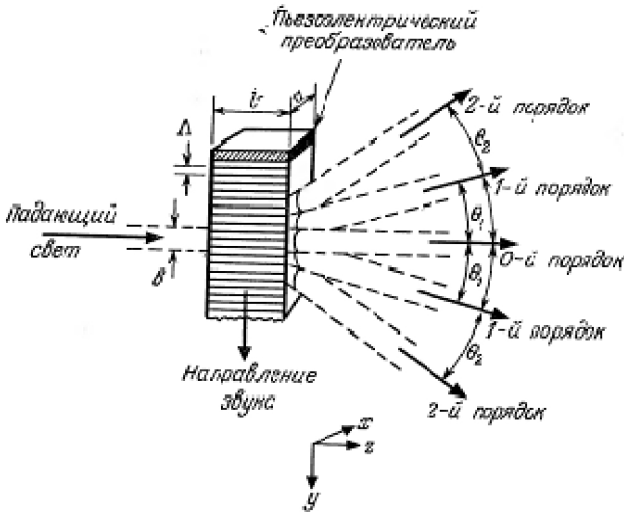


Рисунок 9 - Структура акустооптического модулятора типа Рамана-Ната.

Для получения дифракции типа Рамана-Ната длина взаимодействия должна быть настолько короткой, чтобы не возникало многократной дифракции. Это обеспечивается при выполнении условия:

$$l << \Lambda^2 / \lambda, \qquad (28)$$

где λ - длина волны света внутри материала, из которого изготовлен модулятор. При этом падающий свет дифрагирует на ряд дифракционных пучков разных порядков, причем порядок связан с углом следующим образом:

$$\sin\Theta = m\lambda_0 / \Lambda, \quad m = 0,\pm1,\pm2,\ldots. \qquad (29)$$

Интенсивность излучения в дифракционном пучке соответствующего порядка задается соотношением:

$$I / I_0 = \begin{cases} \left|J_m(\Delta\varphi')\right|^2 / 2, |\, m \,| > 0, \\ \left|J_0(\Delta\varphi')\right|^2, m = 0, \end{cases} \qquad (30)$$

где J_m - функции Бесселя; I_0 - интенсивность оптического пучка, проходящего через устройство в отсутствие акустических волн, а $\Delta\varphi'$ - максимальная величина $\Delta\varphi$, определяемая выражением:

$$\Delta\varphi' = 2\pi l\Delta n / \lambda_0 = (2\pi / \lambda_0)\sqrt{M_2 10^7 P_a l / 2a}\,. \qquad (31)$$

Выходным каналом в модуляторах Рамана-Ната обычно берут дифракционный пучок света нулевого порядка. В таком случае глубина модуляции равна части света, которая уходит за счет дифракции из дифракционного пучка нулевого порядка и задается соотношением:

$$\eta_{PH} = [I_0 - I_{m=0}]/I_0 = 1 - (J_0(\Delta\varphi'))^2\,. \qquad (32)$$

Модуляторы Рамана-Ната, как правило, имеют меньшую глубину модуляции по сравнению с модуляторами Брэгга. Кроме того, модуляторы Рамана — Ната нельзя должным образом использовать в качестве оптических переключателей, так как дифрагированный свет распределен в большое число порядков, расположенных под разными углами. По этим причинам модуляторы Рамана-Ната, хотя и представляют определенный интерес с теоретической точки зрения, но в интегральных оптических схемах используются сравнительно редко. В противоположность этому модуляторы Брэгга достаточно широко исполь-зуются в качестве модуляторов интенсивности пучка, дефлекторов и переключателей оптических пучков.

Модуляторы типа Брэгга

Для осуществления дифракции типа Брэгга длина взаимодействия между оптическим и акустическим пучками излучения должна быть сравнительно

большой, поэтому оптический пучок испытывает многократную дифракцию. Количественное соотношение, которое выражает такое условие, имеет вид:

$$l >> \Lambda^2 / \lambda, \qquad (33)$$

В случае модулятора Брэгга оптимальный входной угол оптического пучка должен равняться углу Брэгга 0б, который определяется соотношением

$$\sin \Theta_Б = \lambda / 2\Lambda. \qquad (34)$$

Дифрагированный пучок света (1-го порядка) выходит под углом $2\Theta_Б$ по отношению к недифрагированному пучку (0-го порядка), как показано на рисунке 10. Обычно в качестве выхода модулятора считают пучок нулевого порядка. В таком случае глубина модуляции задается выражением:

$$(I_0 - I)/I_0 = \sin^2(\Delta\varphi/2), \qquad (35)$$

где I_0 - интенсивность излучения, проходящего через устройство в отсутствие акустических волн, а I - интенсивность в дифракционном спектре 0-го порядка при наличии акустических волн.

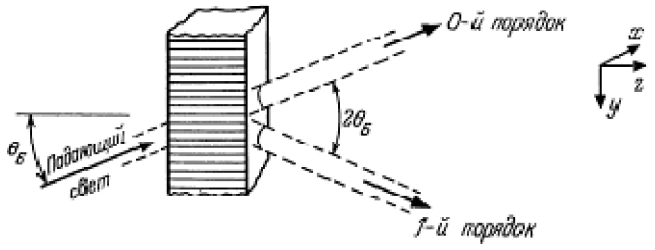


Рисунок 10 – Структура акустооптического модулятора типа Брэгга.

Структуры, представленные на рисунках 9 и 10, можно использовать в качестве объемного или волноводного модулятора в зависимости от отношения толщины *a* к длине волны λ света в материале. Волноводные акустооптические модуляторы,

используемые в составе ОИС, действуют в принципе таким же образом, как и объемные. Но они имеют существенное преимущество, связанное с меньшей потребляемой мощностью, что обусловлено тем, что оптические и акустические волны ограничены одним и тем же сравнительно небольшим объемом. Пример волноводного модулятора Брэгга представлен на рисунке 11.

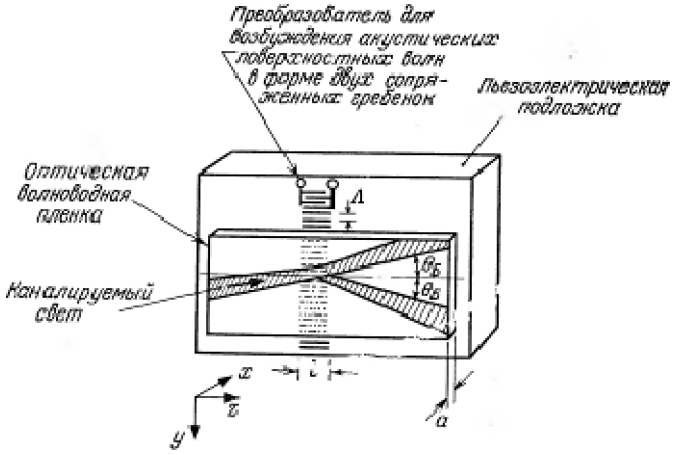


Рисунок 11 - Схема акустооптического волноводного модулятора Брэгга

Дефлекторы и переключатели оптических пучков, основанные на использовании дифракции Брэгга

Модуляторы, основанные на дифракции Брэгга, можно использовать в качестве дефлекторов и переключателей, т.к. направление выходного пучка отличается от направления входного. Если частоту акустического излучения поддерживать постоянной, то оптический пучок можно отклонить на угол $2\Theta_Б$ путем подведения достаточной акустической мощности для достижения 100% дифракции из пучка нулевого порядка в пучок первого порядка. Частоту (длину волны) акустической волны можно изменять для отклонения оптического пучка на различные углы.

При использовании отклонения оптических пучков важным параметром является число N разрешимых позиций. Предположим, что пучок света шириной b падает в область акустического поля и в пределах длины *l* взаимодействует с ним, как показано на рисунке 12.

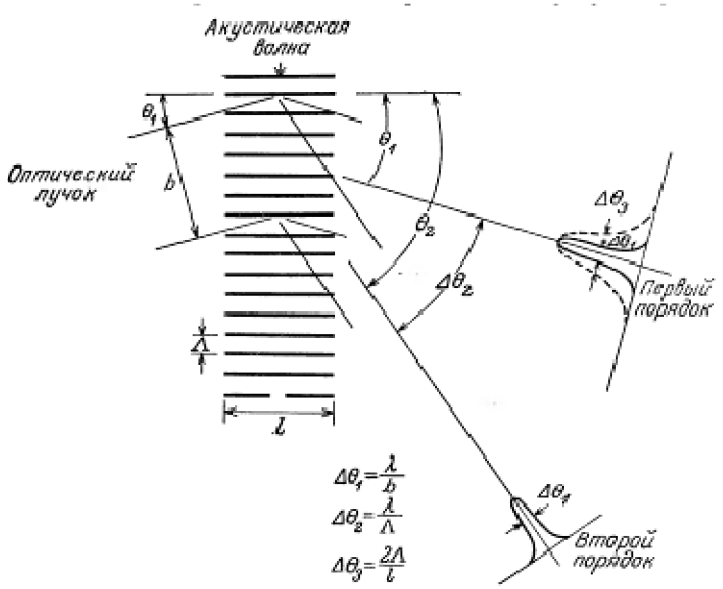


Рисунок 12 – Дифракция оптического пучка на акустической волне

В случае b>>Λ, т.е. прохождения пучка через несколько гребней акустической волны картина дифракции в дальней зон будет содержать ряд дифракционных максимумов, которые имеют ширину, определяемую по спаду интенсивности вдвое:

$$\Delta\Theta_1 = \lambda / b, \qquad (36)$$

с угловым разделением пиков, задаваемым выражением:

$$\Delta\Theta_2 = \lambda / \Lambda, \qquad (37)$$

При падении оптического пучка под углом Брэгга существенным является лишь дифракционный максимум первого порядка. Можно показать, что его интенсивность изменяется по колоколообразному закону и диапазон этого изменения по уровню 0,5 равен:

$$\Delta\Theta_3 = 2\Lambda / l, \qquad (38)$$

Следовательно, число разрешимых позиций составляет:

$$N = \Delta\Theta_3 / \Delta\Theta_1, \qquad (39)$$

Следует отметить, что угловые приращения $\Delta\Theta$ определены для акустооптичского материала, который предполагается изотропным.

Другой важной рабочей характеристикой является частотный диапазон акустического сигнала, в котором работают дефлекторы. Уменьшение интенсивности дифрагирующего оптического пучка, связанное с отклонением от условий Брэгга, определяет верхний предел этого диапазона. Можно показать, что максимальное значение Δf, соответствующее спаду интенсивности вдвое, дается выражением:

$$\Delta f_0 \approx 2\upsilon_a \Lambda / \lambda l, \qquad (40)$$

где υ_a - скорость звука. Количество разрешенных пятен можно выразить через Δf, используя выражения (39) и (40):

$$N = \Delta f_0 t, \qquad (41)$$

где t - время прохождения звуковой волны через оптический пучок шириной b, которое определяется выражением $t=b/\upsilon_a$.

Ограничение этого диапазона оптическими факторами не является единственным эффектом, который необходимо учитывать. Часто этот диапазон ограничивается более низкими частотами, связанными с инерционностью пьезоэлектрических преобразователей, которые используются для возбуждения акустических волн. Полный временной отклик т акустического дефлектора светового пучка задается выражением:

$$\tau = (1/\Delta f_0) + (1/\Delta f_a) + t, \qquad (42)$$

где Δf_a - частотный диапазон работы акустического преобразователя.

Если требуется, чтобы число разрешимых положений было более единицы, то, как видно из выражения (41), $1/\Delta f_0$ будет много меньше L. Следовательно, в таком случае выражение (42) сводится к выражению:

$$\tau \approx (1/\Delta f_a) + t = (1/\Delta f_a) + b/\upsilon_a. \qquad (43)$$

Для достижения максимальной скорости работы прибора величину τ необходимо уменьшить, что достигается уменьшение величины t и увеличением величины Δf_a. Однако отметим из выражений (41) и (43), что имеет место неизбежная взаимосвязь между числом разрешенных положений и скоростью работы, так как можно уменьшить время прохода t за счет уменьшения величины b, однако это приводит к несколько меньшей величине разрешенных положений. При этом, диапазон акустического преобразователя Δf_a является ограничивающим фактором в любом случае.

Лекция 12

Примеры интегрально-оптических схем. Многоканальный переключатель. Когерентный оптический приемник. Анализаторы спектра высокочастотных сигналов. Интегрально-оптические корреляторы, АЦП, ЦАП. Факторы, ограничивающие возможности интегральной оптики.

Многие из монолитных интегрально-оптических устройств имеют эксплуатационные характеристики, сравнимые или даже лучшие по сравнению с аналогами на основе дискретных элементов. Достигнутый относительно невысокий уровень интеграции, по-видимому, связан скорее с инерцией разработчиков, чем с принципиальными техническими ограничениями. Некоторые специалисты предпочитали разрабатывать лазеры на гетеропереходах, в то время как другие свое внимание сосредоточивали на волноводах или детекторах. Подобная специализация была выгодна на ранней стадии развития этой области, поскольку приводила к изобретению и разработке сравнительно сложных элементов, входящих в состав ОИС различного назначения. Такие приборы, как лазеры на гетеропереходах с распределенной обратной связью, акустооптические волноводные модуляторы и переключаемые электроноптические направленные ответвители, являются примерами надлежащим образом проработанной технологии. Это должно привести к разработке нескольких базовых монолитных ОИС, объединяющих лазеры, модуляторы или переключатели, а также детекторы. Такая тенденция, без сомнения, будет продолжаться и в дальнейшем должна привести к значительно большей интеграции. Примеры ОИС приведены на рисунках 1 - 14.

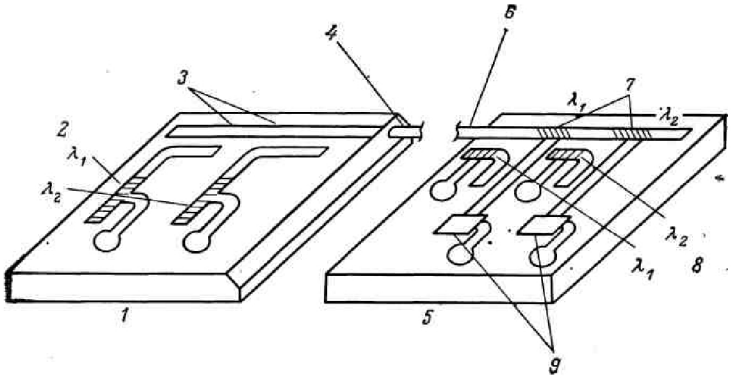


Рисунок 1 – Монолитные приемная и передающая ОИС для оптической связи: 1 - подложка передатчика; 2 - излучатели с распределенной обратной связью; 3 - пассивные направленные ответвители; 4 - элемент связи пленка - оптическое волокно; 5 подложка чип приемника; 6 - элемент связи оптическое волокно - пленка; 7 - селективные частотные фильтры; 8 — местные излучатели с распределенной обратной связью; 9 - интегральные детекторы.

В качестве источника излучения используются интегральные лазеры с распределенной обратной связью, работающие на различных длинах волн. Для простоты показано 2 таких лазера, однаков в реальных схемах их может быть несколько десятков. Излучение различных лазеров с помощью направленных ответвителей поступает в общий волновод, а затем в волоконную линию связи, где распространяется без смешивания информации. В приемном модуля излучение разделяется по длинам волн и поступает к индивидуальному детектору. Для гетеродинного детектирования иногда используют дополнительные «местные» излучатели.

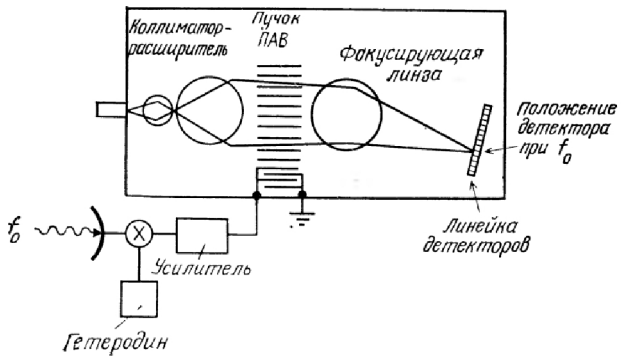


Рисунок 2 – Схема ИОС ВЧ спектроанализатора

Свет от лазерного источника вводится в планарный волновод, в котором он проходит сначала через коллимирующую линзу, а затем через акустооптический модулятор брэгговского типа. ВЧ-сигнал, спектр которого необходимо проанализировать, подается на акустический преобразователь, который генерирует звуковые волны с изменяемым периодом. Таким образом, угол отклонения оптического луча на выходе модулятора зависит от ВЧ-сигнала. Для фокусировки оптического луча на линейку фотодетекторов используется вторая линза. Если в ВЧ-сигнале присутствует более одной частотной компоненты, луч света разделяется на соответствующие компоненты, которые фокусируются на различные детекторные элементы. Каждый детекторный элемент представляет собой определенный частотный канал, и, поскольку фотодиоды обычно имеют квадратичные характеристики, выходной сигнал с каждого канала пропорционален ВЧ-мощности на данной частоте. Преимущество интегрально-оптического спектроанализатора по сравнению с электронным состоит в том, что требуется лишь несколько оптических элементов для выполнения тех функций, которые в противном случае требовали бы тысячи электронных элементов.

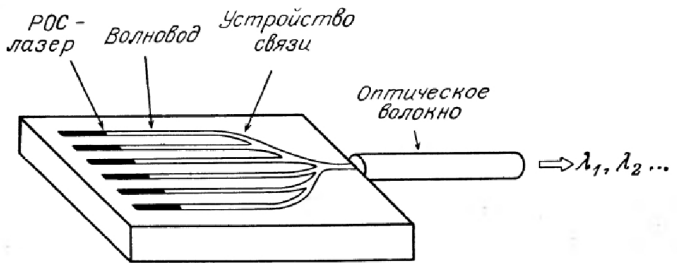


Рисунок 3 – Передающий оптический модуль со спектральным уплотнением

Передающий модуль объединяет несколько интегральных лазеров, снабженных спектрально-избирательными элементами. Излучение с различными

длинами волн вводится в интегральный, а затем в волоконный световод. На каждой длине волны передается своя информация.

Интегрально-оптические корреляторы (ИО-корреляторы) электрических и оптических сигналов могут использоваться для реализации операций типа свертки. В настоящее время существует несколько разновидностей ИО-корреляторов цифровых и аналоговых сигналов. В основе их работы используется дифракция когерентных пучков оптического излучения в планарных ОВ на дифракционных решетках (индуцированных с помощью ПАВ и/или электрооптического эффекта), последующее фурье-преобразование пространственно модулированного оптического сигнала посредством планарных волноводных или внешних линз, а также выделение пространственным фильтром и фотоприемником (или матрицы фотоприемников) корреляционного сигнала или сигнала свертки двух входных сигналов *S1(t)* и *S2(t)*.

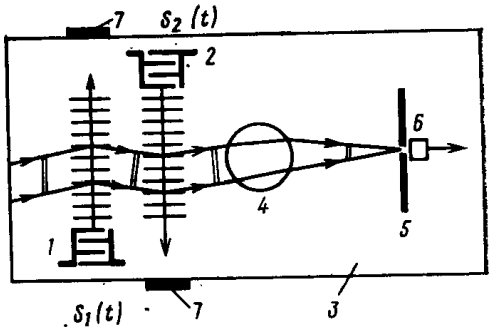


Рисунок 4 – Акустооптический пространственно-интегрирующий коррелятор аналоговых сигналов: 1, 2 – встречно-штыревые преобразователи; 3 – планарный волновод; 4 – линза; 5 – фильтр пространственных частот; 5 – фотодетектор; 6 – поглотитель поверхностной акустической волны.

Из множества разработанных ИО-корреляторов аналоговых сигналов можно выделить два основных типа: пространственно-интегрирующие и время-интегрирующие корреляторы. Схемы быстродействующих ИО-корреляторов цифровых сигналов представлены на рисунках 5 и 6. Для выполнения интегральной операции корреляции или свертки необходимо обеспечить

запоминание фильтрующей функции (эталонного кода) — сигнала *S2(t),* прием изменяющегося во времени информационного сигнала S1*(t),* который складывается из полезного сигнала и шумов, а также умножение и интегрирование сигналов.

В быстродействующем ИО-коррелятоpе цифровых сигналов (рисунок 5) эталонный код формируется с помощью соответствующих электронных устройств электрооптического пространственного модулятора света. Входной информационный сигнал *S1(t)* вызывает в волноводе генерацию радиоимпульсов ПАВ, распространяющихся на участке, длина которого (произведение длительности импульса на скорость ПАВ) согласована с длиной элемента электрооптического пространственного модулятора света. Логическая единица формируется высокочастотной ПАВ (преобразователь *2* с центральной частотой $f_0 2$ = 875 МГц), а нуль (или минус единица, если сигнал биполярный) - низкочастотной ПАВ (преобразователь / с центральной частотой $f_0 i$ =459 МГц). Умножение сигналов осуществляется в результате двукратной дифракции параллельного светового пучка в планарном ОВ, а интегрирование - с помощью оптической линзы.

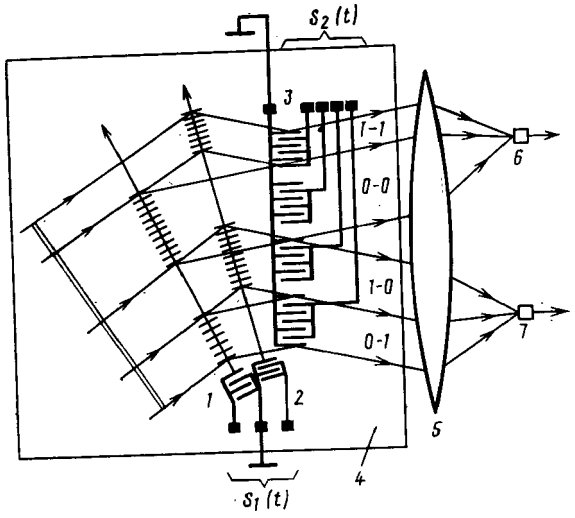


Рисунок 5 - Схема быстродействующего ИО-коррелятора цифровых сигналов:

1, 2 - встречио-штыревые преобразователи; 3 - электрооптический пространственный модулятор света; 4 - планарный ОВ; 5 - линза; 6, 7 – фотодетекторы

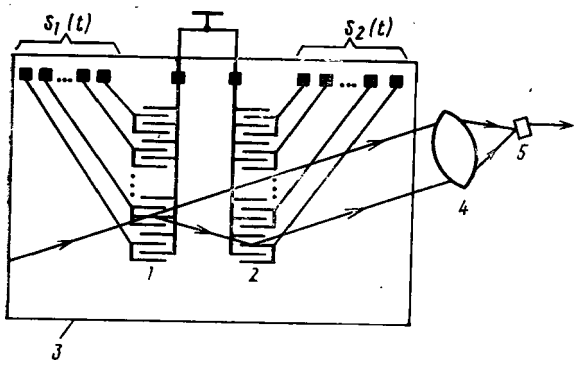


Рисунок 6 - Схема электрооптического брэгговского ИО-коррелятора цифровых сигналов:

1, 2 — система встречно» штыревых сигнальных и кодирующих электродов соответственно: 3 — планарный ОВ; 4 — линза; 5 — фотодетектор

Интегрально-оптические аналого-цифровые и цифро-аналоговые преобразователи

Интегрально-оптические АЦП (ИО-АЦП) являются многообещающими устройствами для таких областей применения, как широкополосная связь, радиолокация, радиопротиводействие, где требуется высокая частота дискретизации (примерно 10^9 отсчет/с) для преобразования широкополосных аналоговых сигналов. В настоящее время разработаны ИО-АЦП различных типов. Наиболее перспективны ИО-АЦП на основе волноводных электрооптических модуляторов интерферометрического типа (интерферометров Маха-Цендера).

Схема 4-разрядного электрооптического АЦП на основе волноводных модуляторов типа Маха-Цендера приведена на рисунке 7. Аналоговый сигнал U(t) подается на матрицу волноводных электрооптических модуляторов, каждый из которых соответствует двоичному разряду преобразователя. Выходные сигналы

модулятора воспринимаются фотоприемниками и с помощью быстродействующих электронных компараторов сравниваются с опорным сигналом, поступающим от общего источника оптического излучения. Для каждого разряда генерация бинарного сигнала (единицы или нуля) зависит соответственно от того, превышает выходной сигнал модулятора значение опорного сигнала, которая выбрана равной $I_0/2$ (где I_0 - максимальная интенсивность выходного сигнала модулятора), или он меньше ее. Аналоговый сигнал дискретизируется оптическим методом с помощью последовательности коротких оптических импульсов с требуемой частотой повторения, которые одновременно поступают на модуляторы и в опорный канал. Цифровое представление аналогового сигнала, показанное на рис. 7, б, является кодом Грея, который используется во многих типах АЦП. В этом коде на каждом шаге квантования цифровой сигнал меняется только в одном разряде, следовательно, малые погрешности при переходе от числа к числу не могут привести к накоплению ошибок преобразования на нескольких шагах.

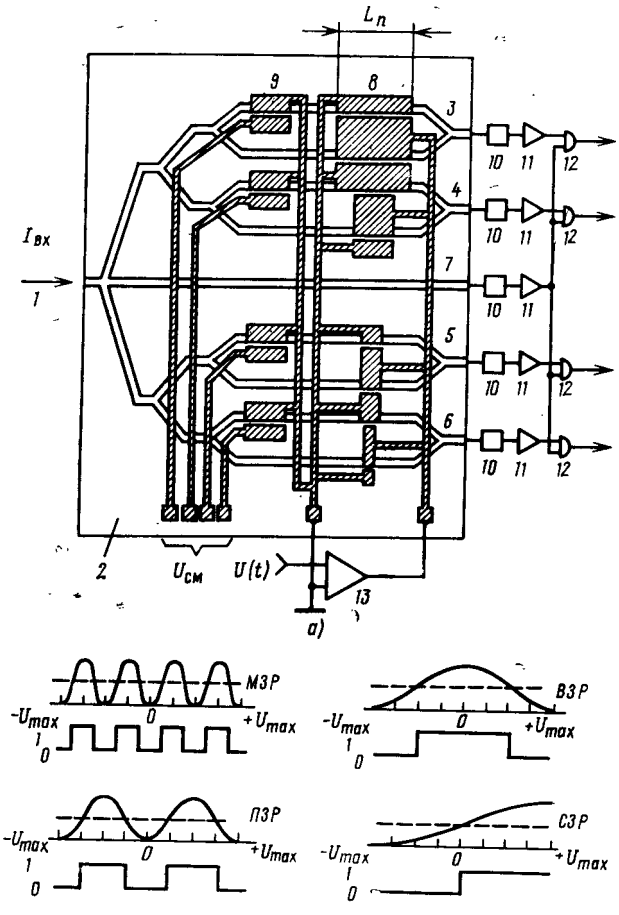


Рисунок 7 - Схема 4-разрядного электрооптического АЦП на основе волноводных интерферометров Маха-Цендера (а) и зависимости интенсивности оптического излучения на выходе модуляторов (вверху) и сигнала компараторов (внизу) от амплитуды аналогового сигнала U для различных разрядов (б): штриховая линия - уровень, соответствующий $f_0/2$; 1 - импульсное лазерное излучение; 2 - подложка с ОИС; 3-6 - модуляторы МЗР, ПЗР, ВЗР н СЗР соответственно; 7 - оптический опорный сигнал; 8 - сигнальные электроды; 9 - электроды смещения; 10 - фотодетектор; 11 - усилитель; 12 - компаратор; 13 - усилитель входного аналогового сигнала.

В этом случае напряжение U_q, соответствующее шагу квантования, для N-разрядного преобразователя равно половине полуволнового напряжения $U_{1/2}$ канала младшего значащего разряда (МЗР):

$$U_q = \lambda d/2 r_{ij} n_1^3 L_N , \quad (1)$$

где L_N - длина сигнальных электродов в канале МЗР. С ростом значимости разряда значение $U_{1/2}$ последовательно увеличивается в 2 раза для каждого последующего разряда, а длина электродов соответственно в 2 раза уменьшается. Таким образом, в общем случае длина сигнальных электродов L_n равна:

$$L_n = 2^{n-N} L_N ,$$

где n=1 соответствует старшему значащему разряду (СЗР), п = 2 - второму значащему разряду (ВЗР), n = N-1 - предпоследнему значащему разряду (ПЗР), а n = N - МЗР. Достоинством кода Грея является и то, что для его реализации АЦП требует в 2 раза меньшее напряжение, чем при эквивалентном двоично-взвешенном коде.

Статические сдвиги фаз Φ_{0n}, необходимые для реализации кода Грея, обеспечиваются подачей к электродам смещений напряжений, удовлетворяющих следующим условиям: $\Phi_{01}=\pi/2$ при n=1, $\Phi_{02} = 0$ при n = 2 и $\Phi_{0n}=\pi$ при n>2. В рассматриваемой схеме преобразования единице соответствует выходная интенсивность $I>I_0/2$. Инвертированием значений каждого разряда и применением сигналов одной полярности можно добиться того, чтобы во всех каналах, кроме СЗР, напряжения смещений были бы равны нулю. На практике постоянные смещения во всех каналах ИО-АЦП используются также для компенсации разбаланса фаз, возникающего в плечах волноводных модуляторов.

Максимальное преобразуемое напряжение аналогового сигнала определяется выражением:

$$U_{\max} = 2^{N-1}\, U_q. \qquad (2)$$

На рисунке 7, б приведены зависимости интенсивности оптического излучения на выходе матрицы модуляторов и сигналов компараторов для каждого канала АЦП от амплитуды аналогового сигнала U в пределах ±Umax. В ИО-АЦП (в отличие от обычных АЦП) напряжение аналогового сигнала U не должно превышать U_{max}, так как при $|U|>U_{max}$ цифровой код становится неоднозначным. В электрооптическом АЦП величина и U_{max} для СЗР не зависит от значения $U_{1/2}$ при $U_{1/2}$ (СЗР)$>U_{max}$. Канал СЗР дает информацию лишь о знаке сигнала, так как интенсивность выходного сигнала модулятора, равная $I_0/2$, соответствует нулевому напряжению сигнала компаратора. В этом случае при выборе $U_{1/2}$ (СЗР) $=U_{max}$ получаем $U_{1/2}$ (СЗР) = $U_{1/2}$ (ВЗР). Тогда сигнальные электроды в каналах СЗР и ВЗР имеют одинаковую длину (рисунок 7,б). Это обеспечивает максимальную крутизну зависимости интенсивности оптического сигнала от напряжения в окрестности точки U = 0, что уменьшает ошибки преобразования аналогового сигнала. Из равенства напряжений $U_{1/2}$(СЗР) и $U_{1/2}$ (ВЗР) следует, что выход за границы диапазона напряжений $|U|>U_{max}$ приводит к многозначности уровней квантования, интервалы между которыми равны $2U_{max}$. Такой большой

фиксированный интервал неоднозначности может оказаться полезным в ряде случаев, так как частичная корреляция последовательных отсчетов сигнала позволяет разрешить указанную неоднозначность.

Не вдаваясь в подробности, сформулируем требования к основным элементам ИО-АЦП рассматриваемого типа. Каждый отсчет в АЦП является средним значением входного напряжения за малый, но конечный интервал времени. В электрооптическом АЦП происходят два процесса усреднения по времени. Первый обусловлен конечным временем прохождения оптического импульса через матрицу модуляторов, а второй — его конечной длительностью. Максимально допустимое время усреднения непосредственно связано с минимальным временем $\tau_{МЗР}$ в течение которого происходит изменение МЗР. При полной амплитуде напряжения аналогового сигнала и максимальной частоте, соответствующей полосе Δf:

$$\tau_{МЗР} = 1/2^N \pi \Delta f. \qquad (3)$$

Величина $\tau_{МЗР}$ при неизменной полосе пропускания Δf уменьшается с увеличением числа битов информации. Максимально допустимое время распространения τ_N должно удовлетворять условию:

$$\tau_N < 2^{2-N/2} / \pi \Delta f = 2^{2+N/2} \tau_{МЗР}. \qquad (4)$$

Для 4-разрядного АЦП при Δf = 500 МГц $\tau_{МЗР}$ =40 пс, τ_N = 640 пс. В реальном ИО-АЦП на основе ниобата лития τ_N =175 пс и, как показывает теоретический анализ, для входного сигнала с частотой Δf =500 МГц спад частотной характеристики не превышает 0,1 дБ, а среднеквадратичная ошибка составляет всего 0,02 U_q. Это следует из (4) в предположении, что время прохождения оптическим сигналом расстояния до середины сигнальных электродов одинаково для всех разрядов. Данное условие легко обеспечивается размещением сигнальных электродов всех модуляторов на общей средней линии (рисунок 7,а). Длительность лазерного

импульса выборки на уровне половины максимальной амплитуды должна удовлетворять условию:

$$\tau_л = 8\,(N-1)^{1/2}\ln 2/2^N \pi^2 \Delta f = 8\,(N-1)^{1/2}\ln 2\cdot\tau_{МЗР}/\pi. \quad (5)$$

Для 4-разрядного АЦП длительность импульса $\tau_л$ должна быть меньше 0,06 периода сигнала, т. е. менее 120 пс для сигнала с частотой f = 500 МГц. На самом деле (5) ограничивает длительность импульса диапазоном значений от половины периода до полного периода колебаний интенсивности в канале МЗР.

Для того чтобы полностью использовать все разряды электрооптического АЦП, аналоговый входной усилитель должен при соответствующей линейности амплитудно-частотной характеристики обеспечивать выходное напряжение с амплитудой ±Umax на емкости всех сигнальных электродов матрицы модуляторов в диапазоне частот до верхней границы спектра аналогового сигнала. Анализ показывает, что применительно к параметрам реальных приборов при высоких частотах дискретизации число разрядов ИО-АЦП ограничивается длительностью лазерных импульсов $\tau_л$ (рисунок 8), а при меньших частотах дискретизации ограничивающим фактором может оказаться реально достижимая мощность входного электрического сигнала (рисунок 9). При этом полагается, что частота дискретизации вдвое больше верхней частоты спектра аналогового сигнала, а значение L_N согласовано с импедансом источника, равным 50 Ом.

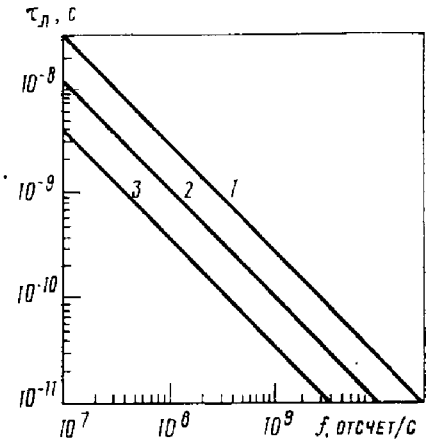


Рисунок 8 – Зависимость максимально допустимой длительности импульса выборки $\tau_л$ от частоты дискретизации f для значений числа разрядов N=2 (1), 4 (2), 6 (3).

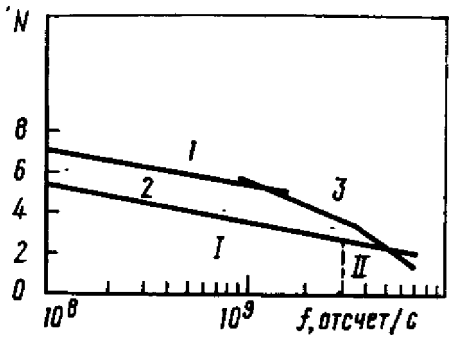


Рисунок 9 – Зависимость достижимого числа разрядов N от частоты дискретизации f для λ=0,85 мкм: 1 – мощность аналогового сигнала 10 Вт; 2 – 1 Вт; 3 – длительность импульсов выборки $\tau_л$=50 пс; I – область, соответствующая модуляторам с сосредоточенными параметрами; II – модуляторам бегущей волны.

Экспериментальные образцы ИО-АЦП, выполненные на основе волноводных модуляторов интерферометрического типа на канальных ОВ в ниобате лития, имеют число разрядов N = 4, скорость выборки порядка 1 ГГц, полосу пропускания 500 МГц, электрическая мощность аналогового сигнала имеет значение до 2,5 Вт. При этом для обеспечения нормальной работы лавинных фотодиодов оптическая мощность (для λ = 0,85 мкм) на выходе каждого модулятора должна быть не менее 1,5мкВт. Электрооптический двухразрядный ИО-АЦП может быть выполнен на волноводных модуляторах, объединенных в одну ОИС с помощью направленных ответвителей. Несмотря на более сложную структуру, такой ИО-АЦП имеет ряд преимуществ перед ранее рассмотренным преобразователем, в частности более низкий уровень вносимых потерь (<10 дБ), высокий контраст, а также более низкий уровень ошибок преобразования.

Электрические ИО-АЦП имеют ряд преимуществ перед другими методами аналого-цифрового преобразования. Во-первых, ИО-АЦП такого типа имеют

достаточно простую конструкцию, так как на каждый двоичный разряд в этом случае требуется только по одному модулятору и одному фотоприемнику. Во-вторых, такой АЦП обеспечивает развязку аналогового сигнала от оптических импульсов, осуществляющих выборку, что устраняет нежелательные искажения сигналов. В-третьих, высокое быстродействие ИО-АЦП (свыше 1 ГГц) с очень короткими импульсами выборки обеспечивается при применении полупроводниковых лазеров с синхронизацией мод. Несмотря на то, что полоса пропускания выходных электронных цепей электрооптического ИО-АЦП должна соответствовать частоте дискретизации, ни к одной цепи не предъявляется требование пропускания более широкой полосы, соответствующей спектру коротких оптических импульсов выборки. Разработанные ИО-АЦП представляют собой гибридную ОИС на основе ниобата лития, а в перспективе они могут быть выполнены как монолитные ОИС на основе полупроводниковых соединений $A^{III}B^{V}$, в частности GaAs.

Разработаны и экспериментально реализованы ИО-АЦП и других типов: ИО-АЦП на основе матрицы волноводных модуляторов поляризационного типа или матрицы волноводных электрооптически управляемых модуляторов Фабри-Перо, а также на основе волноводных акустооптических дефлекторов, которые могут быть выполнены в виде составных ОВ на кремниевой подложке. Общий принцип построения ИО-АЦП остается прежним, изменяется лишь метод преобразования аналогового электрического сигнала в ОИС (рисунок 10). Однако в этом направлении нужны еще дополнительные исследования.

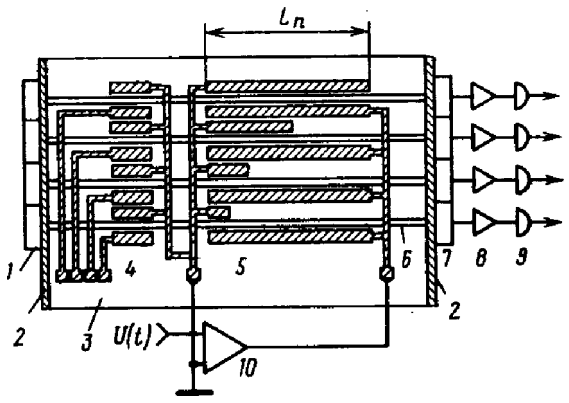


Рисунок 10 - Схема 4-разрядного электрооптического АЦП на основе волноводных модуляторов Фабри-Перо: 1 - линейка полупроводниковых лазеров; 2 - зеркала резонатора Фабри-Перо; 3 - подложка с ОИС; 4 - электроды смещения; 5 - сигнальные электроды; 6 - канальные ОВ; 7 - линейка фотодетекторов; 8 - усилители; 9 - компараторы; 10 - усилитель входного аналогового сигнала.

Наряду с ИО-АЦП разрабатываются и ИО-ЦАП. Для их построения используются волноводные электрооптические модуляторы на основе интерферометра Маха-Цендера и пространственные электрооптические модуляторы на дифракционных решетках в планарном ОВ. Схема 4-разрядного электрооптического ИО-ЦАП на основе волноводного модулятора интерферометрического типа приведена на рисунке 11.

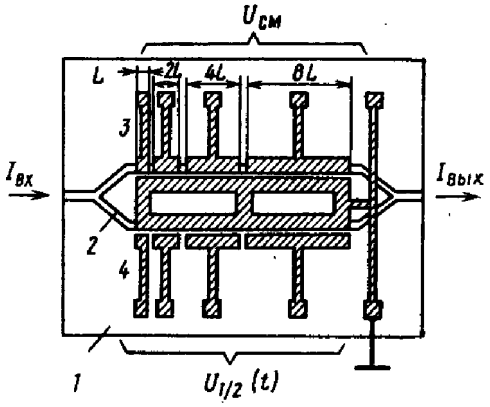


Рисунок 11 - Схема 4-разрядного электрооптического ЦАП на основе волноводного модулятора Маха-Цендера: 1 - подложка с ОИС; 2 - интерферометр; 3 - электроды смещения; 4 - сигнальные электроды.

Такой ИО-ЦАП сформирован на монокристалле ниобата лития Z-среза, имеет длину 16 мм и может управляться 4-битовым словом. Для этого сигнальные электроды смещения модулятора выполнены в виде набора четырех электродов с длинами L, 1L, 4L и 8L соответственно. При управляющем полуволновом напряжении $U_{1/2}$ = 2,6 В, прикладываемом ко всем сигнальным электродам, глубина модуляции оптического излучения на выходе модулятора достигает 93%.

Для осуществления цифро-аналогового преобразования сигналов электрические напряжения 4-битовой цифровой, схемы должны быть преобразованы к напряжению $U_{1/2}$ для всех разрядов и поданы на соответствующие сигнальные электроды ИО-ЦАП. Интенсивность оптического сигнала на выходе модулятора пропорциональна аналоговому представлению входного цифрового сигнала.

В схеме ИО-ЦАП на электрооптических дифракционных решетках электрический вход каждого дифракционного элемента пространственного модулятора выполнен так, что при одном и том же приложенном напряжении дифракционные эффективности смежных элементов различаются в 2 раза. Интенсивность излучения в планарном ОВ пропорциональна аналоговому представлению поступившего на модулятор цифрового кода. Быстродействие таких ИО-ЦАП определяется собственной емкостью системы сигнальных электродов устройства и может достигать нескольких гигагерц. Однако развитие этого направления требует дальнейших исследований и конструктивно-технологических проработок ИО-ЦАП.

Многоканальные переключатели

Коммутация оптических сигналов является одной из основных операций в системах оптической обработки и передачи информации. Разработка коммутирующих ОИС тесно связана с повышением степени интеграции большинства монолитных и гибридных ОИС для обработки информации, а также с возможностью объединения таких ОИС в более сложные многофункциональные гибридные оптические схемы. В настоящее время для разработки коммутирующих ОИС используются три основных типа электрооптически управляемых оптических ответвителей: 1) на основе связанных ОВ - волноводный направленный ответвитель; 2) волноводный брэгговский ответвитель; 3) ответвитель на основе пересекающихся или разветвляющихся ОВ с управляемой связью волноводов или зеркалами полного внутреннего отражения.

Коммутирующие ОИС могут быть выполнены в виде набора последовательно или параллельно соединенных оптических ответвителей.

Рассмотрим методы построения и архитектуру последовательных ИО-коммутаторов. Они выполняются по схеме 1 x N каналов и являются по сути многоканальными переключателями. На рисунке 12, а-в приведены схемы простейших последовательных ИО-коммутаторов 1 x N каналов, выполненных на основе направленных ответвителей и разветвляющихся ОВ в ниобате лития.

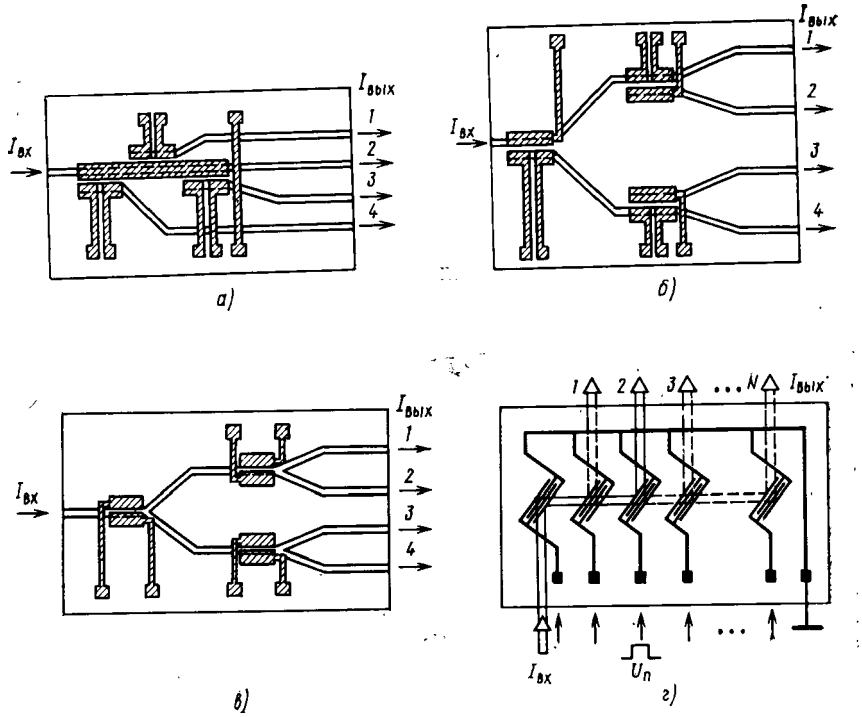


Рисунок 12 - Схемы последовательных ИО-коммутаторов 1 x N каналов на основе направленных ответвителеи (а), разветвляющихся ОВ с направленной связью (б), разветвляющихся ОВ с управляемой связью (е), для N=4, а также на основе брэгговских ответвителеи для N каналов (г)

Они имеют достаточно высокое быстродействие (частота переключения порядка нескольких сотен мегагерц), высокую эффективность переключения (свыше 99%), хорошую защищенность от переходных влияний между соседними каналами (до -25 дБ) и низкие вносимые потери (не более 5 дБ с учетом потерь на стыковку ОВ с одномодовым ВС). Интегрально-оптические коммутаторы на разветвляющихся ОВ могут быть выполнены на основе как направленных ответвителей, так и

электрооптических управляемых отражателей с симметричным и асимметричным разветвлением ОВ. Для построения ИО-коммутатора по схеме 1х4 канала требуются только три управляемых оптических ответвителя. В ИО-коммутаторе с электрооптическими управляемыми отражателями переключение каналов выполняется тремя потенциалами 0 и $\pm U_n$, защищенность от переходных влияний между коммутируемыми каналами для симметричного и асимметричного вариантов коммутатора составляет 15 и 25 дБ соответственно. Аналогичную архитектуру имеют ИО-коммутаторы на основе разветвляющихся ОВ с направленной связью. Однако с увеличением числа коммутируемых каналов N габариты (длина) рассмотренных последовательных ИО-коммутаторов значительно возрастают, что не всегда приемлемо для построения сложных многофункциональных ОИС.

Для построения последовательных ИО-коммутаторов применяются также волноводные брэгговские ответвители (рисунок 12,г), в которых переключение происходит на индуцированной в планарном ОВ (в результате электрооптического эффекта) брэгговской дифракционной решетке прямоугольной или характеристической формы. Такой ИО-коммутатор может осуществлять переключение каналов и пространственное разделение по длинам волн оптического излучения в планарном ОВ со скоростью переключения свыше 100 МГц при напряжении переключения не более 10 В. Однако и в этом случае для увеличения числа коммутируемых каналов требуется увеличивать размеры коммутирующей ОИС.

Матричные коммутирующие ОИС обеспечивают коммутацию произвольных входных и выходных каналов оптической схемы. Для построения матричных коммутирующих ОИС используются те же три основных типа управляемых оптических ответвителеи, что и для последовательных ИО-коммутаторов. Наиболее разработанными являются матричные ОИС на основе электрооптически управляемых направленных ответвителеи.

Анализ волноводных ИО-коммутаторов по схеме 3х3 канала, выполненных на эквидистантно расположенных связанных ОВ с каскадно соединенными

двухсекционными управляемыми направленными ответвителями, показывает, что для коммутации произвольных входных и выходных каналов требуется задавать определенную дискретную длину связи волноводов $L_Д = \pi/(2^{1/2}К)$. Однако технологический разброс параметров направленных ответвителей, возникающий при изготовлении таких коммутирующих ОИС, приводит к разбросу коэффициентов связи, из-за чего возникает проблема подстройки связи ОВ. Расчеты, проведенные в предположении линейной зависимости фазового рассогласования ДР от приложенного напряжения U, показывают, что состояние ИО-коммутатора, при котором возможно переключение между собой любых смежных ОВ, достигается при произвольных длинах связанных ОВ L_n, удовлетворяющих требованию $\pi/2К< L_n<3\pi/2К$. Остальные состояния достигаются каскадным соединением секций управляемых направленных ответвителеи, одна из которых имеет длину $L_Д$, а вторая L_n.

Исследование переключения оптического излучения в попарно связанных ОВ (N>>1) показывает, что построение многоэлементных матричных коммутаторов по схеме NXN каналов на основе таких ОВ с полной коммутацией всех входных и выходных волноводов возможно только при строгом контроле заданных коэффициентов связи и соответствующего фазового рассогласования для каждой пары связанных ОВ. Такая задача является достаточно сложной и трудно реализуемой на практике. Поэтому в настоящее время разработаны и нашли широкое применение матричные коммутирующие ОИС на основе различных комбинаций последовательно и параллельно соединенных электрооптических управляемых направленных ответвителей на связанных ОВ.

На рисунке 13 приведены различные варианты построения матричных коммутирующих ОИС на N x N каналов на основе волноводных направленных ответвителей. Архитектура таких ОИС определяется числом коммутируемых входов и выходов и зависит от допустимого уровня вносимых потерь и защищенности от переходных влияний между коммутируемыми каналами (перекрестных помех).

Экспериментально реализованы различные варианты коммутирующих ОИС на 4 х 4 канала на подложках из ниобата лития с различным числом управляемых направленных ответвителей. В простейшем варианте коммутирующая ОИС на 4 х 4 канала реализована на пяти направленных ответвителях (см. рисунок 13,а). В такой схеме с общей длиной 25 мм и шириной канальных ОВ 3 мкм напряжение переключения на длине волны 0,63 мкм составляет 25 В, вносимые потери не превышают 1 дБ, а перекрестные помехи -18 дБ. Для снижения уровня перекрестных помех необходимо уменьшить потери на излучение на изгибах ОВ. Выбором соответствующей архитектуры в коммутирующей ОИС на 4 х 4 канала удалось не только уменьшить число используемых направленных ответвителей до четырех, но и существенно (до -25 дБ) снизить уровень перекрестных помех (см. рисунок 13,б).

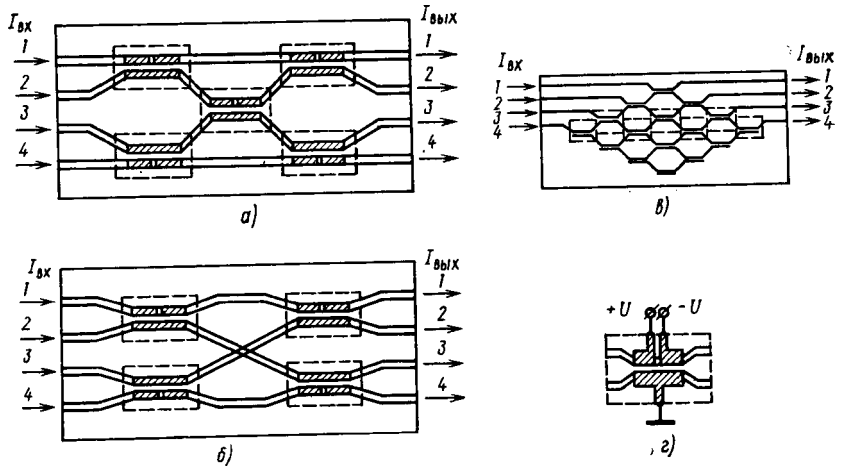


Рисунок 13 - Различные варианты матричных коммутирующих ОИС N x N каналов на основе направленных ответвителей дли N=4 (а-в) и схема электродов элементарного переключателя (г)

В общем случае в коммутирующей ОИС такой архитектуры с числом входных н выходных плеч $N=2^n$ требуется 2^{n-1} n электрооптически управляемых направленных ответвителей (n = 2, 3, 4, ...). Коммутирующая ОИС на 4 х 4 канала, выполненная на подложке из ниобата лития X-среза, имеет общую длину 20 мм и ширину канальных ОВ 5 мкм. На длине волны 1,3 мкм перекрестные помехи в

такой ОИС не превышают -25 дБ, а вносимые потери, с учетом потерь на стыковку с одномодовыми ВС на входе и выходе, составляют 4,2 дБ. Разработанная ОИС имеет практически одинаковые характеристики для обеих ортогональных поляризаций основной моды ОВ. Для создания коммутирующей ОИС на 8х8 каналов аналогичной архитектуры требуется подложка из ниобата лития в 2 раза большей длины.

Максимальное число коммутируемых каналов в матричной коммутирующей ОИС и, следовательно, число элементарных управляемых направленных ответвителей, каскадно соединенных друг с другом, ограничиваются общей длиной подложки ОИС и длиной элементарного волноводного переключателя. Построение более гибких перестраиваемых коммутирующих ОИС требует более сложной их архитектуры (см. рисунок 13,в). В таких ОИС выбором соответствующего набора управляемых направленных ответвителей можно реализовать различные варианты неблокирующих матричных коммутирующих ОИС с заданными характеристиками и независимой коммутацией призвольных входных и выходных каналов при уровне перекрестных помех между каналами -35 дБ.

Для обеспечения коммутации N входных и выходных каналов схема такого типа должна содержать как минимум 2N-1 управляемых ответвителей. Для повышения защищенности между каналами от переходных влияний в коммутирующих ОИС до уровня -40 дБ применяется существенно более сложная архитектура ОИС с пространственно разнесенными входными и выходными канальными ОВ. При этом для вполне определенной архитектуры коммутирующей ОИС с увеличением числа коммутируемых каналов защищенность между каналами от переходных влияний снижается.

Анализ различных вариантов архитектуры волноводных коммутирующих ОИС матричного типа проведен в работах. Сравнение архитектуры таких ОИС проведено с точки зрения: обеспечения заданного типа соединений входных и выходных каналов (блокирующий или неблокирующий и т. п.); необходимого числа элементарных переключателей, в том числе управляемых; общих и дифференциальных вносимых потерь; достижимой защищенности между

каналами от переходных влияний, которая связана с перекрестными помехами на отдельном переключателе; максимально реализуемой размерности матричной коммутирующей ОИС на N x N каналов.

При допустимом уровне защищенности между каналами от переходных влияний для всей схемы не хуже -11 дБ, что соотвествует скорости появления ошибки 10^{-9} бит $^{-1}$, и общих вносимых потерях не более 11 дБ для каждого канала максимальная размерность матричной коммутирующей ОИС может достигать 16 x 16 каналов для архитектуры ОИС по схеме Бениса. Однако реализация на практике архитектуры ОИС с таким числом коммутируемых каналов представляет достаточно сложную в технологическом отношении задачу. В настоящее время на основе ниобата лития Z-среза разработаны и экспериментально реализованы неблокирующие коммутирующие ОИС на 8 x 8 и 1 x 16 каналов с уровнем перекрестных помех до -30 и -10 дБ соответственно.

Представляют также интерес матричные коммутирующие ОИС на основе пересекающихся (канальных и полосковых) ОВ с зеркалами полного внутреннего отражения и на основе волноводных брэгговских ответвителей в планарных ОВ. Архитектура таких ОИС имеет много общего (рисунок 14). Для построения переключателей на зеркалах полного внутреннего отражения могут применяться как одномодовые, так и многомодовые ОВ. С целью улучшения характеристик (уменьшения напряжения переключения и снижения перекрестных помех) в коммутирующих ОИС на основе пересекающихся ОВ коммутационный элемент выполняет с протяженной областью пересечения и связи (непосредственной или распределенной) соединяемых ОВ. В такой ОИС, выполненной по схеме 4x4 канала на шести элементарных переключателях «а подложке из ниобата лития общей длиной 40 мм, на длине волны 0,83 мкм напряжение и время переключения составляют соответственно ±5 В и 1 не, вносимые потери не превышают 6 дБ, а перекрестные помехи -25 дБ.

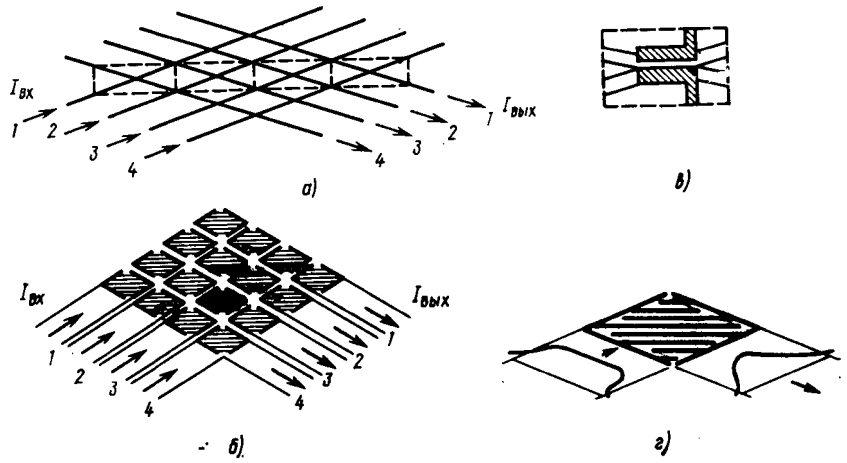


Рисунок 14 - Варианты матричных коммутирующих ОИС N x N каналов на основе ответвителей на пересекающихся ОВ (а) и брэгговских ответвителей характеристической формы (б) для N=4, а также схемы электродов их элементарных переключателей соответственно (в, г).

Матричные коммутирующие ОИС на волноводных брэгговских ответвителях характеристической формы имеют определенные преимущества перед другими типами коммутирующих ОИС. Основное преимущество такого ответвителя заключается в том, что он имеет наименьшую длину при одном и том же управляющем напряжении. Кроме того, в отличие от волноводных направленных ответвителей брэгговские ответвители позволяют создавать неблокирующие матричные коммутирующие ОИС для коммутации произвольных входов и выходов независимо друг от друга, используя только один элементарный переключатель.

В коммутирующих ОИС брэгговские ответвители характеристической формы могут быть выполнены в виде прямоугольника или ромба. Ответвители второго типа предпочтительнее для ОИС, так как имеют пороговую переключательную характеристику и высокую эффективность переключения в отличие от ответвителей прямоугольной формы, которые имеют осциллирующую переключательную характеристику, неоптимальное распределение поля дифрагировавшей волны и не обеспечивают полного переключения световых

потоков. Для повышения скорости переключения брэгговский ответвитель может быть выполнен многосекционным.

Разработана и экспериментально исследована матричная коммутирующая ОИС на 4 х 4 канала, выполненная на 16 брэгговских ответвителях характеристической формы, в которой используются планарно-канальные ОВ шириной 60 мкм в подложке из ниобата лития общей длиной 20 мм . Такая ОИС обеспечивает коммутацию и переключение произвольных входных и выходных каналов за время порядка 100 пс на длине волны 0,63 мкм при напряжении переключения 5 В и уровне перекрестных помех -26 дБ. Основной трудностью, сдерживающей широкое применение коммутирующих ОИС на брэгговских ответвителях характеристической формы, является необходимость дополнительного согласования достаточно широких коммутируемых световых пучков в планарных ОВ с другими волноводными элементами (такими, как канальные и полосковые ОВ и волоконные световоды), имеющими существенно отличающиеся поперечные размеры.

Интересные возможности для построения многофункциональных коммутирующих ОИС с улучшенными характеристиками открывают трехсекционные электрически управляемые направленные ответвители, переключатели на основе ответвителей на трех связанных ОВ с промежуточным волноводом, переключатели на ответвителях с усилением оптического сигнала, на основе нелинейных ОВ, оптически управляемые волноводные переключатели на основе квантоворазмерных структур на полупроводниковых соединениях $A^{III}B^{V}$ и др. В этом направлении сделаны лишь первые шаги, и требуются дальнейшие исследования по изучению и применению новых возможностей, которые открываются для разработки и построения многофункциональных коммутирующих ОИС, способных одновременно как обрабатывать, так и коммутировать оптические и электрические сигналы.

Факторы, ограничивающие возможности интегральной оптики.

Перспективы развития ОИС для обработки информации тесно связаны с предельными возможностями, достоинствами и ограничениями, присущими оптическим волноводным методам обработки информации, возможностью интеграции на общей подложке таких ОИС с электронными интегральными микросхемами на основе как гибридной, так и монолитной технологии, а также с совместимостью оптических компонентов ОИС друг с другом, ОИС с волоконными световодами и с интегральными микросхемами.

Остановимся сначала на основных возможностях интегральной оптики, которые используются в разрабатываемых ОИС. Планарный характер оптических волноводных элементов позволяет обеспечить возможность объединения различных волноводных функциональных элементов в ОИС разной степени интеграции на общей подложке и их совместимость с электронными микросхемами, с устройствами акустооптики и акустоэлектроники в оптоэлектронных приборах. Причем по энергетике и чувствительности волноводные устройства интегральной оптики значительно превосходят их объемные аналоги.

Ограничивающие факторы в интегральной оптике обусловлены, в конечном счете, волноводным характером распространения оптического излучения в ОВ и ОИС, определяют предельные возможности ОИС для обработки информации и связаны с двумерностью волноводных пучков планарной оптики и пространственными фазовыми флуктуациями в ОВ.

Двумерный характер оптических волноводных пучков приводит к тому, что возможно воспроизведение и преобразование лишь одномерных изображений и сигналов. Поэтому интегральная оптика не имеет прямых перспектив для построения приборов, служащих для воспроизведения неискаженных двумерных изображений. Однако применение интегральной оптики для обработки двумерных изображений в принципе не исключено, так как могут быть разработаны и построены аналоговые системы параллельной обработки в виде многослойных планарных структур, в которых управляющие и управляемые волноводные слои последовательно чередуются друг с другом, заполняя весь рабочий объем оптической системы. Подобные системы могут представлять

интерес для картинной логики и адаптивной оптики. Таким образом, одномерность обрабатываемой информации потенциально не является принципиальным ограничением в интегральной оптике, которое может быть устранено в результате ее дальнейшего успешного развития.

Причиной существенных ограничений в интегральной оптике могут стать пространственные флуктуации фазы волн в планарных и других типах ОВ, которые возникают вследствие различного рода неоднородностей в реальных волноводах. Однако в отличие от шумов, которые обусловлены в основном рассеянием света на сосредоточенных неоднородностях, имеющих размеры, сравнимые с длиной волны, величина пространственных фазовых флуктуации является интегральной характеристикой неоднородностей:

$$\Delta\Phi(z) = k \int_0^z \delta n^*(z')\, dz' ,$$

где δn*(z) - отклонение эффективного показателя преломления в ОВ от номинального (задаваемого) значения. Величина ΔΦ(z) определяется в основном медленно изменяющимися вдоль направления распространения поверхностной волны средними значениями δn*(z), т. е. плавными неоднородностями волноводных элементов достаточной протяженности. При современном уровне развития технологии интегральной оптики фазовые флуктуации приводят к значительным фазовым ошибкам взаимодействующих волн в ОИС, искажают волновые фронты и существенно ограничивают предельно достижимые характеристики - в первую очередь разрешающую способность и допустимые размеры планарных волноводных элементов и устройств, для работы которых существенно задание определенных фазовых соотношений.

Ограничивающими факторами, влияющими на динамический диапазон и уровень перекрестных помех в реальных ОИС, являются оптические потери в ОВ и волноводных элементах и связанные с ними пространственные шумы (рассеяние) оптического излучения. Волноводная дисперсия, принципиально неустранимая в интегральной оптике, не позволяет создавать бездисперсионные

элементы и схемы, сохраняющие свои свойства в широкой полосе частот, например широкополосные отражатели с большим коэффициентом отражения, подобные металлизированным зеркалам в объемной оптике.

Предельные возможности оптических интегральных схем

Рассмотренные ОИС для обработки информации построены на основе одномодовых планарных, канальных и полосковых ОВ. В соответствии с теоремой Лиувилля такие ОИС совместимы в общем случае только с одномодовыми волоконными световодами; ОИС, требующие стыковки с волоконными световодами на выходе схемы, совместимы также с многомодовыми ВС. Предельные возможности ОИС для обработки информации обусловлены тремя наиболее важными их характеристиками: минимальными размерами волноводных элементов, быстродействием и уровнем потребляемой мощности.

Существуют два ограничения, препятствующие уменьшению размеров волноводных элементов ОИС. Одно из них - минимальные размеры области локализации оптического излучения из-за конечного значения длины волны излучения. Второе связано с предельной минимальной длиной взаимодействия, необходимой для обеспечения эффективного преобразования и переключения оптических волн в ОИС. Анализ показывает, что в наиболее интересном с практической точки зрения диапазоне длин волн 0,8 ... 1,6 мкм для реальных параметров ОВ минимальные размеры области локализации оптического излучения (минимальный поперечный размер ОВ) - порядка 1 мкм. Этот предел локализации оптического излучения в О В, который вряд ли будет преодолен в будущем, приводит к тому, что по отношению к электронным интегральным схемам на кремнии и приборах на эффекте Джозефсона минимальные поперечные размеры отдельных элементов ОИС оказываются заметно большими. Именно поэтому следует ожидать, что для разработки ОИС можно будет в значительной мере воспользоваться технологическими достижениями микроэлектроники.

Кроме того, для построения высокоэффективных ИО-устройств на основе различных физических эффектов, таких, как электро-, акусто-, магнитооптические, поляризационные, нелинейные и др., необходимо обеспечить накопление этих эффектов на длинах взаимодействия, которые много больше длины волны излучения и соизмеримы с протяженностью волноводного устройства или всей ОИС. Как видно из предыдущего рассмотрения, типичные размеры ОИС варьируются примерно от единиц до десятков миллиметров.

Предельное быстродействие оптических систем, и в том числе ОИС, определяется шириной полосы оптического сигнала, однако соединения оптических компонентов с электронными схемами определенным образом ограничивают быстродействие ОИС. Основным ограничением скорости обработки оптических сигналов является задержка сигнала при его распространении в ОВ, а также уширение оптических импульсов. Задержка оптического сигнала τ при его распространении в ОИС определяется длиной схемы L_c и групповой скоростью оптической волны, которая в первом приближении равна фазовой скорости волны в ОВ, поэтому $\tau/L_c \approx n^*/c$. Для типичных параметров ОВ $\tau/L_c \approx 10$ пс/мм.

Уширение оптических импульсов играет заметную роль в волноводных элементах, обладающих дисперсией. В ОВ, применяемых для построения ОИС, дисперсию практически можно не учитывать из-за малости линейных размеров ОИС. Однако в ОВ с периодической модуляцией параметров дисперсия может приводить к уширению и искажению формы оптических импульсов. Так, в волноводном резонаторе с РБЗ происходит заметное уширение ультракоротких оптических импульсов. При относительной ширине брэгговского резонанса $\Delta\lambda/\lambda=5*10^{-4}$ наблюдалось уширение входного импульса длительностью 100 пс почти в 2 раза. Следует отметить, что быстродействие ОИС практически не ограничивается быстродействием полупроводниковых лазерных диодов - единственных лазерных излучателей, совместимых с ОИС. Полупроводниковые лазеры могут генерировать импульсы излучения пикосекундной длительности при гигагерцевых частотах повторения.

Уровни потребляемой и рассеиваемой мощности для различных типов ОИС могут изменяться в достаточно широких пределах. Так, волноводные фильтры фактически не потребляют никакой мощности, для переключателей в коммутирующих ОИС требуется незначительная мощность, а для полупроводниковых лазеров, ОИС-спектроанализаторов и ИО-АЦП - наибольшая. Энергия, необходимая для одного переключения или для модуляции одного бита информации, равна.

$$E_{\text{п}} = CU_{1/2}^2/2,$$

где С — емкость копланарных электродов переключателя, которая практически не зависит от ширины зазора d и пропорциональна длине электродов L. При межэлектродном зазоре d и ширине W (толщине h) ОВ порядка 7 мкм предельный коэффициент электрооптической добротности $E_{\text{п}}L$ = 1 пДж*см, что на практике получить весьма сложно. Следует отметить, что между энергией $E_{\text{п}}$, необходимой на одно переключение, и длиной ИО-устройства $L_{\text{с}}$ существует непосредственная связь: чем больше $L_{\text{с}}$ и, следовательно, задержка τ, тем меньше $E_{\text{п}}$. Нетрудно показать, что при рассматриваемых условиях предельное значение произведения $E_{\text{п}}\tau=10^2$ пДж*с. Уровни потребляемой мощности ОИС-спектроанализаторами на ПАВ и ИО-АЦП - порядка нескольких ватт. В ОИС-спектроанализаторах потребляемая мощность определяется в основном электрической мощностью, необходимой для возбуждения ПАВ в ОВ, которая зависит от частоты и ширины полосы принимаемого сигнала и акустооптической добротности материалов ОВ. В ИО-АЦП уровень потребляемой мощности определяется необходимой мощностью входного аналогового сигнала для заданного числа разрядов преобразования.

Приведенные оценки для энергии переключения $E_{\text{п}}$ неприменимы к резонансным электрооптическим структурам, например, такого типа, которые используются в бистабильных устройствах. Достигнутая в эксперименте энергия

переключения в бистабильных ИО-устройствах составляет примерно 1 пДж и, как показывают оценки, может быть уменьшена примерно на 2-3 порядка.

Принципиальные физические ограничения характеристик специализированных ОИС для обработки информации, например оптически управляемых цифровых оптических переключателей и логических элементов, обусловлены природой взаимодействия оптического излучения с нелинейными материалами, применяемыми для переключателей, и позволяют определить потенциальные области применения цифровых ОИС для обработки информации. В таких ОИС с увеличением быстродействия возрастает уровень потребляемой мощности.

В настоящее время параметры многих разработанных оптически управляемых ОИС приближаются к фундаментальным пределам. Такие устройства при временах переключения 10^{-7} ... 10^{-11} с по уровню затрачиваемой энергии на одно переключение уступают полупроводниковым устройством, однако при временах переключения 10^{-12} ... 10^{-14} с оптические устройства оказываются вне конкуренции. В настоящее время еще трудно сказать, какие типы оптически управляемых ОИС и переключателей найдут наиболее широкое применение на практике. Однако ясно, что именно эти ОИС позволяют работать в субпикосекундном диапазоне, что недостижимо для всех других типов ОИС, а также полупроводниковых и джозефсоновских устройств.

Основными ограничивающими факторами предельных характеристик ОИС-спектроанализаторов, выполненных на основе диффузионных ОВ в ниобате лития, являются рассеяние оптического излучения в плоскости планарного волновода, термооптический и фоторефрактивный эффекты и акустооптическая интермодуляция третьего порядка. Анализ показывает, что для достижения частотного разрешения 1 МГц в полосе свыше 1 ГГц в ОИС-спектроанализаторах на основе ниобата лития необходимо учитывать влияние термооптического и фоторефрактивного эффектов. При этом уровень перекрестных помех, обусловленных интермодуляцией, составляет -27 дБ по отношению к уровню интенсивности оптического излучения в первом порядке дифракции. Характе-

ристики планарных дифракционно-решетчатых линз с переменным периодом в планарных ОВ также имеют принципиальные ограничения, что необходимо учитывать при разработке ОИС на их основе.

Для оптических систем передачи и обработки информации представляют интерес, как правило, одномодовые ИО-устройства. Поэтому уровень шумов, вносимых такими устройствами в общий шум информационной системы, обусловлен главным образом тепловыми флуктуациями фазы световой волны и его практически можно не учитывать.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Семенов, А.Б. Волоконно-оптические системы современных СКС [Текст] / А.Б.Семенов. – М.: Академия АйТи; ДМК Пресс, 2007. – 632 с.

2. Бутусов, М.М. Волоконная оптика и приборостроение [Текст] / М.М.Бутусов. – М.: Машиностроение, 1988. – 328 с.

3. Гуртов, В.А. Оптоэлектроника и волоконная оптика [Текст] / В.А.Гуртов. – Петрозаводск: изд-во ПетрГУ, 2005. – 100 с.

4. Ильин, В. Стратегические бомбардировщики и ракетоносцы зарубежных стран [Текст] / В.Ильин, И.Кудишин. – М.: АСТ, 2002. – 384 с.

5. Сайт программы DARPA [Электронный ресурс] : http://www.darpa.mil/NewsEvents/Releases/2011/2011/01/13_NEW-HIP_Program_Advances_Avionics_Networks.aspx

6. Афанасьев В.Н., Букреев В.Г. и др. Электроприводы промышленных роботов с адаптивным управлением. Томск: Изд. томского университета, 1987. – 164с.

7. Белов М.П., Новиков В.А, Рассудов Л.Н. Автоматизированный электропривод типовых производственных механизмов и технологических комплексов. М.: ACADEMA, 2004. – 575 с.

8. Кацман М.М. Электрические машины автоматических устройств. М.: ФОРУМ, ИНФРА-М, 2002. – 264 с.

9. Ишанин, Г.Г. Источники и приемники излучения [Текст] / Г.Г.Ишанин и др. – СПб: Политехника, 1991. – 240 с.

10. Иванов, А.Б. Волоконная оптика: компоненты, системы передачи, измерения [Текст] / М.: компания САЙРУС СИСТЕМС, 1999. – 671 с.

11. Джексон, Р.Г. Новейшие датчики: пер. с англ. [Текст] / Р.Г.Джексон. – М.: Техносфера, 2007. – 384 с.

12. Ермаков, О. Н. Прикладная оптоэлектроника [Текст] / О. Ермаков. - М. : Техносфера, 2004. - 414 с. - (Мир электроники). - ISBN 5-94836-023-7.

13. Розеншер, Э. Оптоэлектроника [Текст] : пер. с фр. / Э. Розеншер, Б. Винтер ; под ред. О. Н. Ермакова. - [2-е испр. изд]. - М. : Техносфера, 2006. - 589 с. - (Мир электроники ; VII-04). - ISBN 5-94836-031-8.

14 Кизель, В.А. Отражение света [Текст] / В.А.Кизель. – М.: Наука, главная редакция физико-математической литературы, 1973. – 352 с.

15 Соколов, А.В. Оптические свойства металлов [Текст] / А.В.Соколов. – М.: государственное издательство физико-математчисекой литературы, 1961. – 315 с.
16 Путилин, Э.С. Оптические покрытия. Учебное пособие [Текст] / Э.С.Путилин. – СПб: СПБГУ ИТМО, 2010. – 227 с.
17 Пат. 3923380 США, МПК^2 G02F 1/32, Electrooptic modulator utilizing multiple interference [Текст] / S.Hattori, T.Hiramatsu; заявитель и патентообладатель Higawara, № 449522 заявл. 8.03.1974; опубл. 2.11.1975.
18 Сонин, А.С. Электрооптические кристаллы [Текст] / А.С.Сонин, А.С.Василевская. – М.: Атомиздат, 1971. – 328 с.
19 Кузьминов, Ю.С. Сегнетоэлектрические кристаллы для управления лазерным излучением [Текст] / Ю.С.Кузьминов. – М.: Наука, 1982. – 400 с.
20 Гурзадян, Г.Г. Нелинейно-оптические кристаллы. Свойства и применение в квантовой электронике: справочник [Текст] / Г.Г.Гурзадян, В.Г.Дмитриев, Д.Н.Никогосян. – М.: Радио и связь, 1991. – 160 с.
21 Weber, J. Handbook of optical materials [Текст] / J.Weber и др. – CRC Press, 2002.
22 Лайнс, М. Сегнетоэлектрики и родственные им материалы [Текст] / М.Лайнс, А.Гласс. – М.: Мир, 1981. – 736 с.
23 Васильев, А.А. Пространственные модуляторы света [Текст] / А.А.Васильев и др. – М.: Радио и связь, 1987. – 320 с.
24 Рез, И.С. Диэлектрики. Основные свойства и применение в электронике [Текст] / И.С.Рез, Ю.М.Поплавко. – М.: Радио и связь, 1989. – 288 с.