

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Томский государственный университет систем управления и радиоэлектроники»

Кафедра электронных приборов

Основы оптоинформатики

ОСНОВНЫЕ ПРИНЦИПЫ, МОДЕЛИ, МЕТОДЫ И СРЕДСТВА ОПТИЧЕСКОЙ
ОБРАБОТКИ ИНФОРМАЦИИ

Методические указания к самостоятельной работе
для студентов направления «Фотоника и оптоинформатика»

2012

Слядников, Евгений Евгеньевич

Основные принципы, модели, методы и средства оптической обработки информации: методические указания к самостоятельной работе для студентов направления «Фотоника и оптоинформатика» / Е.Е. Слядников; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования Томский государственный университет систем управления и радиоэлектроники, Кафедра электронных приборов. - Томск : ТУСУР, 2012. - 19 с.

Цель данной работы: формирование у студентов представлений об основах оптоинформатики - области фотоники, в которой изучаются модели, методы, средства, технологии передачи, приёма, обработки, хранения и отображения информации на основе фотонов.

Предназначено для студентов очной и заочной форм, обучающихся по направлению «Фотоника и оптоинформатика» по курсу «Основы оптоинформатики».

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Томский государственный университет систем управления и радиоэлектроники»

Кафедра электронных приборов

УТВЕРЖДАЮ
Зав.кафедрой ЭП
_____ С.М. Шандаров

«__» _____ 2012 г.

Основы оптоинформатики

ОСНОВНЫЕ ПРИНЦИПЫ, МОДЕЛИ, МЕТОДЫ И СРЕДСТВА ОПТИЧЕСКОЙ
ОБРАБОТКИ ИНФОРМАЦИИ

Методические указания к самостоятельной работе
для студентов направления «Фотоника и оптоинформатика»

Разработчик

д-р физ.-мат. наук, проф. каф.ЭП
_____ Е.Е. Слядников

«__» _____ 2012 г

Содержание

1. Пределы электронной техники и их преодоление на основе оптических альтернатив.
2. Принципы работы полупроводниковых лазеров, лазеры на гетероструктурах, лазеры и усилители на основе квантоворазмерных эффектов, вертикально-излучающие полупроводниковые лазеры, волоконные лазеры и усилители, планарные лазеры и усилители.
3. Формирование, распространение, поглощение и дисперсия световых импульсов в волоконно-оптических линиях, спектральное и временное уплотнение информационных потоков, элементная база оптических линий связи, передача оптических сигналов в атмосфере и космосе.
4. Локальная и распределенная запись информации, оптические дисковые системы записи и хранения информации, магнитооптические технологии, голографические технологии, регистрирующие среды и механизмы записи, быстроедействие, считывание информации в реальном времени - динамическая голография, ассоциативная голографическая память. Аналоговые оптические вычисления, Фурье-голография, голографическая коммутация, мультиплексирование и демultipлексирование сигналов, оптическая би- и мультистабильность, цифровая оптическая обработка сигналов. Технологии создания и перспективы применения.
5. Перспективы использования и ограничения квантовой криптографии, квантовых вычислений. Когерентно-оптические системы распознавания образов, оптические нейронные сети, оптические системы нечеткой и нейро-нечеткой логики. Голографическая парадигма в искусственном интеллекте, реализация принципов информатики мозга методами оптоинформатики.

1. Пределы электронной техники и их преодоление на основе оптических альтернатив

Компьютеры появились очень давно в нашем мире, но только в последнее время их начали так усиленно использовать во многих отраслях человеческой жизни. Ещё десять лет назад было редкостью увидеть какой-нибудь персональный компьютер — они были, но были очень дорогие, и даже не каждая фирма могла иметь у себя в офисе компьютер. А теперь? Теперь в каждом третьем доме есть компьютер, который уже глубоко вошёл в жизнь самих обитателей дома.

Сама идея создания искусственного интеллекта появилась очень давно, но только в 20 столетии её начали приводить в исполнение. Сначала появились огромные компьютеры, которые были зачастую размером с огромный дом. Использование таких машин, как вы сами понимаете, было не очень удобно. Но что поделаешь? Но мир не стоял на одном месте эволюционного развития — менялись люди, менялась их Среда обитания, и вместе с ней менялись и сами технологии, всё больше совершенствуясь. И компьютеры становились всё меньше и меньше по своим размерам, пока не достигли сегодняшних размеров.

Современные вычислительные машины представляют одно из самых значительных достижений человеческой мысли, влияние, которого на развитие научно-технического прогресса трудно переоценить. Области применения ЭВМ непрерывно расширяются. Этому в значительной степени способствует распространение персональных ЭВМ, и особенно микроЭВМ.

За время, прошедшее с 50-х годов, цифровая ЭВМ превратилась из “волшебного”, но при этом дорогого, уникального и перегретого нагромождения электронных ламп, проводов и магнитных сердечников в небольшую по размерам машину - персональный компьютер - состоящий из миллионов крошечных полупроводниковых приборов, которые упакованы в небольшие пластмассовые коробочки.

В результате этого превращения компьютеры стали применяться повсюду. Они управляют работой кассовых аппаратов, следят за работой автомобильных систем зажигания, ведут учёт семейного бюджета, или просто используются в качестве развлекательного комплекса, но это только малая часть возможностей современных компьютеров. Более того, бурный прогресс полупроводниковой микроэлектроники, представляющей собой базу вычислительной техники, свидетельствует о том, что сегодняшний уровень как самих компьютеров, так и областей их применения является лишь слабым подобием того, что наступит в будущем. Постепенно изучение компьютерной техники пытаются вводить в программы школьного обучения как обязательный предмет, чтобы ребёнок смог уже с довольно раннего возраста знать строение и возможности компьютеров. А в самих школах (в основном на западе и в Америке) уже многие годы компьютеры применялись для ведения учебной документации, а теперь они используются при изучении многих учебных дисциплин, не имеющих прямого отношения к вычислительной технике.

Перспективы развития Intel

В этой работе я постараюсь заглянуть в ближайшие планы Intel.

27 сентября 1999 года - 133 МГц FSB

Итак, с этого дня начинается жизнь систем с частотой шины 133 МГц. Казалось бы, VIA уже давно выпустил свой чипсет Apollo Pro133, который имеет возможность использования этой частоты, однако процессоров поддерживающих такую FSB не было, потому о полноценных 133 МГц говорить не приходилось. В конце сентября ситуация изменилась - на рынок выпустили первый процессор, рассчитанный на эту частоту. Правда, к сожалению, этим процессором пока не станет давно ожидаемый Coppermine, представляющий собой Pentium III, сделанный на базе технологии 0.18 мкм и имеющий интегрированный в ядро и работающий на полной частоте процессора кеш второго уровня размером 256 Кбайт. Ошибки, допущенные при проектировании этого ядра, не дают возможности представить этот процессор в конце сентября - его появлением будет ознаменован последний квартал этого года.

Но одними процессорами Intel, ясное дело, не ограничится - в этом случае создалась бы достаточно нелепая ситуация - новинки поддерживались бы только материнскими платами на чипсетах VIA. В этот же день свет увидят и два новых чипсета i820 и i810e. Выход i820 - своего рода эпохальное событие - этот чипсет откладывался и переделывался несметное количество раз - первой официальной датой его выхода был июнь этого года. Но, наконец-то разработчики и потенциальные потребители пришли к какому-никакому соглашению, что и позволит вывести i820 на рынок.

Самым большим плюсом, и самым большим минусом i820 является поддерживаемый им совершенно новый для PC тип памяти - Direct Rambus DRAM. В общем, самое сомнительное звено - первые материнские платы на чипсете i820 будут требовать от пользователя полностью сменить используемую память, с модулей DIMM перейти на RIMM. Что, в сочетании с их дороговизной и немалыми объемами памяти, требуемыми сегодняшними приложениями, выльется в копеечку, и вряд ли вызовет массовый энтузиазм.

25 октября 1999 года - Coppermine

Технологию 0.18 мкм - в жизнь! Этот девиз однозначно описывает все события, которые произошли 25 октября. В этот день начаты массовые продажи Pentium III-процессоров, выпущенных по новой технологии и начиненные новым ядром - Coppermine. Наличие в нем 256-килобайтного встроенного кеша второго уровня, работающего на частоте ядра и подобного тому, что мы имеем сегодня в Celeron, гарантирует новое увеличение производительности.

А дальше?

Что будет потом, зная гибкость самой любимой компании, точно сказать уже достаточно тяжело. Однако некие общие тенденции можно описать.

Что касается процессоров, то помимо дальнейшего наращивания скоростей, нас будет ждать и еще один ребенок из семейки Coppermine. Это новый Celeron, сделанный на этом ядре, который был запущен где-то в районе первого квартала 2000 года. Главные отличия от существующих Celeron будут скрываться в поддержке частоты системной шины 100 МГц и долгожданной поддержке набора интеловских SIMD-инструкций SSE.

После этого каких-то кардинальных событий с x86 процессорами от Intel не случилось аж до 2001 года, когда миру было представлено новое процессорное ядро - Willamette, обеспечивающее безпроблемную работу на частотах более гигагерца, преодолеть который существующие архитектуры вряд ли смогут. Willamette будет иметь L1-кеш объемом 256 Кбайт и L2-кеш как минимум 1 Мбайт. При этом данный

CPU начнет выпускаться по технологии 0.18 мкм с последующим переходом на 0.13 мкм и медную технологию, вводиться которую на 0.18 мкм Intel, в отличие от AMD, считает нецелесообразным.

Вторая интересность, поджидающая нас в том же втором квартале - Timna. Это немного напоминает Cyrix MediaGX, поскольку является Pentium III процессором с интегрированным L2-кешем 128 Кбайт, графическим контроллером и контроллером памяти, поддерживающим Direct Rambus DRAM. Timna, по идее, будет выпускаться также в виде FC-PGA, устанавливаемым в новый сокет - PGA370-S. Впрочем, и это пока только проект, кремния еще нет, потому все может измениться.

Не остановится на месте и направление Mainstream-чипсетов. Выход Camino2 произойдет во втором-третьем квартале 2000 года. Этот чипсет, представляющий собой усовершенствованный i820, будет специально оптимизироваться под Coppermine. В его состав помимо всего прочего будет входить четырехпортовый контроллер USB, контроллер Ultra ATA/100 (еще бы знать, что это такое, многоканальный AC97 кодек, интегрированный LAN-контроллер и некоторые другие возможности, которые к тому времени давно уже появятся в чипсетах VIA. В общем, скучать не придется.

Оптические компьютеры

Развитие вычислительной техники представляет собой постоянно сменяющие друг друга физические способы реализации логических алгоритмов - от механических устройств (вычислительная машина Бэббиджа) к ламповым (компьютеры 40-50-х годов Марк I и Марк II), затем к транзисторным и, наконец, к интегральным схемам. И уже на рубеже XXI века идут разговоры о скором достижении пределов применения полупроводниковых технологий и появлении вычислительных устройств, работающих на совершенно ином принципе. Все это свидетельствует о том, что прогресс не стоит на месте, и с течением времени ученые открывают новые возможности создания вычислительных систем, принципиально отличающихся от широко применяемых компьютеров. Существует несколько возможных альтернатив замены современных компьютеров, одна из которых - создание так называемых оптических компьютеров, носителем информации в которых будет световой поток.

Проникновение оптических методов в вычислительную технику ведется по трем основным направлениям. Первое основано на использовании аналоговых интерференционных оптических вычислений для решения отдельных специальных задач, связанных с необходимостью быстрого выполнения интегральных преобразований. Второе направление связано с использованием оптических соединений для передачи сигналов на различных ступенях иерархии элементов вычислительной техники, т.е. создание чисто оптических или гибридных (оптоэлектронных) соединений вместо обычных, менее надежных, электрических соединений. При этом в конструкции компьютера появляются новые элементы - оптоэлектронные преобразователи электрических сигналов в оптические и обратно. Но самым перспективным направлением развития оптических вычислительных устройств является создание компьютера, полностью состоящего из оптических устройств обработки информации. Это направление интенсивно развивают с начала 80-х годов ведущие научные центры (MTI, Sandia Laboratories и др.) и основные компании-производители компьютерного оборудования (Intel, IBM).

В основе работы различных компонентов оптического компьютера (трансфазаторы-оптические транзисторы, триггеры, ячейки памяти, носители информации) лежит явление оптической бистабильности. Оптическая бистабильность - это одно из проявлений взаимодействия света с веществом в нелинейных системах с обратной связью, при котором определенной интенсивности и поляризации падающего на вещество излучения соответствуют два (аналог 0 и 1 в полупроводниковых системах) возможных стационарных состояния световой волны, прошедшей через вещество, отличающихся амплитудой и (или) параметрами поляризации. Причем предыдущее состояние вещества однозначно определяет, какое из двух состояний световой волны реализуется на выходе. Для большего понимания явление оптической бистабильности можно сравнить с обычной петлей магнитного гистерезиса (эффект, используемый в магнитных носителях информации). Увеличение интенсивности падающего на вещество светового луча до некоторого значения I_1 приводит к резкому возрастанию интенсивности прошедшего луча; на обратном же ходе при уменьшении интенсивности падающего луча до некоторого значения $I_2 < I_1$ интенсивность прошедшего луча остается постоянной, а затем резко падает. Таким образом, интенсивности падающего пучка I , значение которой находится в пределах петли гистерезиса, соответствуют два значения интенсивности прошедшего пучка, зависящих от предыдущего оптического состояния поглощающего вещества.

Весь набор полностью оптических логических устройств для синтеза более сложных блоков оптических компьютеров реализуется на основе пассивных нелинейных резонаторов-интерферометров. В зависимости от начальных условий (начального положения пика пропускания и начальной интенсивности оптического излучения) в пассивном нелинейном резонаторе, нелинейный процесс завершается установлением одного из двух устойчивых состояний пропускания падающего излучения. А из нескольких нелинейных резонаторов можно собрать любой, более сложный логический элемент (триггер).

Элементы памяти оптического компьютера представляют собой полупроводниковые нелинейные оптические интерферометры, в основном, созданными из арсенида галлия (GaAs). Минимальный размер оптического элемента памяти определяется минимально необходимым числом атомов, для которого устойчиво наблюдается оптическая бистабильность. Это число составляет ~ 1000 атомов, что соответствует 1-10 нанометрам.

К настоящему времени уже созданы и оптимизированы отдельные составляющие оптических компьютеров – оптические процессоры, ячейки памяти), однако до полной сборки еще далеко. Основной проблемой, стоящей перед учеными, является синхронизация работы отдельных элементов оптического компьютера в единой системе, поскольку уже существующие элементы характеризуются различными параметрами рабочей волны светового излучения (интенсивность, длина волны), и уменьшение его размера. Если для конструирования оптического компьютера использовать уже разработанные компоненты, то обычный РС имел бы размеры легкового автомобиля. Однако применение оптического излучения в качестве носителя информации имеет ряд потенциальных преимуществ по сравнению с электрическими сигналами, а именно:

1. световые потоки, в отличие от электрических, могут пересекаться друг с другом;
2. световые потоки могут быть локализованы в поперечном направлении до нанометровых размеров и передаваться по свободному пространству;
3. скорость распространения светового сигнала выше скорости электрического;
4. взаимодействие световых потоков с нелинейными средами распределено по всей среде, что дает новые степени свободы (по сравнению с электронными системами) в организации связи и создании параллельных архитектур.

Вообще, создание большего количества параллельных архитектур, по сравнению с полупроводниковыми компьютерами, является основным достоинством оптических компьютеров, оно позволяет преодолеть ограничения по быстродействию и параллельной обработке информации, свойственные современным ЭВМ. Развитие оптических технологий все равно будет продолжаться, поскольку полученные результаты важны не только для создания оптических компьютеров, но также и для оптических коммуникаций и сети Internet.

Квантовый компьютер

Создание качественно новых вычислительных систем с более высокой производительностью и некоторыми характеристиками искусственного интеллекта, например с возможностью самообучения, - очень актуальная тема. Последние десять лет такие разработки ведутся во многих направлениях - наиболее успешными и быстро развивающимися из них являются квантовые компьютеры, нейрокомпьютеры и оптические компьютеры, поскольку современная элементная и технологическая база имеет все необходимое для их создания. Хотя при этом возникают определенные проблемы. Но не будем забегать вперед - все перечисленные виды компьютеров и проблемы с их разработкой подробно описаны в этой статье. Начнем с предпосылок создания более скоростных, а значит, и более высокопроизводительных вычислительных систем.

Итак, что же такое квантовый компьютер? Основной его строительной единицей является кубит (qubit, Quantum Bit). Классический бит имеет лишь два состояния - 0 и 1, тогда как состояний кубита значительно больше. Для описания состояния квантовой системы было введено понятие волновой функции, ее значение представляется в виде вектора с большим числом значений. Существуют волновые функции, которые называются собственными для какой-либо определенной величины. Квантовая система может находиться в состоянии с волновой функцией, равной линейной комбинации собственных функций, соответствующих каждому из возможных значений (такое состояние называется сложным), т. е. физически - ни в возбужденном, ни в основном состоянии. Это означает, что кубит в одну единицу времени равен и 0, и 1, тогда как классический бит в ту же единицу времени равен либо 0, либо 1. Как для классических, так и для квантовых компьютеров были введены элементарные логические операции: дизъюнкция, конъюнкция и квантовое отрицание, при помощи которых будет организована вся логика квантового компьютера.

Как работает квантовый компьютер? Согласно законам квантовой механики, энергия электрона, связанного в атоме, не произвольна. Она может иметь лишь определенный прерывный (дискретный) ряд значений E_0, E_1, \dots, E_n называемых уровнями энергии. Этот набор называют энергетическим спектром атома. Самый нижний уровень энергии E_0 , при котором энергия атома наименьшая, называется основным. Остальные уровни (E_1, E_2, \dots, E_n) соответствуют более высокой энергии атома и называются возбужденными. Излучение и поглощение атомом электромагнитной

энергии происходит отдельными порциями - квантами, или фотонами. При поглощении фотона энергия увеличивается - он переходит "вверх" - с нижнего на верхний уровень, при излучении фотона атом совершает обратный переход вниз.

Если атом в данный момент времени находится в одном из возбужденных состояний E_2 , то такое состояние атома неустойчиво, даже если на него не влияют другие частицы. Через очень короткое время атом перейдет в одно из состояний с меньшей энергией, например E_1 . Такой самопроизвольный (спонтанный) переход с одного уровня на другой и сопровождающее его спонтанное излучение столь же случайны во времени, как радиоактивный распад ядра атома. Предсказать точно момент перехода принципиально невозможно - можно лишь говорить о вероятности того, что переход произойдет через такое-то время. Но атом может перейти с уровня E_2 на E_1 не спонтанно, а под действием электромагнитной волны, если только частота этой волны достаточно близка к частоте перехода атома. Такая резонансная волна как бы "расшатывает" электрон и ускоряет его "падение" на уровень с меньшей энергией. Переходы, происходящие под действием внешнего электромагнитного поля, называются вынужденными (или стимулированными). При создании квантового компьютера основное внимание уделяется вопросам управления кубитами при помощи вынужденного излучения и недопущении спонтанного излучения, которое нарушит работу всей квантовой системы. От рассказа о физике происходящих в квантовом компьютере процессов перейдем к тому, как эти свойства реализуются в экспериментальном образце квантового компьютера.

Для того чтобы практически реализовать квантовый компьютер, существуют несколько важных правил, которые в 1996 г. привел Дивиченцо (D.P. Divincenzo). Без их выполнения не может быть построена ни одна квантовая система.

1. Точно известное число частиц системы.
2. Возможность приведения системы в точно известное начальное состояние.
3. Высокая степень изоляции от внешней среды.
4. Умение менять состояние системы согласно заданной последовательности элементарных преобразований.

Выполнение этих требований вполне реально с помощью существующих квантовых технологий.

Зная все возможности квантовых компьютеров, можно предположить, что будущее вычислительных систем предрешено, однако, несмотря на все плюсы, которые нам дают квантовые компьютеры, - это не совсем так...

Нейрокомпьютер

Для решения некоторых задач требуется создание эффективной системы искусственного интеллекта, которая могла бы обрабатывать информацию, не затрачивая много вычислительных ресурсов. И разработчиков "осенило": мозг и нервная система живых организмов позволяют решать задачи управления и эффективно обрабатывать сенсорную информацию, а это огромный плюс для создаваемых вычислительных систем. Именно это послужило предпосылкой создания искусственных вычислительных систем на базе нейронных систем живого мира. Специалисты, добившись нужных результатов в этой области, создадут компьютер с большими возможностями.

Создание компьютера на основе нейронных систем живого мира базируется на теории перцептронов, разработчиком которой был Розенблатт. Он предложил понятие перцептрона - искусственной нейронной сети, которая может обучаться распознаванию

образов. Предположим, что есть некоторая зенитно-ракетная установка, задача которой - распознать цель и определить наиболее опасную из них. Также есть два самолета вероятного противника: штурмовик и бомбардировщик. Зенитно-ракетная установка, используя оптические средства, фотографирует самолеты и отправляет полученные снимки на вход нейронной сети (при полностью сфотографированном самолете нейронная сеть быстро распознает его). Но если снимок получился плохо, то именно здесь используются основные свойства нейронной сети, одно из которых - возможность к самообучению. Например, на снимке отсутствует одно крыло и хвостовая часть самолета. Через некоторое (приемлемое) время нейронная сеть сама дорисовывает отсутствующие части и определяет тип этого самолета и дальнейшие действия по отношению к нему. Из распознанных штурмовика и бомбардировщика оператор данной зенитно-ракетной установки выберет для уничтожения более опасный самолет.

Перспективность создания компьютеров по теории Розенблатта состоит в том, что структуры, имеющие свойства мозга и нервной системы, имеют ряд особенностей, которые сильно помогают при решении сложных задач:

1. Параллельность обработки информации.
2. Способность к обучению.
3. Способность к автоматической классификации.
4. Высокая надежность.
5. Ассоциативность.

Нейрокомпьютеры - это совершенно новый тип вычислительной техники, иногда их называют биокомпьютерами. Нейрокомпьютеры можно строить на базе нейрочипов, которые функционально ориентированы на конкретный алгоритм, на решение конкретной задачи. Для решения задач разного типа требуется нейронная сеть разной топологии (топология - специальное расположение вершин, в данном случае нейрочипов, и пути их соединения). Возможна эмуляция нейрокомпьютеров (моделирование) - как программно на ПЭВМ и суперЭВМ, так и программно-аппаратно на цифровых супербольших интегральных схемах.

Искусственная нейронная сеть построена на нейроподобных элементах - искусственных нейронах и нейроподобных связях. Здесь важно заметить, что один искусственный нейрон может использоваться в работе нескольких (приблизительно похожих) алгоритмов обработки информации в сети, и каждый алгоритм осуществляется при помощи некоторого количества искусственных нейронов.

Заключение

Так на какой же основе будет построена вычислительная система будущего? Попробуем ответить на этот вопрос. В данной работе рассматривались три вида компьютеров: квантовые компьютеры, которые построены на основе явлений, возникающих в квантовой физике и дающих мощный вычислительный агрегат при решении задач сложных вычислений; нейрокомпьютеры и оптические компьютеры, которые построены на различной теоретической базе, но схожи в том, что и те и другие

занимаются обработкой информации. С достоверностью известно, что уже сейчас существуют системы обработки информации, построенные на объединении оптических и нейронных компьютеров, - это так называемые нейроно-оптические компьютеры. Для того чтобы создать мощную систему обработки информации, пришлось разработать гибридную систему, т. е. имеющую свойства как оптических, так и нейронных компьютеров. Можно предположить, что объединение квантовых и нейроно-оптических компьютеров даст миру самую мощную гибридную вычислительную систему. Такую систему от обычной будут отличать огромная производительность (за счет параллелизма) и возможность эффективной обработки и управления сенсорной информацией. Но это лишь предположение, которое никакими фактическими доказательствами в настоящее время не подкреплено. Но технология создания вычислительных систем не стоит на месте, и в ближайшем будущем на рынке возможно появление новых вычислительных систем.

Список литературы

1. Глазер В. "Световодная техника" М. Энергоатомиздат 1995г.
2. Оокоси Е. Оптоэлектроника и оптическая связь. М.: Мир, 1998г.

Контрольные вопросы по теме:

1. Основные принципы построения и функционирования оптических компьютеров
2. Физические основы квантового компьютера
3. Основные принципы работы нейрокомпьютера

2. Принципы работы полупроводниковых лазеров, лазеры на гетероструктурах, лазеры и усилители на основе квантоворазмерных эффектов, вертикально-излучающие полупроводниковые лазеры, волоконные лазеры и усилители, планарные лазеры и усилители

Полупроводниковый лазер

Полупроводниковый лазер, полупроводниковый квантовый генератор, лазер с полупроводниковым кристаллом в качестве рабочего вещества. В П. л., в отличие от лазеров др. типов, используются излучательные квантовые переходы не между изолированными уровнями энергии атомов, молекул и ионов, а между разрешенными энергетическими зонами кристалла (см. Твёрдое тело). В П. л. возбуждаются и излучают (коллективно) атомы, слагающие кристаллическую решётку. Это отличие определяет важную особенность П. л. — малые размеры и компактность (объём кристалла $\sim 10^{-6}$ — 10^{-2} см³). В П. л. удаётся получить показатель оптич. усиления до 10^4 см⁻¹ (см. Усиления оптического показатель), хотя обычно для возбуждения генерации лазера достаточны и меньшие значения (см. ниже). Другими практически важными особенностями П. л. являются: высокая эффективность преобразования электрической энергии в энергию когерентного излучения (до 30—50%); малая инерционность, обуславливающая широкую полосу частот прямой модуляции (более 10⁹ Гц); простота конструкции; возможность перестройки длины волны λ излучения и наличие большого числа полупроводников, непрерывно перекрывающих интервал длин волн от 0,32 до 32 мкм.

Люминесценция в полупроводниках. При рекомбинации электронов проводимости и дырок в полупроводниках освобождается энергия, которая может испускаться в виде квантов излучения (люминесценция) или передаваться колебаниями кристаллической решётки, т. е. переходить в тепло. Доля излучательных актов рекомбинации у таких полупроводников, как Ge и Si, очень мала, однако в некоторых полупроводниках (например, GaAs, CdS) при очистке и легировании она может приближаться к 100%.

Для наблюдения люминесценции необходимо применить какой-либо способ возбуждения (накачки) кристалла, т. е. способ генерации избыточных электронно-дырочных пар (светом, быстрыми электронами или электрическим полем). При малой скорости образования избыточных электронно-дырочных пар излучательная рекомбинация носит беспорядочный (спонтанный) характер и используется в нелазерных полупроводниковых источниках света (см. Светоизлучающий диод). Чтобы получить генерацию когерентного излучения, т. е. лазерный эффект, необходимо

создать особое состояние люминесцирующего кристалла — состояние с инверсией населённостей.

Рекомбинация электронно-дырочной пары может сопровождаться испусканием кванта излучения, близкого по энергии к ширине запрещенной зоны DE полупроводника (рис. 1, а); при этом длина волны $l \approx hc/DE$, где h — Планка постоянная, c — скорость света.

Инверсия населённостей в полупроводниках. Оптическое квантовое усиление в полупроводнике может наблюдаться в том случае, если зона проводимости вблизи её дна E_c заполнена электронами в большей степени, чем валентная зона вблизи её потолка E_v . Преобладание числа переходов с испусканием квантов над переходами с их поглощением обеспечивается тем, что на верхних уровнях находится больше электронов, чем на нижних, тогда как вероятности вынужденных переходов в обоих направлениях одинаковы. Заполнение зон принято описывать с помощью т. н. квазиуровней Ферми, отделяющих состояния с вероятностью заполнения уровней больше $1/2$ от состояний с вероятностью заполнения меньше $1/2$. Если μ_n — квазиуровни Ферми для электронов и дырок, то условие инверсии населённостей относительно переходов с энергией $h\nu$ (где ν — частота излучения) выражается формулой:

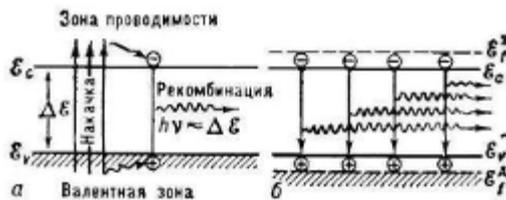


Рис 1. Полупроводниковый лазер. Энергетические схемы

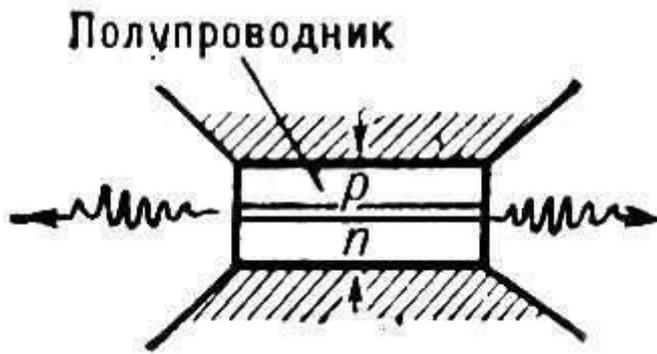


Рис 2. Инжекционный лазер на р-п-переходе

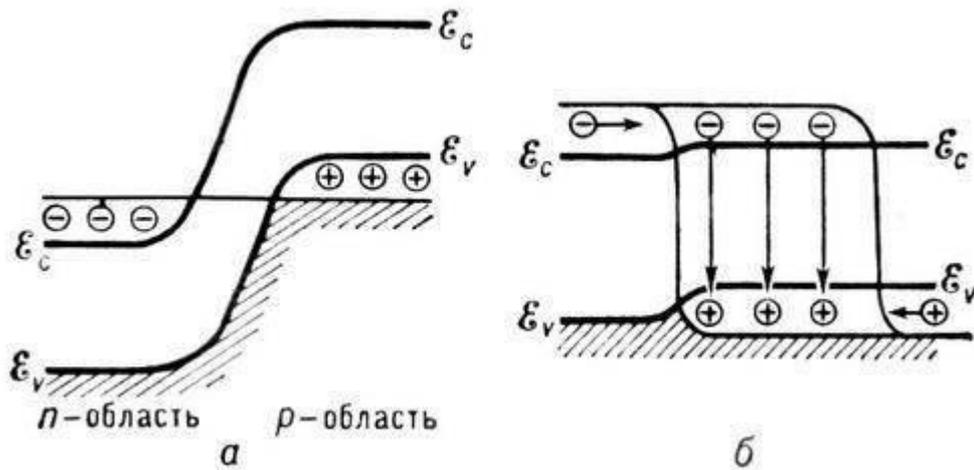


Рис 3. Энергетические зоны в р-п-переходе

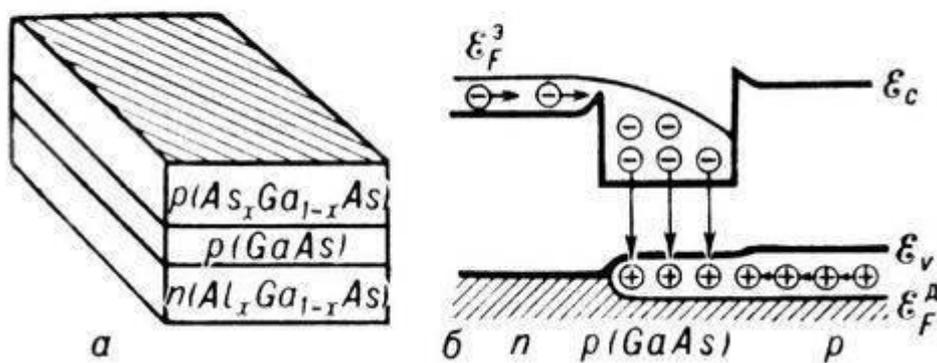


Рис 4. Лазер на гетеропереходе и его энергетическая схема

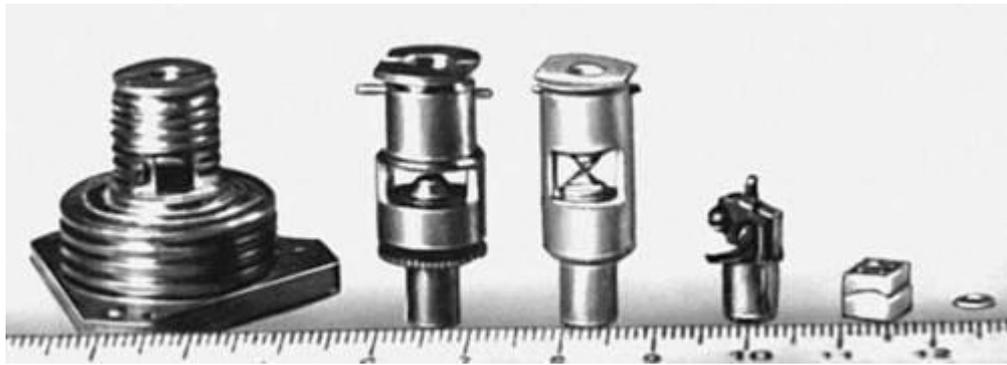


Рис 5. Образцы инжекционных лазеров

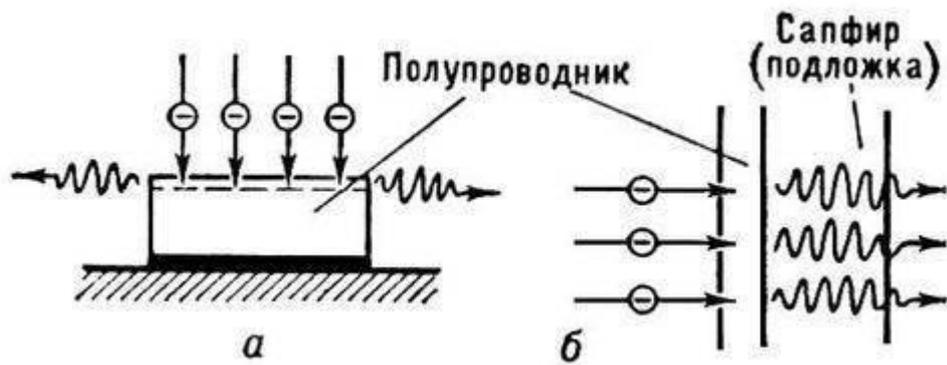


Рис 6 Полупроводниковые лазеры

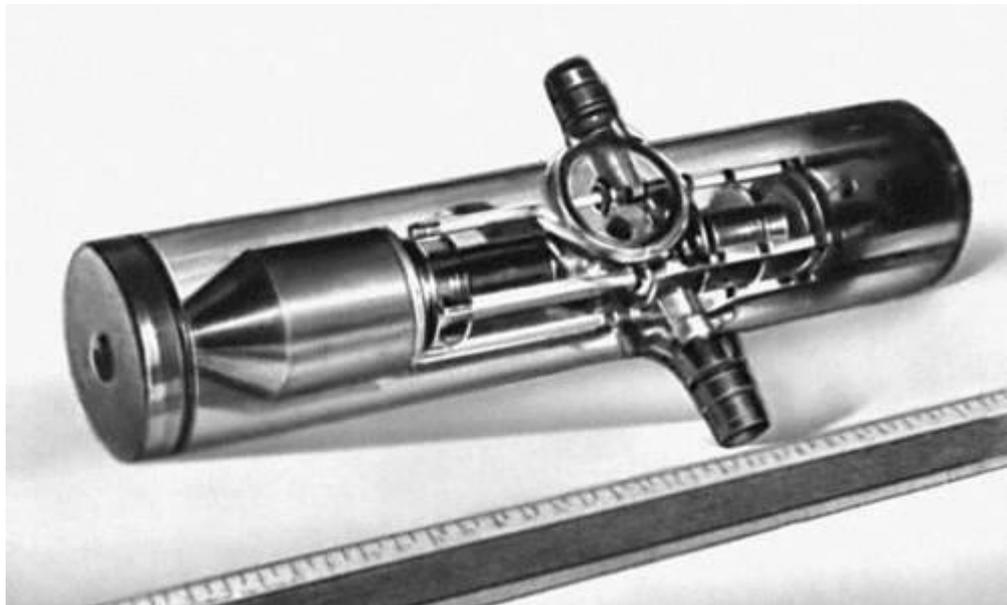


Рис 7 Полупроводниковый лазер

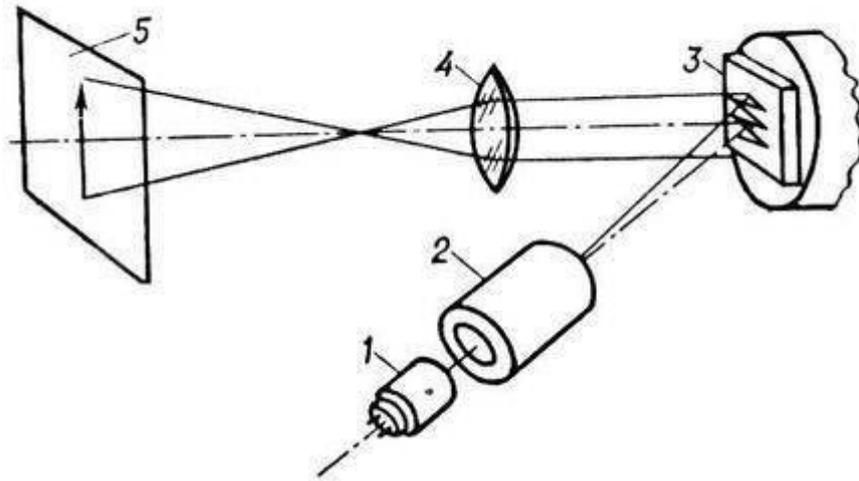


Рис 8 Проекционный лазерный телевизор

Лазеры на гетеропереходах.

Полупроводниковые лазеры отличаются от газовых и твердотельных тем, что излучающие переходы происходят в полупроводниковом материале не между дискретными энергетическими состояниями электрона, а между парой широких энергетических зон. Поэтому переход электрона из зоны проводимости в валентную зону с последующей рекомбинацией приводит к излучению, лежащему в относительно широком спектральном интервале и составляющему несколько десятков нанометров, что намного шире полосы излучения газовых или твердотельных лазеров.

Рассмотрим собственный полупроводник. В условиях термодинамического равновесия валентная зона полупроводника полностью заполнена электронами, а зона проводимости пуста. Предположим, что на полупроводник падает поток квантов электромагнитного излучения, энергия которых превышает ширину запрещенной зоны $h\nu > E_g$. Падающее излучение поглощается в веществе, так как образуются электронно-дырочные пары. Одновременно с процессом образования электронно-дырочных пар протекает процесс их рекомбинации, сопровождающийся образованием кванта электромагнитного излучения. Согласно правилу Стокса - Люммы энергия излученного кванта меньше по сравнению с энергией генерирующего кванта. Разница между этими энергиями преобразуется в энергию колебательного движения атомов кристаллической решетки. В условиях термодинамического равновесия вероятность перехода с поглощением фотона (валентная зона - зона проводимости) равна вероятности излучательного перехода (зона проводимости - валентная зона).

Предположим, что в результате какого-то внешнего воздействия полупроводник выведен из состояния термодинамического равновесия, причем в нем созданы одновременно высокие концентрации электронов в зоне проводимости и дырок в валентной зоне. Электроны переходят в состояние с некоторой энергией F_n вблизи потолка валентной зоны. Рассматриваемая ситуация иллюстрируется диаграммами,

приведенными на рис. 1. Так как все состояния вблизи дна зоны проводимости заполнены электронами, а все состояния с энергиями вблизи потолка валентной зоны заполнены дырками, то переходы с поглощением фотонов, сопровождающиеся увеличением энергии электронов становятся невозможными. Единственно возможными переходами электронов в полупроводнике в рассматриваемых условиях являются переходы зона проводимости - валентная зона, сопровождающиеся рекомбинацией электронно-дырочных пар и испусканием электромагнитного излучения. В полупроводнике создаются условия, при которых происходит усиление электромагнитной волны. Иными словами, коэффициент поглощения получается отрицательным, а рассматриваемая ситуация отвечает состоянию с инверсной плотностью населенности.

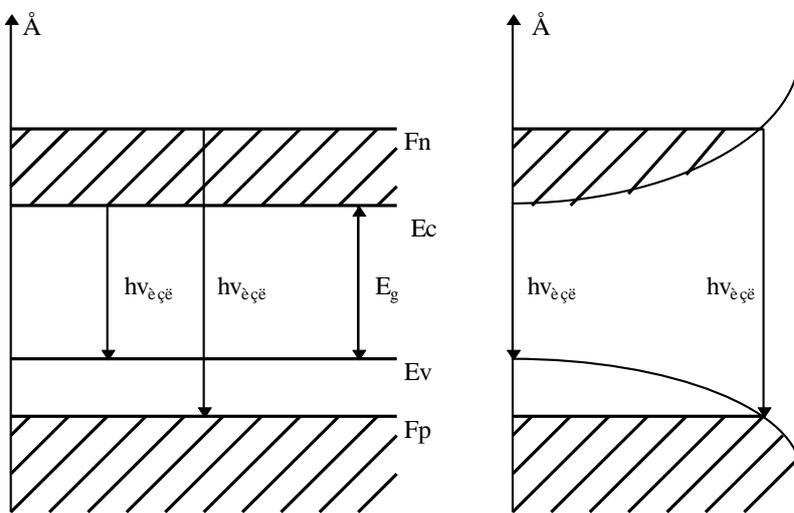


Рис. 1

Поток квантов излучения, энергия которых находится в пределах от $h\nu = E_c - E_v$ до $h\nu = F_n - F_p$, распространяется через возбужденный полупроводник беспрепятственно.

Для реализации процесса излучательной рекомбинации необходимо выполнить два условия. Во-первых, электрон и дырка должны локализоваться в одной и той же точке координатного пространства. Во-вторых электрон и дырка должны иметь одинаковые по значению и противоположно направленные скорости. Иными словами, электрон и дырка должны быть локализованы в одной и той же точке \mathbf{k} -пространства. Так как импульс образующегося в результате рекомбинации электронно-дырочной пары фотона значительно меньше по сравнению с квазиимпульсами электрона и дырки, то для выполнения закона сохранения квазиимпульса требуется обеспечить равенство квазиимпульсов электрона и дырки, участвующих в акте излучательной рекомбинации.

Оптическим переходам с сохранением квазиимпульса соответствуют вертикальные в \mathbf{k} -пространстве (прямые) переходы. Сохранение квазиимпульса в процессе излучательного перехода может рассматриваться как квантомеханическое правило отбора (в том случае, когда в акте излучательной рекомбинации не принимают участие

третьи частицы, например, фононы или атомы примесей). Невертикальные в k -пространстве (непрямые) переходы имеют значительно меньшую вероятность по сравнению с прямыми переходами, так как в этом случае требуется сбалансировать некоторый разностный квазиимпульс $d\mathbf{k}$ (рис. 2).

Таким образом для получения излучательной рекомбинации необходим прямозонный полупроводник, например, GaAs. Вообще, придерживаясь строгой теории можно доказать, что инверсная населенность возможна лишь при условии $E_c - E_g < F_n - F_p$.

Широко используемыми на практике способами создания инверсной населенности являются: 1) возбуждение за счет инжекции неосновных носителей через p-n - переход; 2) возбуждение электронным лучом; 3) возбуждение в сильном электрическом поле.

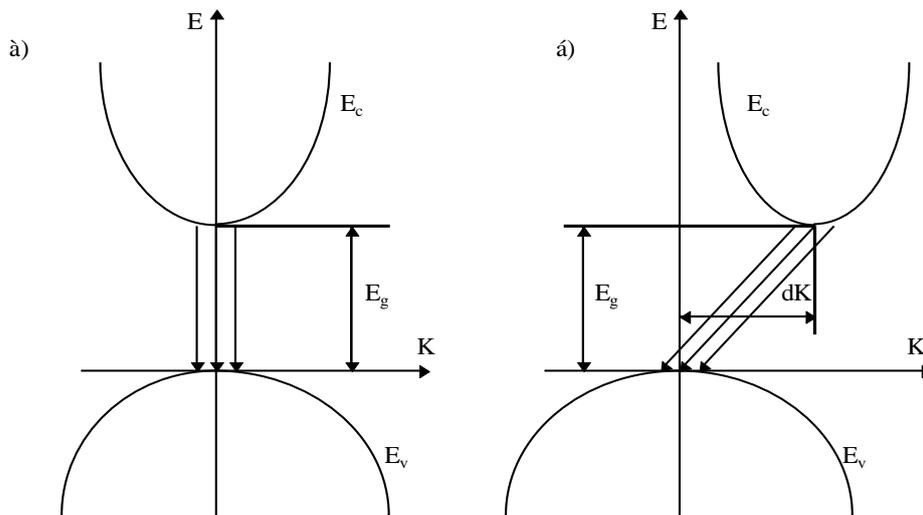


Рис. 2.

Рис. 2. Энергетические зоны (а) и (а') полупроводников.

Наиболее легко и эффективно инверсия населенности достигается в p—n-переходах за счет инжекции электронов.

Известно, что в сильнолегированных (вырожденных) полупроводниках, когда одному и тому же значению энергии соответствуют различные электронные или дырочные состояния, в p- и n-областях уровни Ферми находятся в пределах разрешенных зон и при тепловом равновесии эти уровни для электронов и дырок совпадают (рис. 3, а). В области p—n-перехода образуется потенциальный барьер, не позволяющий переходить основным носителям из зоны в зону. Если же к переходу приложить напряжения U в прямом направлении, то потенциальный барьер в области p—n-перехода уменьшается на значение энергии, соответствующей этому напряжению. Как правило, это напряжение оказывается приложенным к переходу, вследствие чего равновесие носителей тока нарушается. Если при тепловом равновесии распределение электронов и дырок можно было описать с помощью квазиуровня Ферми, то при наличии приложенного электрического поля заполнение состояний нужно рассматривать отдельно для зоны проводимости и отдельно для валентной зоны. При включении прямого смещения возникает диффузионный поток электронов через p—n-переход, который стремится поднять квазиуровень Ферми F_n для электронов в p—n-

области до его уровня в n-области. Инжектированные электроны после диффундирования на небольшое расстояние, определяемое диффузионной длиной, рекомбинируют с дырками; в результате возникает стационарное состояние, при котором скорость рекомбинации электронов в точности сбалансирована скоростью их инжекции. Совершенно аналогичны рассуждения и для дырок в валентной зоне. При наличии стационарного состояния положение квазиуровней Ферми для двух типов носителей в области перехода меняется (рис. 3, б). Основные носители вытягиваются из контакта, чтобы обеспечить условие нейтральности. В настоящее время лазерные диоды в основном изготавливают из GaAs или $Ga_{1-x}Al_xAs$. Структура лазерного диода на p—n-переходе представлена на рис. 4

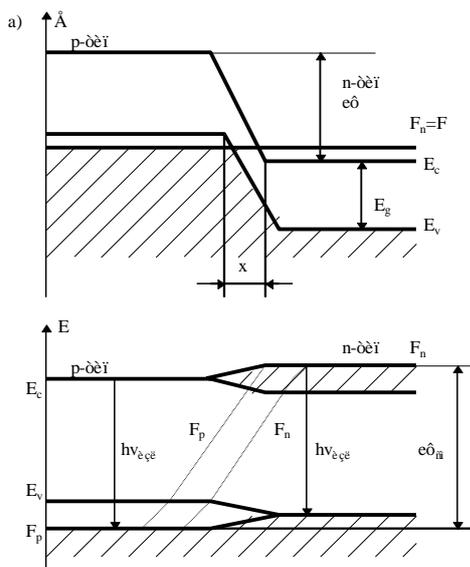


Рис. 3.

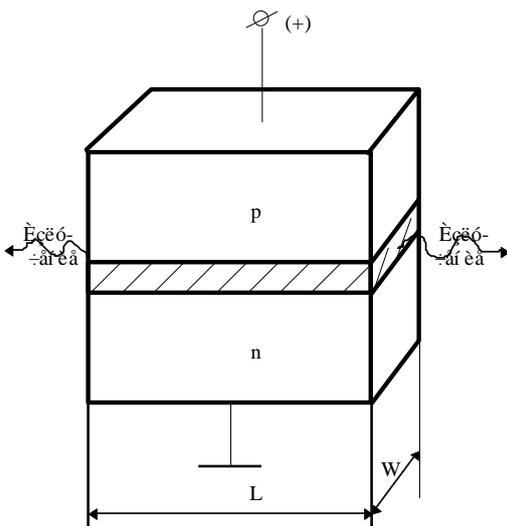


Рис. 4.

Рис. 4

Обычно р-п переход формируется путем эпитотсиального выращивания слоя р-типа на подложке п-типа. Электрический ток является источником энергии накачки, необходимой для создания инверсии населенности в активной зоне, примыкающей к р—п-переходу. Две параллельные торцевые поверхности изготавливаются путем скола по кристаллографической оси для работы в качестве зеркал резонатора и создания положительной оптической обратной связи, необходимой для генерации излучения. В силу большого показателя преломления полупроводникового материала коэффициент отражения от граней составляет 30—35%. Боковые грани лазерного кристалла имеют неровности, для того чтобы подавить поперечное нежелательное распространение света.

К основным параметрам лазерного диода относятся спектр частот излучения (оптические моды), пороговый ток, выходная мощность излучения и эффективность работы. Когда ток проходит через лазерный диод, то свет генерируется за счет инверсии населенности посредством спонтанного и стимулированного излучений. Вследствие отражения от торцов свет многократно проходит через активную область и преимущественно усиливается стимулированным излучением. Внутри лазерного диода устанавливается стоячая волна с целым числом полувольт между торцевыми поверхностями. Модовое число m задается числом полувольт

$$m=2Ln/L_w,$$

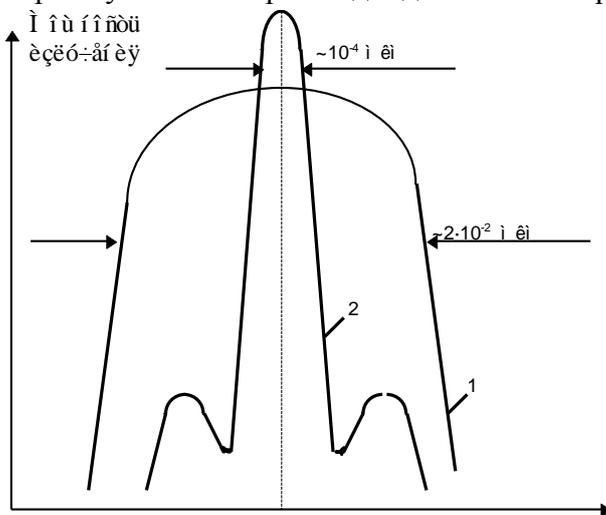
где L — расстояние между торцами; n — показатель преломления; L_w — длина волны излучения в вакууме. Модовое разделение можно установить, взяв производную dm/dL_w . Тогда

$$dm/dL_w=-2Ln/L_w^2+(2L/L_w)(dn/dL_w).$$

При $dm=-1$, что соответствует потере одной полувольты в резонаторе, получим выражение для модового разделения:

$$dL_w= dL_w^2/\{2L[n- L_w(dn/dL_w)]\}.$$

Спектр излучения лазерного диода показан на рис. 5. Обычно существует несколько



Ññ. 5.

продольных мод, имеющих длины волн вблизи пика спонтанной эмиссии. Модовое разделение для полупроводникового лазера на основе GaAs составляет $dL_w = 0.3$ нм. Для того чтобы лазер работал в одномодовом режиме, необходимо каким-либо способом подавить нежелательные боковые моды, оставив основную центральную.

Лазерный диод не сразу начинает излучать при приложении к нему напряжения от внешнего источника. При малом токе имеет место спонтанное излучение (рис. 5) с шириной спектра излучения в несколько сот микрометра. По мере нарастания тока накачки в области р—п-перехода создается высокая степень инверсии населенности и излучается больше света. Отдельные фотоны многократно проходят строго в плоскости р—п-перехода и перпендикулярно к торцам диода усиливаются. С возрастанием тока накачки испускаемое диодом излучение существенно сужается одновременно по ширине спектра и по пространственной расходимости. Когда возникает индуцированное излучение, интенсивность излучения увеличивается за счет образования большого количества электронно-дырочных пар в единицу времени. Спонтанное излучение подавляется вследствие того, что образовавшиеся первоначально фотоны повторяют себя при прохождении через активную область. Излучение лазерного диода, полученное при плотностях тока выше порогового, являются когерентными. При этом форма кривой спектрального распределения резко изменяется от широкой кривой распределения спонтанной эмиссии 1 к кривой с несколькими узкими модами 2 (рис. 5).

Значение порогового тока в зависимости от природы материала и геометрических параметров можно получить из следующих рассуждений. Пусть в области р—п-перехода существует светоизлучающий слой толщины D , который больше толщины d слоя с инверсной населенностью. Тогда можно предположить, что из всех существующих электронно-дырочных пар только часть d/D остается в активной области и может участвовать в индуцированном излучении.

Положим, что световая волна распространяется в кристалле и на каждую торцевую поверхность падает световой поток мощностью P_s , а коэффициент отражении от торца p . При наличии лазерного излучения произведение pP_s экспоненциально увеличивается в зависимости от длины активной зоны L . Существующие потери световой волны значительно перекрываются лазерным усилением за счет индуцированного излучения. Каждый торец диода излучает свет мощностью $P_{\text{вых}}/2 = (1-p)P_s$. Если μ [см⁻¹] — коэффициент потерь для волны при ее распространении в кристалле, а H [см⁻¹] — коэффициент усиления, то мощность в зависимости от пройденного волной расстояния вдоль активной области будет

$$P = pP_s \exp[H(d/D) - \mu]z.$$

Усиление волны происходит только в области с инверсной населенностью, поэтому величину H необходимо умножить на d/D , в то время как потери имеют место по всему объему и поэтому коэффициент μ не имеет такого множителя. Тогда при прохождении кристалла длиной L будем иметь:

$$P = pP_s \exp[H(d/D) - \mu]L;$$

$$\ln(1/p)=[H(d/d)-\mu]L.$$

Таким образом, условие лазерного излучения имеет вид

$$H(d/D)=\mu+(1/L) \ln(1/p). \quad (1)$$

Коэффициент усиления H связан с плотностью инжектированного тока. Выражение для величины H будет

$$H=gL_w^2 I/(8qen^2 dV), \quad (2)$$

где для GaAs при комнатной температуре квантовая эффективность $g=0.7$, длина волны излучения в вакууме $L_w=9.0 \cdot 10^{-6}$ см, показатель преломления $n=3.34$ при L_w ; V — ширина полосы спонтанного излучения, $V=1.5 \cdot 10^{13}$ с⁻¹; e — заряд электрона; d — толщина активной области, $d=10^{-4}$ см; I — плотность инжектируемого тока.

Выражение (2) справедливо для допорогового тока. Подставляя (2) в (1), получим

$$(gL_w^2 I)/(8qen^2 VD)=\mu+(1/L) \ln(1/p). \quad (3)$$

Левая часть в выражении (3) описывает усиление волны за один проход, а правая часть — потери. Из (3) найдем значение порогового тока, достаточное для покрытия потерь:

$$I=(8qen^2 VD)/(gL_w^2 I)(\mu+(1/L) \ln(1/p)). \quad (4)$$

Слагаемое $(1/L) \ln(1/p)$ определяет потери на излучение. Коэффициент отражения может быть выражен через коэффициент пропускания $T=1-p$, и тогда разложение

$\ln[1/(1-T)]$ в ряд имеет вид

$$(1/L) \ln(1/p)=(1/L) \ln[1/(1-T)]=(1/L) [T-(T^2/2)+ (T^3/3)- (T^4/4)+...].$$

Приближая членами высокого порядка по T , найдем

$$(1/L) \ln(1/p)=T/L.$$

Тогда выражение (4) представим в виде

$$I=(8qen^2 VD)/(gL_w^2 I)(\mu+T/L). \quad (5)$$

Формула (5) справедлива для приближенных расчетов. Из формулы (5) также следует, что для уменьшения I необходимо уменьшать D и наиболее оптимальным условием будет $D=d$. Но практически это условие трудно осуществить на обычном лазерном диоде, так как генерируемая в окрестности р—n-перехода световая волна распространяется не только в активной области, но и за ее пределами, где не выполняются условия инверсии населенности. Еще одной причиной является то, что часть инжектируемых электронов, обладая большой длиной свободного пробега, протаскивает активную часть р—n-перехода и не участвует в образовании электронно-дырочных пар. По этим причинам необходимо ограничить зону распространения генерируемого света и инжектируемых электронов и обеспечить условия, чтобы эти

процессы протекали только в активной области. Желаемые свойства оптического ограничения могут быть получены на гетеропереходных структурах. Самым простым из них является лазер с одинарным гетеропереходом (ОГ), представленный на рис. 6, а. Излучающий р—n-переход образуется между GaAs и $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ посредством специальной технологической обработки. Если концентрации примесей примерно одинаковы на обеих сторонах р—n-перехода, то инжекционный ток будет существовать за счет электронов, инжектируемых в слой р-типа, поскольку эффективная масса электронов почти на порядок меньше эффективной массы дырок. Поэтому слой с инверсной населенностью будет находиться в р-GaAs, толщина которого соизмерима с длиной диффузии инжектируемых электронов. Таким образом, область инверсии населенности ограничена толщиной, где в основном и происходит рекомбинация электронов с последующим излучением.

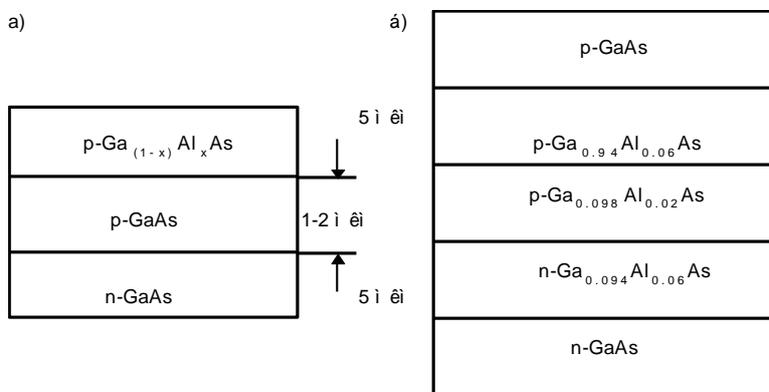


Рис. 6.

В ОГ-лазере оптическое ограничение происходит с одной стороны, отсюда желаемый результат т. е. повышение эффективности работы гетеролазера, реализуется частично, а поэтому у ОГ-лазера значение порогового тока выше, чем у лазера с двойной гетероструктурой (рис. 6, б). Поскольку удалось уменьшить значение порогового тока у ОГ-лазера, это дало возможность использовать его работу при комнатной температуре, но только в импульсном режиме накачки. В непрерывном режиме накачки при комнатной температуре работают лазеры с двойной гетероструктурой (ДГ).

Толщина активного слоя ДГ-лазера составляет не менее 1 мкм. При этом по всему слою создается инверсная населенность. Если в ОГ-лазерах толщина активного слоя соизмерима с длиной диффузии инжектируемого электрона, то в ДГ-лазерах толщина меньше этой длины. Кроме того, в ДГ-лазерах обеспечивается оптическое ограничение с двух сторон активной зоны. Эти обстоятельства приводят к тому, что ДГ-лазеры являются высокоэффективными приборами и характеризуются минимальным пороговым током, что позволяет осуществлять непрерывную накачку электрическим током при комнатной температуре.

Для улучшения выходных характеристик гетероструктурного лазера в процессе получения гетероструктуры создают условия, обеспечивающие ограничение носителей заряда в активной области. Для структуры, изображенной на рис. 6, б, диаграмма

энергитических зон приведена на рис. 7. Из-за того, что ширина запрещенной зоны у полупроводника больше в области с увеличением концентрацией атомов Al, возникают смещения в зоне проводимости на $p-p^+$ -переходе (dE_c) и в валентной зоне на $n-p$ - и n^+p -переходах (dE_v).

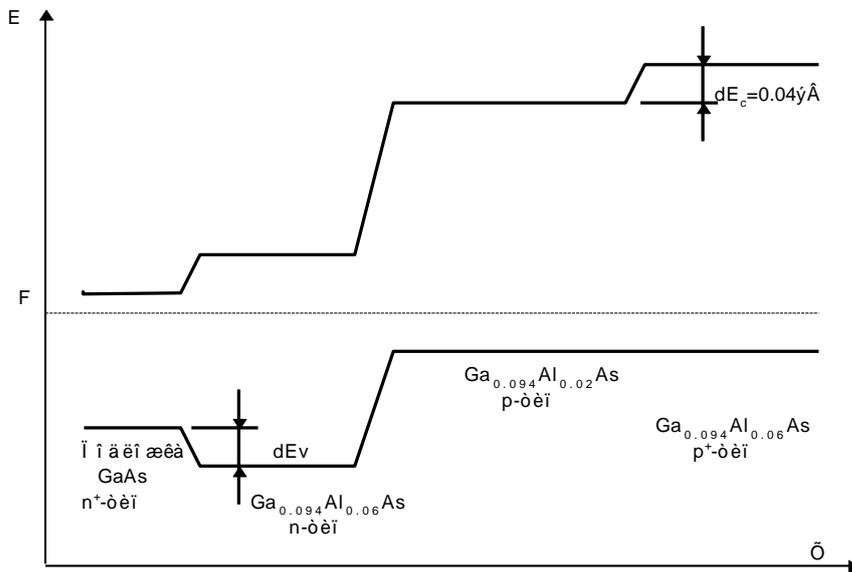


Рис. 7.

Когда к такой структуре прикладывается прямое напряжение смещения, электроны инжектируются из n^+ в p -область. Скачок зоны проводимости на $p-p^+$ -границе раздела на dE_c обеспечивает энергетический барьер для инжектируемых электронов, производя тем самым ограничение их в p -области и увеличивая вероятность их рекомбинации с дырками. Скачок валентной зоны на $n-p$ -переходе dE_v повышает уже существующий потенциальный барьер, препятствующий инжекции дырок в n -область, улучшая тем самым инжекционную эффективность. Таким образом, у двойной гетероструктуры имеет место тенденция ограничения как основных, так и инжектируемых неосновных носителей в активной зоне. Это обеспечивает хорошие условия для получения более эффективной инверсной населенности. Значит ДГ-лазеры обеспечивают более высокие выходные характеристики по сравнению с ОГ-лазерами, и тем более по сравнению с гомопереходными лазерами. Сравнение технических характеристик показывает, что если у гомоструктурного лазера пороговая плотность тока равна 10^4 А/см² при квантовой эффективности 10%, то у ОГ-лазеров пороговая плотность тока равна 10^3 А/см² при квантовой эффективности 40%. Эти лазеры работают только в импульсном режиме. У ДГ-лазеров пороговая плотность тока равна 700—800 А/см², а квантовой эффективность составляет 55%. Эти лазеры работают в непрерывном режиме.

Однако у ДГ-лазеров большая угловая расходимость луча (20—40°) в плоскости, перпендикулярной к плоскости перехода, из-за дифракции света в тонком активном слое, в то время как у гомоструктурных и ОГ-лазеров угловая расходимость составляет 15—20°. У всех рассмотренных типов лазеров угловая расходимость луча в плоскости перехода составляет не более 10°.

4. Литература.

1). К. И. Крылов, В. Т. Прокопенко, В. А. Тарлыков “Основы лазерной техники “. Машиностроение 1990 год.

2). П. Г. Елисеев “Введение в физику инжекционных лазеров”.

Контрольные вопросы:

1. Основные принципы работы полупроводниковых лазеров.
2. Основные принципы работы лазера на гетероструктурах.
3. Основные принципы работы лазера лазеры и усилители на основе квантоворазмерных эффектов.
4. Структура вертикально-излучающего полупроводникового лазера.
5. Волоконные лазеры и усилители, планарные лазеры и усилители.

3. Формирование, распространение, поглощение и дисперсия световых импульсов в волоконно-оптических линиях, спектральное и временное уплотнение информационных потоков, элементная база оптических линий связи, передача оптических сигналов в атмосфере и космосе

Основные характеристики компонентов волоконно-оптических систем передачи

За последний период развития в области связи, наибольшее распространение получили оптические кабели (ОК) и волоконно-оптические системы передачи (ВОСП), которые по своим характеристикам намного превосходят все традиционные кабели системы связи. Связь по волоконно-оптическим кабелям, является одним из главных направлений научно-технического прогресса. Оптические системы и кабели используются не только для организации телефонной городской и междугородней связи, но и для кабельного телевидения, видеотелефонирования, радиовещания, вычислительной техники, технологической связи и т.д.

Для любой системы важное значение имеют три фактора:

- информационная емкость системы, выраженная в числе каналов связи, или скорость передачи информации, выраженная в бит в секунду;
- затухание, определяющее максимальную длину участка регенерации;
- стойкость к воздействию окружающей среды.

Волоконно-оптические линии связи по сравнению с обычными кабельными линиями имеют следующие преимущества:

- Высокая помехоустойчивость, нечувствительность к внешним электромагнитным полям и практически отсутствие перекрестных помех между отдельными волокнами, уложенными вместе в кабель.
- Значительно большая широкополосность.
- Малая масса и габаритные размеры. Что уменьшает стоимость и время прокладки оптического кабеля.
- Полная электрическая изоляция между входом и выходом системы связи, поэтому не требуется общее заземление передатчика и приемника. Можно производить ремонт оптического кабеля, не выключая оборудования.
- Отсутствие коротких замыканий, вследствие чего волоконные световоды могут быть использованы для пересечения опасных зон без боязни коротких замыканий, являющихся причиной пожара в зонах с горючими и легковоспламеняющимися средами.
- Потенциально низкая стоимость. Хотя волоконные световоды изготавливаются из ультра чистого стекла, имеющего примеси меньше чем несколько частей на миллион, при массовом производстве их стоимость не велика. Кроме того, в производстве световодов не используются такие дорогостоящие металлы, как медь и свинец, запасы которых на Земле ограничены. Стоимость же электрических линий коаксиальных кабелей и волноводов постоянно увеличивается как с дефицитом меди, так и с удорожанием энергетических затрат на производство меди и алюминия.

К недостаткам оптических кабелей можно отнести:

- подверженность волоконных световодов радиации, за счет которой появляются пятна затемнения и возрастает затухание;
- водородная коррозия стекла, приводящая к микротрещинам световода и ухудшению его свойств.

Сравнение волоконно-оптические линии связи с обычными кабельными линиями представлено в таблице 1.

Табл. 1 Сравнение по методу аналогий

Волоконно-оптическая линия связи	Радио/беспроводная/проводная линия связи	Комментарий

Электрооптический преобразователь	Модулятор или формирователь сигнала	Все три случая требуют какого-то преобразования формы сигнала, напр., AMI в NRZ
Источник оптического сигнала	Источник сигнала (передатчик или модем)	Выход источника сигнала, как правило низкого уровня
Волоконно-оптическая среда передачи	Передача радиосигнала через атмосферу или радио / аудио сигнала по медным проводам	
Детектор оптического сигнала	Приемник или демодулятор модема	Порог срабатывания приемника во всех 3 случаях определяет показатели ошибок
Схема формирования выходного сигнала	Выход приемника или модема и формирователь сигнала	

Причины и степень их влияния на ухудшение показателей систем передачи перечислены в таблице 2.

Табл. 2 Сравнение влияния разных причин на ухудшение показателей различных систем передачи

Показатель/ причина	Радиосистемы и беспроводные системы	Проводные системы	ВОПС
BER	1×10^{-9}	1×10^{-10}	1×10^{-12}
Потери линии (дБ)	Принципиальные ухудшения	Принципиальные ухудшения	Принципиальные ухудшения
Дисперсия	Могут быть ухудшения при большой скорости	Не являются первопричиной ухудшения	Могут быть ухудшения при большой скорости
Замирания	Влияют	Нет	Нет

Накопленный джиттер	Умеренно влияет	Сильно влияет	Мало влияет
Незащищенность	Низкая	Средняя	Высокая
Емкость канала	Низкая/средняя	Низкая/средняя	Очень высокая
Потери поглощения при дожде	Основные потери на частоте 10 ГГц	Нет	Нет
ЭМС: чувствительность к электромагнитному излучению	Существует	Существует	Нет
ЭМС: генерация электромагнитного излучения	Существует	До некоторой степени	Нет

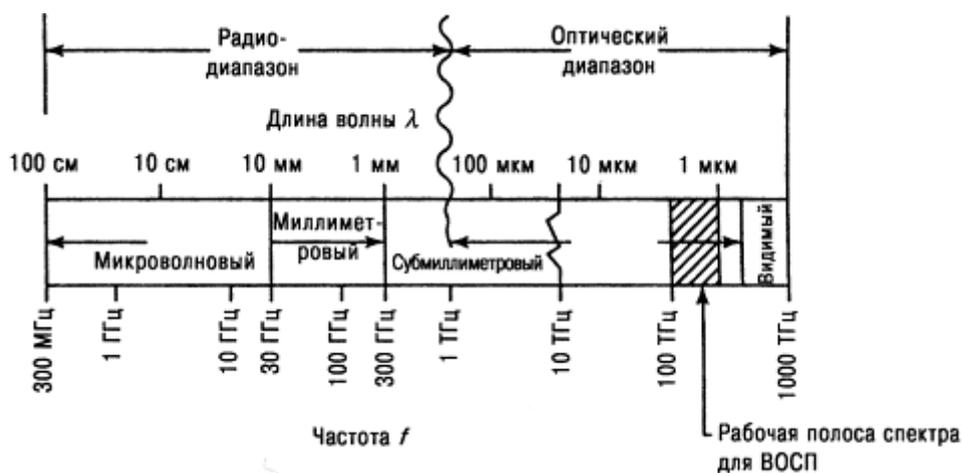


Рис. 1. Частотный спектр выше 300 МГц, где показано положение рабочей области ВОСП

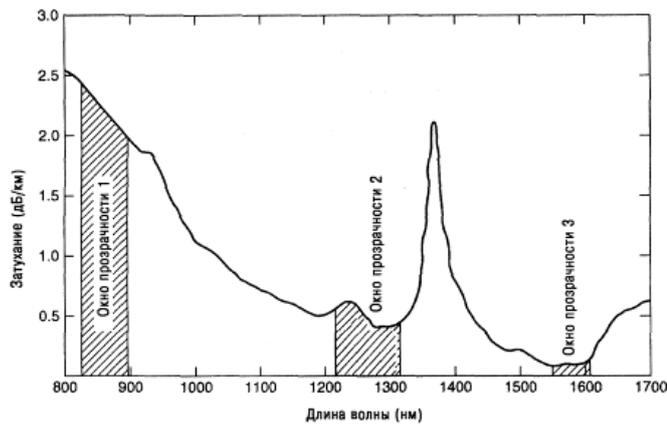


Рис. 2. Затухание оптического волокна в зависимости от длины волны (показаны три окна прозрачности, используемые для ВОСП)

Элементная база современных ВОСП

Основой современных систем ВОЛС являются оптические излучатели, представляющие собой оптические квантовые генераторы света, оптические квантово-электронные фотодетекторы, преобразующие кванты света (фотоны) в поток электронов (электрический ток) и среда распространения оптического излучения — световоды или оптические волокна.

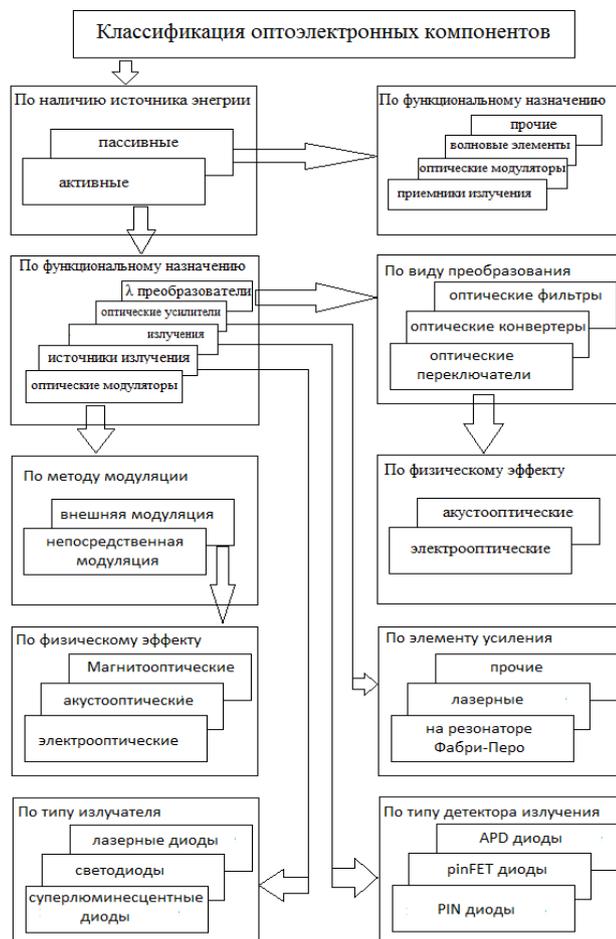


Рис. 3 Классификация оптоэлектронных компонентов

Источники излучения

Основными компонентами источников излучения в ВОСП являются светоизлучающие диоды (СИД) и полупроводниковые инжекционные лазеры. Эти устройства излучают свет в диапазоне длин волн от 0,75 до 1,6 мкм.

Сами устройства представляют кристалл в виде параллелепипеда, торцевые грани которого образуют резонатор Фабри-Перо. Размеры резонатора весьма малы: длина резонатора $L=100\dots 500$ мкм и ширина 100 мкм, ширина активной части области излучения 10 мкм, толщина 1 мкм. Малые габариты этих устройств позволяют выполнять передающие оптические модули в интегральном исполнении, кроме того источники света ВОСП должны обладать большой выходной мощностью, допускать возможность разнообразных типов модуляции света, иметь малые габариты и стоимость, большой срок службы, КПД и обеспечить возможность ввода излучения в оптическое волокно с максимальной эффективностью.

При выборе источника излучения следует учитывать, что диаграмма направленности у лазерных диодов уже, чем у СИД и, соответственно, потери при вводе в волокно излучения выше у СИД. Однако стоимость и срок службы у СИД на два порядка лучше, чем у лазерных диодов.

Светоизлучающие диоды

Светоизлучающие диоды – СИДы, используемые в связи, излучают свет в близкой ИК области. Они недороги, по сравнению с большинством лазеров. Первоначально СИДы использовались с многомодовым волокном, учитывая, что они излучали свет в широком конусе, который мог быть захвачен эффективно только многомодовым волокном, имеющим большую числовую апертуру.

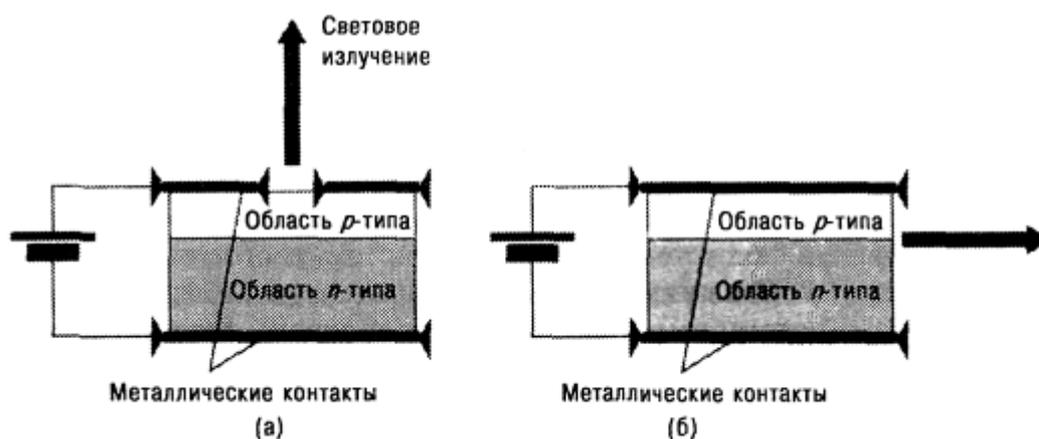


Рис. 4 Поперечный разрез СИДа с излучающей поверхностью (а) и СИДа с излучающим срезом (б)

Лазерные диоды

Если системные требования не так строги, то в качестве источника света обычно выбирается СИД. **Лазерные диоды** обычно используются для линий связи большой длины или при большой скорости передачи (выше 155 Мбит/с).

Существует несколько типов ЛД:

- многомодовые (MLM) или с резонаторами Фабри–Перо;
- одномодовые (SLM);
- одномодовые с распределенной обратной связью (DFB), часто называемые DFB-лазерами;
- DFB-лазеры с внешним модулятором;
- лазеры с вертикальной резонаторной полостью и излучающей поверхностью (VCSEL).

Принцип действия полупроводниковых лазеров (ППЛ) основан на вынужденной излучательной рекомбинации электронно-дырочных пар, в активных полупроводниковых структурах, получаемых при прохождении через такие структуры электрического тока накачки. Наибольшее распространение получили лазеры на гетероструктурах (гетеролазеры), лазеры с распределенной обратной связью (РОС-лазеры) и лазеры на квантоворазмерных структурах (КРС-лазеры).

Современные ППЛ, применяемые в системах оптической связи, обычно работают в спектральных диапазонах высокой прозрачности кварцевого оптоволокна – 0.82–0.90 мкм, 1.30 – 1.33 мкм и около 1.55 мкм. Типичная мощность излучения таких ППЛ от 1 до 5 мВт; увеличение выходной мощности ППЛ для магистральных ВОСПИ сверх 5–10 мВт нецелесообразно, так как срок действия мощных лазеров сравнительно невелик. Кроме этого, при больших плотностях мощности в одномодовом волокне заметную роль начинают играть нелинейно-оптические явления – ВКР, ВРМБ и др., приводящие к искажениям передаваемых сигналов. Ширина спектра излучения лучших образцов промышленных полупроводниковых лазеров около 0.1 нм при уровне боковых частот ниже 20 дБ. В одночастотных ППЛ, используемых в системах когерентной оптической связи, полуширина спектра генерации менее 500 МГц. Величина порогового тока накачки РОС-лазеров составляет несколько десятков мА, у ППЛ на основе квантоворазмерных структур пороговый ток накачки существенно меньше – от 3 до 15 мА. Как правило, ППЛ для ВОЛС выпускаются в виде компактных оптоэлектронных модулей, содержащих сервисную электронику (усилитель подводимого сигнала, систему авторегулировки мощности, температуры и др.), фотоприемник для контроля мощности выходного излучения, терморезистор и

полупроводниковый термоэлемент (элемент Пельтье) – «холодильник», управляемый специальным электронным устройством и поддерживающий стабильную рабочую температуру внутри модуля. В таком модуле излучение из активной области ППЛ с использованием микролинз вводится в выходное одномодовое или многомодовое оптоволокно. В последние годы выпуск отдельных излучательных полупроводниковых модулей для ВОСПИ становится ограниченным, и гораздо большее распространение в технике оптической связи находят приемо-передающие оптоэлектронные модули, содержащие в едином компактном блоке полупроводниковые излучатель и фотоприемник.

Частота модуляции современных коммерческих высокоскоростных ППЛ составляет от нескольких десятков – сотен МГц до примерно 1.5 – 2.5 ГГц. В уникальных образцах ППЛ достигнута скорость передачи сигналов свыше 25 Гбит/с.

В последние годы повышенный интерес разработчиков ВОСПИ вызывают «викселы» – полупроводниковые лазеры с вертикальным резонатором (VCSEL – vertical-cavity surface-emitting lasers). В таких лазерах резонатор образован двумя объемными дифракционными решетками Брэгга и излучение генерируется в направлении, перпендикулярном плоскости подложки, являющейся основанием гетеро- и квантоворазмерных слоев полупроводников. Благодаря сверх-короткой длине L резонатора Фабри-Перо, викселы генерируют на одной продольной моде, при этом диаметр выходного пучка лазера достигает 20–30 мкм, что позволяет осуществлять его эффективную фокусировку в одномодовое волокно. Пороговый ток викселов крайне мал – до 2–5 мА, мощность излучения около 1 мВт.

Приложением электрического поля в направлении оси резонатора в некоторых (консольных) вариантах VCSEL удастся осуществлять плавную перестройку длины волны генерации в полосе длин волн 1530–1560 нм. Это создает возможность эффективного применения викселов в перспективных ВОСПИ со спектральным уплотнением (WDM и DWDM).

Табл. 3 Сравнительные характеристики настраиваемых лазеров

Тип лазера	Преимущества	Недостатки	Применение
С распределенной обратной связью (DFB)	стабильность излучения в процессе производства	относительно низкая выходная мощность; ограниченный диапазон настройки	для узкого диапазона настройки; широко используется в длинных секциях
С распределенным брэгговским отражателем (DBR)	высокая скорость переключения	большая ширина линии; нестабильность длины волны	сети доступа; оптические мультиплексоры ввода-вывода
Типа (DBR) с выбранными решетками	широкий диапазон настройки; высокая скорость переключения	низкая выходная мощность; большая ширина линии; отсутствие непрерывной настройки	сети доступа; региональные сети; оптические мультиплексоры ввода – вывода
С вертикальной полостью и излучающей поверхностью (VCSEL)	малая ширина линии при О/Р; низкое потребление; круговой луч; широкий диапазон настройки	низкая выходная мощность при Е/Р; ограничен коротковолновыми диапазонами 850/1310 нм	региональные сети и сети доступа
С внешними резонансными микрополостями	высокая мощность; малая ширина линии; низкий уровень RIN; непрерывность и широкий диапазон настройки	скорость настройки чувствительна к ударам и вибрациям	длинные и ультра-длинные секции; OADM для региональных сетей; сети с коммутацией

Основные параметры промышленных источников светового излучения

Светоизлучающие диоды (СИД)

Выходная мощность СИД, в зависимости от конструкции и производителя, лежит в пределах между 0,01 и 0,1 мВт (от -20 до -10 дБм).

Существуют две нежелательные особенности у СИД, которые проектировщик системы должен принимать во внимание:

1. Угловая ширина излученного пучка:
 - а) для СИД с излучающей поверхностью – порядка 120°;
 - б) для СИД с излучающим срезом – порядка 30°.
2. Ширина спектра излучения: 30–80 нм.

СИДы производятся для работы в трех окнах прозрачности: 850, 1310 и 1550 нм. Однако, чаще всего они используются в первых двух окнах: 850 и 1310 нм.

СИДы существенно дешевле в производстве, чем ЛД; они имеют значительно больший срок службы и не требуют стабилизации температуры при работе в нормальных условиях. Что же касается условий светопередачи (на стыке СИД-волокно), то их эффективность мала. Однако, из двух указанных типов, СИД с излучающим срезом значительно более эффективен, чем СИД с излучающей поверхностью. Их применение обычно ограничено системами, работающими на скорости 155 Мбит/с и ниже.

В ВОСП небольшой (0.1–1.0 км) протяженности, а также в низкоскоростных (не боле 10 Мбит/с) ВОСП в качестве источников света нередко используются полупроводниковые светодиоды, отличающиеся сравнительно малой мощностью излучения (до 0.5 мВт) и большой (около 20 – 30 нм) шириной спектра излучения. В основе действия светоизлучающих полупроводниковых диодов (СИД) лежит спонтанная рекомбинация электронно-дырочных пар в активной области гетеро-либо квантоворазмерной полупроводниковой структуры.

Диапазон рабочих длин волн современных СИД весьма широк – от 0.4 до 1.6 мкм. СИД активно используются и в технике систем индикации, в осветительной и сигнальной технике; разработаны СИД с мощностью излучения до 30 мВт в диапазонах длин волн около 690, 590, 470 нм, а также различные типы СИД белого свечения. Степень поляризации выходного излучения СИД близка к нулю. СИД обычно используются в линиях передачи сигналов на основе многомодовых волоконных световодов с большим (десятки-сотни мкм) диаметром сердцевины, в том числе изготовленных из недорогих полимерных материалов. Удобно применение СИД и в открытых системах связи с дальностью действия в пределах 100 м, используемых на промышленных объектах и строительных площадках. Характерная особенность СИД,

применяемых в системах связи, – линейная ватт-амперная характеристика в широком диапазоне токов накачки, что делает их весьма удобными в аналоговых оптоэлектронных системах передачи и обработки сигналов. Основные достоинства СИД – малая потребляемая электрическая мощность, дешевизна и значительная долговечность (около 10^5 часов).

Табл. 4 Сводка основных параметров СИД с излучающим срезом

Параметр	Значение
Выходная мощность, излучаемая в одномодовое волокно (25°C)	2–50 мкВт 0,1 – 0,6
Числовая апертура (NA)	3 нс (максимум)
Время нарастания / спада импульса	30–60 нм
Ширина полосы излучения на уровне половины мощности	1,2%/°C (типовое)
Температурный коэффициент мощности	0,5–0,8 нм/°C
Изменение центральной длины волны с температурой	0,4 нм/°C (типовое)
Спектральное уширение	

Лазерные диоды (ЛД)

Полупроводниковые лазерные диоды, как правило, используются в качестве источников света в высокоскоростных (> 155 Мбит/с) системах дальней связи. Выходная мощность лазерных источников для большинства производителей составляет порядка +3 – +10 дБм (1–10 мВт). Лазерные диоды с большей мощностью (напр., +20 дБм) также начали появляться на рынке. Сейчас достаточно поставить на выходе лазерного источника оптический усилитель, например, типа EDFA чтобы увеличить его выходную мощность до 500 мВт или выше. Одной из причин, заставляющих производителей выпускать лазеры с большей мощностью на выходе, является широкое использование систем DWDM (плотного мультиплексирования по длине волны). Линейные компоненты систем DWDM имеют высокие вносимые потери. Использование таких высоких уровней выходной мощности вместе с оптическими усилителями помогает преодолеть потери в системах DWDM. Лазеры, работающие на таких больших уровнях мощности, принесли ряд проблем, в частности проблему влияния большой мощности на оптические компоненты тракта передачи. Изоляторы, используемые в качестве универсального устройства снижения световой энергии, отраженной в сторону такого лазерного источника, имеют возвратные потери > 80 дБ, что позволяет эффективно снижать отраженную мощность.

Табл. 5 Сравнение характеристик СИД с характеристиками некоторых типов лазерных диодов

Параметр	СИД типа ELED	FP (MLM) – лазер	DFB-лазер с внешним модулятором	Лазер типа VCSEL
Длина волны, нм	850/1310	1310/1550	1550	850/1310
Мощность на стыке с ОВ, дБм	-10 ... -15	0,0	0,0	0,0
Спектральная ширина линии, нм	30–60	< 3	<0,1	< 3
Максимальная скорость, Гбит/с	< 0,155	>2	> 10	2
Тип волокна	многомодовое	одномодовое	со сдвигом дисперсии	многомодовое или одномодовое
Цена	низкая	умеренно высокая	высокая	умеренная
MTBF, часы	10^9	10^8	10^7	10^8

Сравнение ширины спектра излучения СИД и ЛД

На рис. 5 приведены для сравнения формы спектральных линий СИД и лазерного диода.

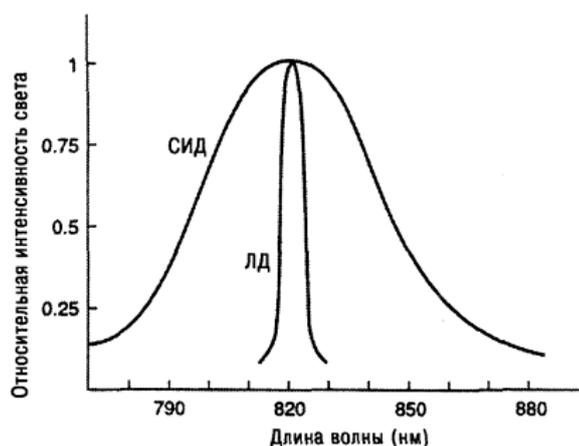


Рис. 5 Типичные спектры линий, излучаемых СИД и ЛД. Амплитуды спектральных линий были нормализованы (приведены к одному значению).

Фактически пиковая интенсивность ЛД много больше, чем у СИД.

Фотодетекторы ВОСП

Фотодетекторы – это устройства, которые преобразуют оптические сигналы в электрические идентичной формы. Существуют различные типы детекторов, работающие на основе пиро-, термо- или фотоэлектрических эффектов. Приемники для ВОЛС обычно представляют собой фотодетекторы, т.е. фотоэлектрические устройства. В волоконно-оптической связи в качестве фотоприемников используют только рпн- и лавинные фотодиоды.

Основными требованиями предъявляемыми к фотодетектору являются: фотодетектор должен воспроизводить форму принимаемого оптического сигнала, не внося дополнительного шума, то есть обладать требуемой широкополосностью, динамическим диапазоном и чувствительностью. Кроме того, он должен иметь малые размеры (но достаточные для надежного соединения с оптическим волокном), большой срок службы и быть не чувствительным к изменениям параметров внешней среды.

Существующие фотодетекторы далеко не полно удовлетворяют перечисленным требованиям. Наиболее подходящими среди них для применения в волоконно-оптических системах передачи являются полупроводниковые р-і-п фотодиоды и лавинные фотодиоды (ЛФД). Они имеют малые размеры и достаточно хорошо стыкуются с волоконными световодами. Достоинством ЛФД является высокая чувствительность (может в 100 раз превышать чувствительность р-і-п фотодиода), что позволяет использовать их в детекторах слабых оптических сигналов. Однако, при использовании лавинных фотодиодов нужна жесткая стабилизация напряжения источника питания и температурная стабилизация, поскольку коэффициент лавинного

умножения, а следовательно фототок и чувствительность ЛФД, сильно зависит от напряжения и температуры.

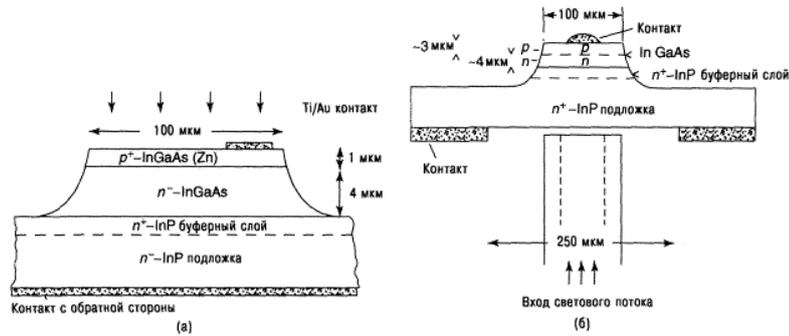


Рис. 6 Обобщенная схема PIN-диодного детектора на основе InGaAs. (а) вход с фронта; (б) вход с подложки (с тыла).

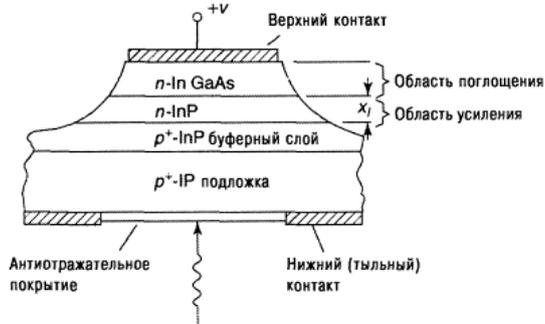


Рис. 7 Схема поперечного сечения структуры APD.

Оптические кабели

Современные проводные системы передачи строятся с использованием практически только оптических кабелей (ОК). Основа оптических кабелей – оптические волокна. В зависимости от назначения, условий прокладки и эксплуатации разработаны и производятся ОК разных типов и конструкций.

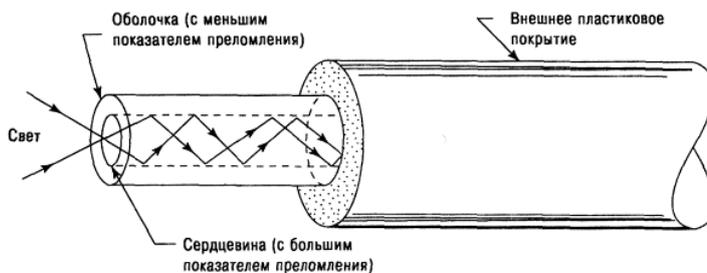
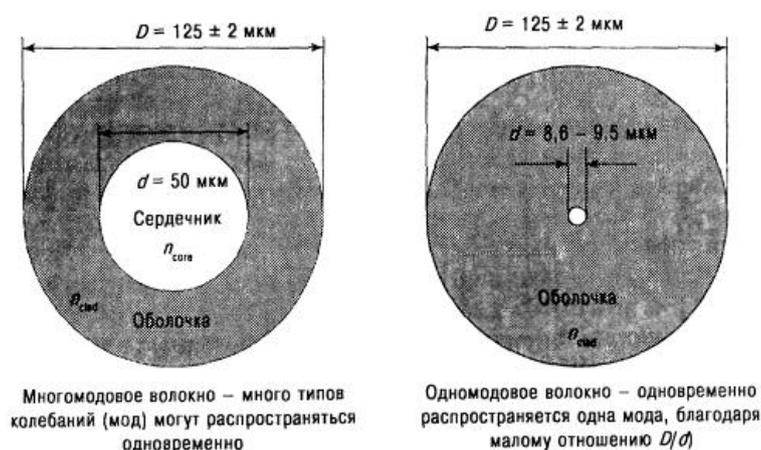


Рис. 8 Структура оптического кабеля

Существуют три основных типа оптического волокна (ОВ), отличающихся числом мод и своими физическими свойствами (считают, что этих типов два: одномодовое и многомодовое):

- одномодовое волокно;
- многомодовое волокно со ступенчатым профилем показателя преломления;
- многомодовое волокно с градиентным профилем показателя преломления.

Одномодовые волокна подразделяются на ступенчатые одномодовые волокна (step index single mode fiber) или стандартные волокна SF (standard fiber), на волокна со смещенной дисперсией DSF (dispersion-shifted single mode fiber) и на волокна с ненулевой смещенной дисперсией NZDSF (non-zero dispersion-shifted single mode).



Обратим внимание на то, что внешний диаметр обоих типов волокон одинаков и составляет номинально 125 мкм. Однако существует огромная разница в диаметрах сердцевин: 50 мкм для многомодового волокна и 8,6–9,5 мкм для одномодового волокна. На практике существуют и другие значения диаметров многомодового волокна, наиболее используемым из них является 62,5 мкм.

Каждое волокно состоит из сердцевины и оболочки с разными показателями преломления. Сердцевина, по которой происходит распространение светового сигнала, изготавливается из оптически более плотного материала. При обозначении волокна указываются через дробь значения диаметров сердцевины и оболочки. Волокна отличаются диаметром сердцевины и оболочки, а также профилем показателя преломления сердцевины. У многомодового градиентного волокна и одномодового волокна со смещенной дисперсией показатель преломления сердцевины зависит от

радиуса. Такой более сложный профиль делается для улучшения технических характеристик или для достижения специальных характеристик волокна.

Если сравнивать многомодовые волокна между собой, то градиентное волокно имеет лучшие технические характеристики, чем ступенчатое, по дисперсии. Главным образом это связано с тем, что межмодовая дисперсия в градиентном многомодовом волокне – основной источник дисперсии – значительно меньше, чем в ступенчатом многомодовом волокне, что приводит к большей пропускной способности у градиентного волокна.

Одномодовое волокно имеет значительно меньший диаметр сердцевины по сравнению с многомодовым и, как следствие, из-за отсутствия межмодовой дисперсии, более высокую пропускную способность. Однако оно требует использования более дорогих лазерных передатчиков.

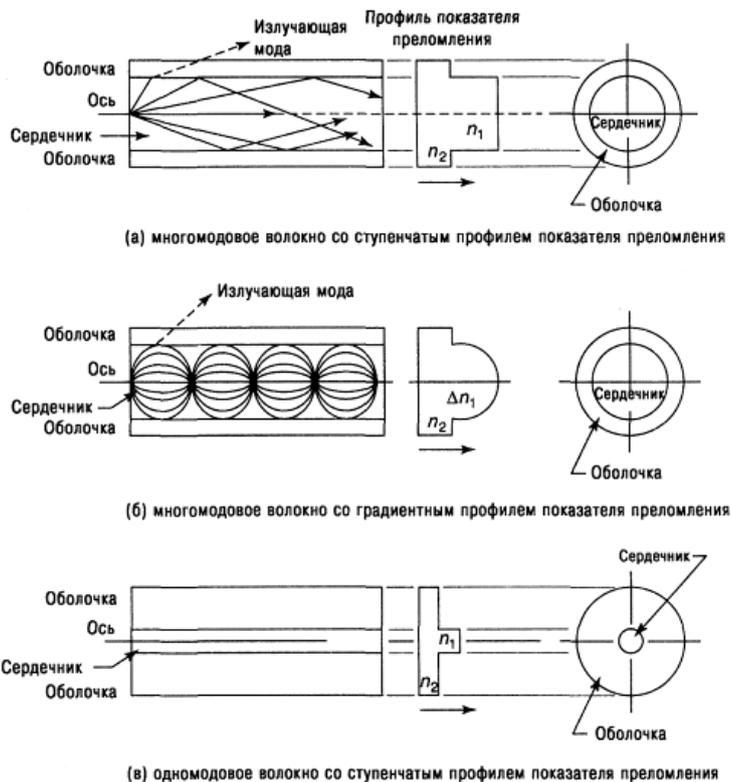


Рис. 10 Профили показателей преломления и моды, распространяющиеся в трех типах ОВ

В волоконно-оптических линиях связи ВОЛС наиболее широко используются следующие стандарты волокон:

- многомодовое градиентное волокно 50/125;
- многомодовое градиентное волокно 62,5/125;
- одномодовое ступенчатое волокно SF (волокно с несмещенной дисперсией или стандартное волокно) 8–10/125;
- одномодовое волокно со смещенной дисперсией DSF 8–10/125;
- одномодовое волокно с ненулевой смещенной дисперсией NZDSF (по профилю показателя преломления это волокно схоже с предыдущим типом волокна).

В стандартном многомодовом градиентном волокне (50/125 или 62,5/125) диаметр световедущей жилы 50 или 62,5 мкм, что на порядок больше длины волны передачи. Это приводит к распространению множества различных типов световых лучей – мод – во всех трех окнах прозрачности. Два окна прозрачности 850 и 1610 нм обычно используют для передачи света по многомодовому волокну.

В ступенчатом одномодовом волокне (SF) диаметр световедущей жилы составляет 8 – 10 мкм и сравним с длиной световой волны. В таком волокне при достаточно большой длине волны света распространяется только один луч (одна мода). Одномодовый режим в одномодовом волокне реализуется в окнах прозрачности 1310 и 1550 нм. Распространение только одной моды устраняет межмодовую дисперсию и обеспечивает очень высокую пропускную способность одномодового волокна в этих окнах прозрачности. Наилучший режим распространения с точки зрения дисперсии достигается в окрестности длины волны 1310 нм, когда хроматическая дисперсия обращается в ноль. С точки зрения потерь это не самое лучшее окно прозрачности. В этом окне потери составляют 0,3 – 0,4 дБ/км, в то время как наименьшее затухание 0,2 – 0,25 дБ/км достигается в окне 1550 нм.

Типичные характеристики оптического волокна высокого качества

Компания Corning Fiber Systems выпустило волокно LEAF, имеющее большую эффективную активную площадь волокна. Оно является идеальным для использования в широкополосных системах DWDM, работающих в окне прозрачности 1550 нм.

Табл. 6 Оптические характеристики волокна LEAF компании Corning

Характеристика/параметр	Значение	Комментарий
Ослабление при длине волны 1550 нм при длине волны 1625 нм	$\leq 0,25$ дБ/км $\leq 0,25$ дБ/км	
Точка разрыва непрерывности	не больше, чем 0,1 дБ при 1550 нм	
Затухание при длине волны 1383 нм	$\leq 1,0$ дБ/км	
Прирост затухания в диапазоне 1525–1575 нм	$\leq 0,05$ дБ/км	По отношению к длине волны 1550 нм
Прирост затухания за счет изгиба	$\leq 0,05$ дБ/км	На оправке диаметром 32 мм, 1 оборот, по отношению к 1550 и 1625 нм
Прирост затухания за счет изгиба	$\leq 0,50$ дБ/км	На оправке диаметром 75 мм, 100 оборотов, по отношению к 1550 и 1625 нм
Диаметр поля моды	9,2–10,0 мкм при 1550 нм	
Полная дисперсия	2,0–6,0 пс/нм/км	В диапазоне 1530–1565 нм
Поляризационная модовая дисперсия (PMD)	4,5–11,2 пс/нм/км	В диапазоне 1565–1625 нм
Дисперсия PMD для протяженной линии	$\leq 0,08$ пс (км) ^{1/2}	
Максимальная PMD для отдельного волокна	$\leq 0,20$ пс (км) ^{1/2}	

Табл. 7 Физические размеры волокна LEAF

Характеристика/параметр	Значение
Стандартная длина	4,4 – 25,2 км/катушку
Радиус собственной кривизны волокна	$\geq 4,0$ м
Диаметр оболочки	125 ± 1 мкм
Неконцентричность сердцевины и оболочки	$\leq 0,5$ мкм $\leq 1,0\%$
Некруглость оболочки	245 ± 5 мкм
Диаметр покрытия	$< 12,0$ мкм
Неконцентричность покрытия и оболочки	

Оптические разветвители

Каждое устройство, используемое в схеме передачи светового сигнала, является источником вносимых потерь. Оно также будет источником отражений, обычно характеризующихся потерями на отражение. Эти потери обычно измеряются в децибелах. За исключением аттенуаторов, хотелось бы иметь как можно более низкие вносимые потери и как можно более высокие потери на отражение (возвратные потери).

Одним из наиболее важных устройств, относящихся к пассивным компонентам ВОЛС является оптический разветвитель.

Оптический разветвитель (сплиттер – комбайнер)

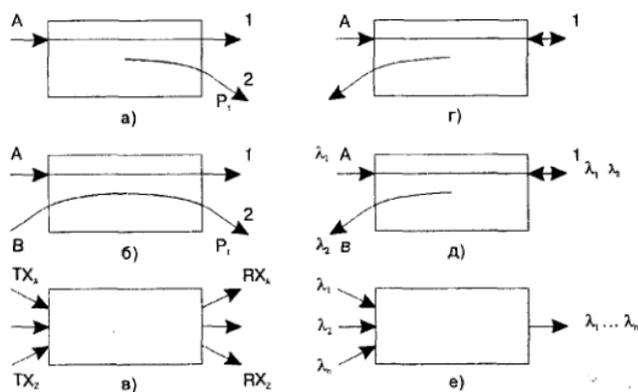


Рис. . Наиболее распространенные конфигурации ответвителей

Ответвителем является неселективный пассивный элемент, обладающий тремя или более портами и распределяющий мощность между ними в определенном соотношении без какого-либо усиления, переключения или какой-то модуляции.

Различают направленные и двунаправленные разветвители, а также разветвители чувствительные к длине волны и нечувствительные. В двунаправленном оптическом разветвителе каждый полюс может работать на прием сигнала или на передачу, или осуществлять прием и передачу одновременно, так что группы приемных и передающих полюсов могут меняться местами в функциональном смысле.

Аттенюатор

Аттенюатор – пассивный элемент, осуществляющий управляемое ослабление сигнала в волоконно-оптической линии передачи (ВОЛП).

Аттенюатор – устройство, которое уменьшает интенсивность светового сигнала, прошедшего через него. Аттенюаторы часто используются в качестве звена в схеме после лазерного передатчика, чтобы согласовать его выходную мощность с уровнем, требуемым следующими за ним в этой схеме устройствами, такими как усилители EDFA. Нужно тщательно выбирать тип аттенюаторов так, чтобы они имели отличные параметры возвратных потерь (их уровень д. б. > 40 дБ), чтобы быть уверенным, что уровень света, отражаемого обратно в направлении передатчика, будет очень низок.

Другими ключевыми параметрами аттенюаторов являются:

- стабильность;
- надежность;
- оптические возвратные потери (ORL);
- потери, зависящие от поляризации (PDL);
- точность;
- повторяемость;
- вносимые потери;
- поляризационная модовая дисперсия (PMD).

Волоконно-оптический фильтр

Фильтр – пассивный элемент, используемый для модификации проходящего через него оптического излучения, как правило, за счет изменения спектрального распределения мощности. В частности, волоконно-оптические фильтры обычно используют для отсека или поглощения оптического излучения в определенных областях длин волн и пропускания излучения на других длинах волн.

Оптические фильтры используют механизм селекции длин волн и могут быть грубо разделены на две большие категории в зависимости от того, какой физический механизм положен в их основу: оптической интерференции или дифракции. Существуют фильтры, рассчитанные на выделение фиксированной длины волны, и настраиваемые фильтры.

Для настраиваемых фильтров желательно иметь следующие свойства:

- широкий диапазон настройки для максимизации числа каналов, которые могут быть выбраны;
- незначительные перекрестные помехи, чтобы избежать помех от соседних каналов;
- быстрая настройка для минимизации времени доступа;
- малые вносимые потери.

Фильтры выполняют исключительно важную роль в оборудовании WDM/DWDM на стороне демультиплексора.

Стандарт рекомендует следующие значения параметров фильтров:

Вносимые потери: максимум 1,5 дБ в полосе пропускания;

Вносимые потери: минимум 40 дБ в полосе задерживания;

Оптическая отражательная способность: -40 дБ.

Волоконно-оптический изолятор

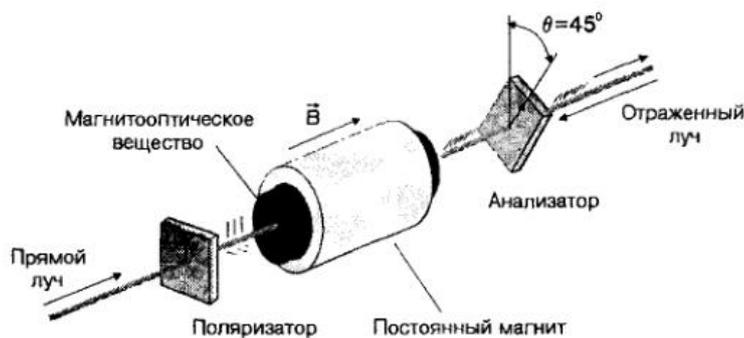


Рис. Оптический изолятор

Изолятор – оптическое устройство, не обладающее свойством взаимности, предназначенное для подавления обратного отражения в ВОЛП и имеющее минимальные вносимые потери в прямом направлении.

Изоляторы – устройства, для которых потери света малы при распространении в одном направлении и велики в противоположном направлении. Изоляторы обычно

устанавливают в выходных схемах устройств с высоким уровнем интенсивности света, таких как передатчики на лазерных диодах и усилители EDFA. Их функция – уменьшить уровень сигнала, отраженного назад в используемый лазерный диод или усилитель EDFA.

Характеристики изолятора определяются следующими критическими параметрами:

- спектральной зависимостью, особенно для так называемых узкополосных изоляторов, которые проектируются для работы в спектральном диапазоне уже, чем 20 нм. Изоляторы описываются пиком ослабления обратного излучения и шириной полосы, лежащей в области 3 дБ ослабления уровня изоляции от максимума этого пика.

- малым уровнем вносимых потерь, <1 дБ в прямом направлении, и большим уровнем потерь в обратном направлении: больше 35 дБ (при одноступенчатой изоляции) и 60 дБ (при двухступенчатой изоляции), и слабой зависимостью от поляризации.

- поляризационной модовой дисперсией (PMD). Изоляторы обычно проектируются на основе использования элементов с высоким уровнем двойного лучепреломления, а они весьма склонны к высокому уровню PMD (типичное значение – 50–100 фс, $1 \text{ фс} = 10^{-15} \text{ с}$), в особенности для одноступенчатого изолятора. Двухступенчатые изоляторы могут быть спроектированы так, что PMD, вносимая первой ступенью, во многом компенсируется второй ступенью.

- потерями, зависящими от поляризации (PDL). Они ухудшают характеристики оптического изолятора.

В лучших образцах ОИ оптические потери в прямом направлении равны около 1.0–1.5 дБ, величина развязки – до 30–40 дБ. Оптические изоляторы обычно действуют в сравнительно узкой (30–80 нм) полосе спектра рабочих длин волн и согласованы с входным и выходным световодами, изготовленными из оптоволокна с сохранением поляризации излучения.

Волоконно-оптический терминатор

Терминатор (оконечный элемент) – элемент, используемый для терминирования оптоволокна (оконцованного или нет) с целью подавления отражения.

Волоконно-оптический переключатель (коммутатор)

Переключатель (коммутатор) – пассивный элемент, имеющий один или больше портов, которые передают, блокируют или перенаправляют оптическую мощность в одно из волокон ВОЛП.

Ключевыми параметрами, определяющими показатели коммутаторов и, следовательно, их пригодность для тех или иных приложений, являются:

- вносимые потери и потери на разветвление;
- возвратные потери;
- потери, зависящие от поляризации;
- перекрестные помехи и уровень изоляции;
- надежность;
- время переключения;
- стабильность;
- степень сложности.

Оптические коммутаторы, в зависимости от принципа действия, подразделяются на восемь типов: механические ОК, электрооптические, термооптические, интегрально-оптические, полупроводниковые (SOA), ОК на фотонных кристаллах и ОК на многослойных жидкокристаллических матрицах.

Технические характеристики современных ОК значительно различаются. Например, механические коммутаторы имеют время переключения от 10 до 500 мс, вносимые потери около 0.5 дБ, переходное затухание до 80 дБ, число входных и выходных портов от 50 до 1600. Быстродействие электрооптических ОК существенно выше, 10–100 пс, но их емкость обычно мала, например, 1x2 и 2x2. При последовательном включении нескольких таких ОК емкость составного ОК, очевидно, возрастает. В последние годы внимание разработчиков оптических коммутаторов привлекают микромеханические коммутаторы (MEMS-типа), в которых использованы технологии как микроэлектроники, так и интегральной оптики.

Табл. 8 Параметры передачи, рекомендуемые для оптических коммутаторов.

Параметры	Переключатели типа 1×n		Переключатели типа 2×2	
	Максимум	Минимум	Среднее	Стандартное
Вносимые потери, дБ	2,5 //log ₂ n	не применимо	изучается	не применимо
Отражательная способность, дБ	-40	не применимо	-40	не применимо
Рабочий диапазон волн, нм	изучается	изучается	изучается	изучается
Потери,	изучается //	не применимо	изучается	не применимо

зависящие от поляризации, ΔдБ	0,1 (1 + log ₂ n)			
Время переключения, мс	20 // 10	не применимо	изучается	не применимо
Повторяемость, дБ	0,25	не применимо	изучается	не применимо
Однородность, дБ	изучается // 0,4 log ₂ n	не применимо	изучается	не применимо
Переходные помехи, дБ	не применимо	Изучается	изучается	не применимо
Направленность, дБ	не применимо	50	изучается	не применимо

Пассивный компенсатор (хроматической) дисперсии

Компенсатор – пассивный элемент, используемый для компенсации хроматической дисперсии одного из оптических трактов.

Волоконно-оптический соединитель (оптический разъем)

Оптический разъем – элемент, прикрепленный к одному из оптических кабелей или отдельной части оборудования для осуществления частых соединений / разъединений оптических волокон или кабелей.

Оптические мультиплексоры и демультиплексоры.

Оптические мультиплексоры и демультиплексоры (МДМ) применяются в оптических системах связи со спектральным уплотнением каналов (с одновременной передачей по оптоволокну сигналов с различными длинами волн – WDM-systems) и их можно отнести к категории спектрально – селективных оптических разветвителей. Мультиплексоры выполняют функции объединения оптических несущих с разными длинами волн (обычно они имеют несколько входов и один выход), демультиплексоры – обратную процедуру разделения оптических несущих по длинам волн (один вход, несколько выходов).

Основные характеристики МДМ – тип (одномодовый либо многомодовый), число рабочих спектральных каналов, оптические потери в каждом канале, величина переходного затухания, уровень обратного рассеяния, уровень вносимых поляризационных искажений. Диапазон рабочих длин волн и величина разноса спектральных несущих также являются важными параметрами МДМ. В наиболее

освоенных сегодня WDM-системах число рабочих каналов обычно составляет 2, 4 или 6, разнос длин волн несущих – несколько единиц нм, оптические потери в каналах 0.5 – 2 дБ (отметим, что теоретически потери в спектрально-селективных разветвителях, в отличие от широкополосных, могут приближаться к нулю). В разрабатываемых «плотных» DWDM-системах спектрального уплотнения число рабочих каналов – несколько десятков (до 30 и более), разнос частот несущих несколько единиц-десятков ГГц, оптические потери в канале на уровне 1 дБ.

Конструктивно мульти-демультиплексоры нередко выполнены на основе объемных микрооптических элементов (микролинз, градианов) и содержат спектрально-селективные устройства – отражательную вогнутую либо плоскую дифракционную решетку, интерференционный фильтр, объемную голограмму, дисперсионную призму и др. Волноводные МДМ, например, на основе интегрально-оптических волноводов, часто представляют собой аналоги классических интерферометров – Фабри-Перо, Маха-Цендера, эшелона Майкельсона и др. В последнее время внимание разработчиков

МДМ привлекают волоконные решетки Брэгга (FBG), являющиеся, в сущности, объемными отражательными голограммами, сформированными в сердцевине оптоволокна.

Внешние оптические модуляторы.

Наиболее распространенным способом амплитудной модуляции оптического сигнала в ВОСПИ является изменение тока накачки излучателя. При больших, свыше 1–2 ГГц, частотах модуляции тока накачки в полупроводниковых лазерах заметную роль начинает играть квазирезонансный эффект, сопровождающийся амплитудными шумами, значительным уширением спектра оптического сигнала и ограничением величины частоты модуляции. Поэтому при создании ВОСПИ с наивысшей, в десятки-сотни и более ГГц/с, пропускной способностью наиболее эффективным способом модуляции сигнала является применение в ВОСПИ внешних оптических модуляторов (ВОМ). В этом случае полупроводниковый лазер работает в непрерывном режиме, а внешний оптический модулятор установлен в оптическом тракте непосредственно за оптическим изолятором и источником света.

Действие наиболее распространенных внешних оптических модуляторов основано на электрооптическом (Поккельса), магнитооптическом (Фарадея), электроабсорбционном и акустооптическом (Брэгга или Рамана-Ната) эффектах. Такие ВОМ работоспособны практически во всех диапазонах оптической связи с частотами модуляции свыше 1 ГГц. Амплитудные ВОМ могут быть выполнены на основе

интегрально-оптического интерферометра Маха-Цендера с фазовой модуляцией в одном из плеч интерферометра.

Разрабатываются высокоскоростные ВОМ и на основе полупроводниковых многослойных структур, а также МОП-структур в виде тонких слоев металла и диэлектрика. Сегодня созданы экспериментальные ВОЛС протяженностью 160 км на основе ВОМ со скоростью передачи в одном канале 111 Гбит/с, что при использовании принципа спектрального уплотнения обеспечивает эквивалентную пропускную способность ВОСПИ порядка 10 Тбит/с.

Оптические усилители.

Оптический усилитель (ОУ) увеличивает мощность оптических сигналов в ВОСПИ без их промежуточного преобразования в электронные сигналы. Такой способ усиления особенно удобен в ВОСПИ со спектральным уплотнением, благодаря возможности одновременного усиления нескольких информационных сигналов с различными длинами волн. Наиболее распространены два типа оптических усилителей – полупроводниковый квантовый усилитель (ПКУ) и волоконно-оптический (например, эрбиевый) усилитель.

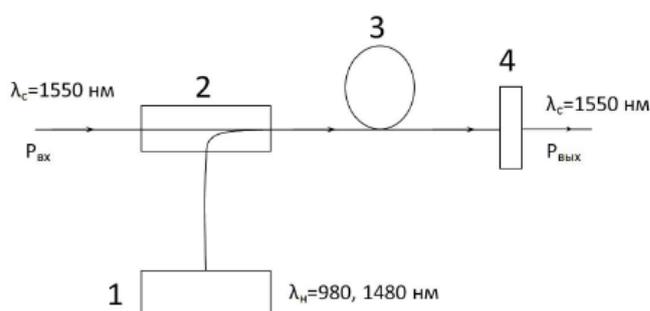


Рис. . Принципиальная схема волоконно-оптического усилителя. 1 – источник накачки (мощный полупроводниковый лазер с длиной волны излучения 980 нм или 1480 нм), 2 – мультиплексор, 3 – эрбиевое оптоволокно, 4 – полосовой оптический фильтр с пропусканием на длине волны усиливаемого сигнала 1550 нм.

Определение функциональных параметров

Вносимые потери (ИЛ)

Вносимые потери – уменьшение оптической мощности между входным и выходным портами пассивного элемента в дБ, определяемое как

$$IL = -10 \log(P_1 / P_0)$$

где P_0 – оптическая мощность, вводимая во входной порт, а P_1 – оптическая мощность, полученная из выходного порта.

Возвратные потери (RL)

Возвратные потери – часть входной мощности, которая возвращается из входного порта пассивного элемента. Они определяются, как

$$RL = -10\log(P_r / P_i)$$

где P_i – оптическая мощность, вводимая во входной порт, а P_r – оптическая мощность, полученная обратно из того же порта.

Отражательная способность

Отражательная способность – отношение R отраженной мощности P_r к падающей мощности P_i , определенное для данного порта пассивного элемента при заданных условиях спектрального распределения, поляризации и геометрического распределения, выраженное в дБ, а именно:

$$R = -10\log(P_r / P_i)$$

Рабочий диапазон длин волн

Это диапазон длин волн от $\lambda_{i\min}$ до $\lambda_{i\max}$ в пределах, задаваемых от номинального λ_j , внутри которого пассивные элементы должны работать с определенными показателями ошибок.

Потери, зависящие от поляризации (PDL)

Эти потери соответствуют максимальной вариации вносимых потерь, вызванной вариацией состояния поляризации, рассматриваемой на множестве всех возможных состояний поляризации.

Зависимость отражательной способности от поляризации

Эти потери соответствуют максимальной вариации отражательной способности, вызванной вариацией состояния поляризации, рассматриваемой на множестве всех возможных состояний поляризации.

Обратные потери (степень изоляции) волоконно-оптических изоляторов

Обратные потери – мера уменьшения оптической мощности (в дБ), распространяющейся в обратном направлении, в результате установки изолятора.

Излучающим портом здесь является выходной порт изолятора, а приемным портом – входной порт изолятора. Потери определяются следующей формулой:

$$BL = -10 \log P_{ob} / P_{ib}$$

где P_{ob} – оптическая мощность, измеренная на входном порте изолятора, когда мощность P_{ib} излучается в рабочий порт. При нормальной работе P_{ib} – оптическая мощность, отраженная от устройств, установленных на удаленном конце оптической линии, и направленная обратно так, что попадает в выходной порт изолятора, потери которого и измеряются.

Направленность

Для волоконно-оптических элементов ветвления, направленность представлена значением a_{ij} – элемента логарифмической матрицы передачи между двумя изолированными портами.

Однородность

Логарифмическая матрица передачи элементов ветвления может содержать определенный набор коэффициентов, который конечен и одинаков. В этом случае диапазон изменения этих коэффициентов a_{ij} (выраженный в дБ) именуется однородностью элементов ветвления.

Заключение

Оптические системы передачи информации являются одним из наиболее перспективных современных направлений в области техники связи, вобравшим в себя лучшие достижения микроэлектроники, волоконной оптики, интегральной оптоэлектроники, физики и техники полупроводников. Научные проблемы освоения оптического диапазона связи к настоящему времени, в значительной степени, решены и дальнейшее развитие оптических систем передачи информации существенно зависит от уровня и состояния технологии производства оптических и оптико-электронных компонент таких систем. Это не исключает возможности выдвижения и реализации новых идей в области физики и техники оптических систем передачи информации, основанных на весьма разнообразных свойствах как оптического излучения, так и применяемых в таких системах оптических материалов, их сложных композиций и структур.

Требования к полосе пропускания удваиваются каждые три года. Только оптоволокно может удовлетворить транспортировке требуемой полосы.

Список использованной литературы

1. Р. Фриман «Волоконно-оптические системы связи» М. 2003 Техносфера
2. Иванов А.Б. «Волоконная оптика: компоненты, системы передачи, измерения» М. 1999
3. А.Л. Дмитриев «Оптические системы передачи информации» Спб. 2007
4. В.В. Богатырева, А.Л. Дмитриев «Оптические методы обработки информации» Спб. 2009

Контрольные вопросы:

- 1.Формирование световых импульсов в волоконно-оптических линиях.
- 2.Распространение световых импульсов в волоконно-оптических линиях.
3. Поглощение и дисперсия световых импульсов в волоконно-оптических линиях.
- 4.Спектральное и временное уплотнение информационных потоков.
- 5.Элементная база оптических линий связи.
6. Передача оптических сигналов в атмосфере и космосе.

4. Локальная и распределенная запись информации, оптические дисковые системы записи и хранения информации, магнитооптические технологии, голографические технологии, регистрирующие среды и механизмы записи, быстроедействие, считывание информации в реальном времени - динамическая голография, ассоциативная голографическая память

Аналоговые оптические вычисления, Фурье-голография, голографическая коммутация, мультиплексирование и демultipлексирование сигналов, оптическая би- и мультистабильность, цифровая оптическая обработка сигналов Технологии создания и перспективы применения

Нынешний век – век ускоренного развития технологий. В эту эпоху информационных технологий каждый из нас хотя бы раз задумывался о том, как же все-таки работает то или иное новое устройство. Все эти изобретения сменяют друг друга с огромной скоростью. Однако технологии и принципы работы устройств довольно долгое время остаются неизменными, лишь периодически совершенствуясь.

Наряду с научными и техническими применениями лазеры используются в информационных технологиях для решения специальных задач, причем эти применения широко распространены или находятся в стадии исследований. Наиболее распространенными примерами таких применений являются оптическая цифровая память, оптическая передача информации, лазерные печатающие устройства, кроме того, они применяются в вычислительной технике в качестве различных устройств.

Целью данного реферата является раскрытие физики явления такого процесса как оптический принцип записи и считывания информации. Задачей является ознакомление с принципом оптической записи и считывания информации с точки зрения физики.

1. Лазеры в вычислительной технике

Принципиально достигнутые малые времена переключения делают возможным применение лазеров и комбинаций с лазерами, включая интеграцию в микроэлектронных переключательных схемах (оптоэлектроника):

- в качестве логических элементов (да-нет, или);
- для ввода и считывания из запоминающих устройств в вычислительных машинах.

В этих целях рассматриваются исключительно инжекционные лазеры.

Преимущества таких элементов: малые времена переключения и считывания, очень маленькие размеры элементов, интеграция оптических и электрических систем.

Достижимыми оказываются времена переключения примерно 10-10 с (соответственно этому быстрые времена вычисления); емкости запоминающего устройства 10^7 бит/см², и скорости считывания 10⁹ бит/с.

2. Оптическая цифровая память

Для становящейся все более тесной связи между обработкой данных, текста и изображения необходимо применять новые методы записи информации, к которым предъявляются следующие требования:

- более высокая емкость запоминающего устройства;
- более высокая эффективность хранения архивных материалов,
- лучшее соотношение между ценой и производительностью.

Это может быть достигнуто с помощью записи и считывания цифровой информации.

Принцип действия. Информация (речь, музыка, изображения, данные), содержащиеся в виде электрических сигналов, преобразуется в цифровые величины и выражается тем самым в виде последовательности импульсов, которая записывается в различной форме (в виде углублений или отверстий различной длины и расстояний между ними или магнитным способом) на диске запоминающего устройства.

При считывании считывающий свет, отраженный (рассеянный в обратном направлении) от этих углублений (отверстий), модулируется и с помощью фотоприемника преобразуется в соответствующий электрический сигнал.

Лазерно-оптическое считывание информации. С помощью этого способа в приборе, аналогичном проигрывателю, воспроизводится неконтактным способом записанная на диске информация (диаметр дисков до 30 см), причем применяются лазерные диски только для считывания, например видеодиски, компакт-диски.

Принцип действия. Кодирование информации происходит путем создания информационных микроуглублений, имеющих различную длину и различные расстояния между ними. Информация на диске сохраняется, таким образом, в цифровой форме, записанной по спирали, которая состоит из информационных ямок (рис. 1).

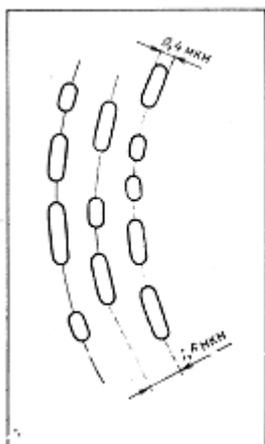


Рис. 1. Схематическое изображение микроуглублений на лазерном диске; ширина углублений 0,4 мкм, расстояние между дорожками 1,6 мкм.

Лазерный видеодиск характеризуется следующими параметрами:

- расстояние между двумя профилирующими дорожками 1,6 мкм;
- ширина углубления 0,4 мкм;
- максимальная длина углубления 3,3 мкм;
- минимальная длина углубления 0,9 мкм;
- максимальное расстояние между углублениями 3,3 мкм;
- минимальное расстояние между углублениями 0,9 мкм.

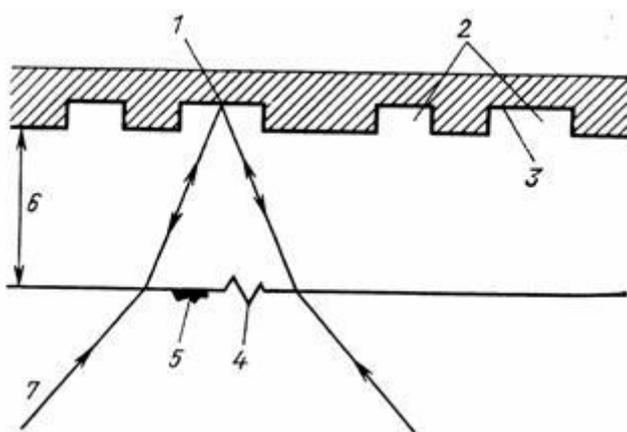


Рис. 2. Сечение видеодиска и грампластинки с лазерной записью:

1 - фокальное пятно ($\lambda \gg 1$ мкм); 2 - структура микроуглублений; 3 - зеркальное покрытие; 4 - царапина; 5 - частица пыли; 6 - прозрачный защитный слой; 7 - луч от лазера

При изготовлении видеодисков нанесенный прежде на подложку из стекла фотолак экспонируется с помощью специальной оптической системы излучением коротковолнового лазера (криптоновый лазер, $\lambda = 0,35$ мкм). После этого следует многоступенчатый процесс проявления, в результате которого образуется образцовый диск, который используется затем для изготовления других дисков путем оттиска. На полученные после отделения от образцового диска оттиски наносится зеркальное покрытие и слой лака, так что полученные при записи микроуглубления не могут быть закрыты частицами пыли. Пыль и царапины на защитном слое не мешают, поскольку они находятся вне плоскости фокусировки считывающей оптики (рис. 2).

При считывании микроскопических маленьких структур используются эффекты дифракции и интерференции света. Оптическая считывающая система для видеодисков состоит из:

- He-Ne-лазера (мощность мВт), который излучает линейно поляризованный свет;
- делителя пучка, который разделяет свет на три пучка с соотношениями интенсивностей 1:3:1 (дифракционная решетка. Работающая на просвет с минус первым, нулевым и плюс первым порядками дифракции);
- призмы Волластона (оптическая длина пути зависит от направления поляризации);
- пластинки $1/4$;
- считывающего объектива, перемещаемого по принципу катушки с подвижным сердечником в направлении оптической оси (ограниченный дифракцией микрообъектив очень малой массы);
- системы фотоприемников (квадратных приемников), а также цилиндрической линзы.

Рассеянный в обратном направлении от диска свет лазерного пучка отображается на квадратном приемнике, лучи, использованные для слежения за дорожкой, попадают на приемники (рис. 3)

Таким образом, становится возможным формирование управляющих сигналов для корректной фокусировки считывающих лучей на информационной дорожке и обеспечение слежения за дорожкой.

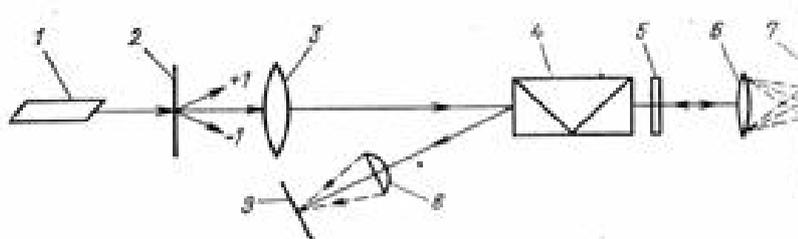


Рис. 3. Оптическая схема считывающей головки для считывания информации, записанной на видеодиске

1 - He-Ne-лазер; 2 - решетка; 3 - согласующая оптика; 4 - призма Волластона; 5 - пластинка $1/4$; 6 - считывающий объектив; 7 - видеодиск; 8 - цилиндрическая линза; 9 - плоскость приемника.

Оптическая считывающая головка для цифрового лазерного проигрывателя. Обратно рассеянный от лазерной пластинки свет попадает на фотодиоды F1-F4. Возникающие при этом фототоки комбинируются друг с другом таким образом, что становится возможным получение как управляющих сигналов для радиальной коррекции, так и управляющего сигнала для установки на резкость считывающей оптики (рис. 4).

Радиальный управляющий сигнал формируется комбинацией токов фотодиодов $(F1+F2) - (F3+F4)$. Если считывающий объектив сфокусирован на информационную плоскость диска, то после призм 4 появляются два резких изображения между фотодиодами F1, F2, а также F3, F4. Если фокальная плоскость считывающего объектива находится за или перед информационной плоскостью, то изображения становятся нерезкими и движутся друг к другу или друг от друга. Тогда с помощью комбинации токов фотодиодов $(F1+F2) - (F3+F4)$ может быть получен управляющий сигнал для установки на резкость считывающей головки.

Однократная запись информации. Этот принцип позволяет осуществить однократную запись и многократные считывания информации. Для этого на нижней стороне очень плоской стеклянной пластины наносится слой теллура. Две круглые

стеклянные пластины юстируются относительно друг друга таким образом, что слои теллура защищены снаружи стеклянными пластинами.

На слоях теллура, находящихся на внутренних сторонах пластин, записывается информация. Пластины снабжены спиральной дорожкой (спиральной канавкой глубиной примерно $1/4$), которая служит для юстировки считывающего или записывающего луча. При записи одного бита информации в слое теллура импульсно повышается мощность полупроводникового лазера за время 50 нс до 12 мВт, при этом в слое возникает отверстие диаметром примерно 1 мкм. Запись и считывание осуществляются с помощью одинакового устройства, причем при считывании мощность полупроводникового лазера уменьшается до 1 мВт (рис. 5).

С помощью таких методов записи и считывания достигаются емкости запоминающего устройства (диаметр диска 30 см) 10^{10} бит информации (передняя и задняя сторона); свободно выбираемые времена доступа составляют 150 мс.

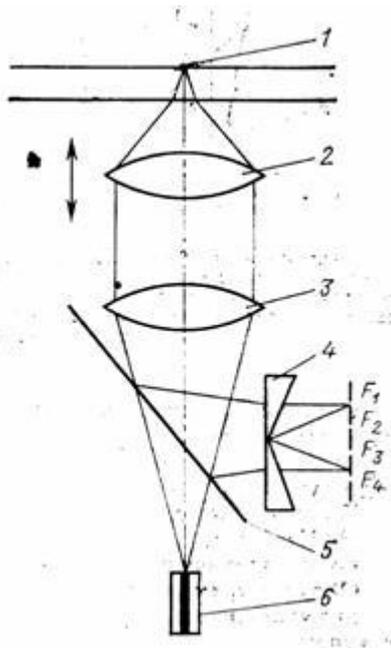


Рис. 4. Схема оптической считывающей головки для лазерных пластинок.

1 - считывающее пятно; 2 - считывающий объектив; 3 - оптическая система для преобразования излучаемого полупроводниковым лазером волнового поля в плоское волновое поле; 4 - призма; 5 - полупрозрачное зеркало; 6 - полупроводниковый лазер; F1 - F4 - фотоприемники.

Применяемые лазеры:

- He-Ne-лазер;
- полупроводниковый лазер (все более часто).

Области применения:

- запоминающее устройство для хранения банка данных с частым доступом;
- запоминающее устройство для хранения архивных данных с отсроченным доступом;
- внешнее дополнительное запоминающее устройство со свободно выбираемой адресацией в ЭВМ;
- видеодиски для обучения;
- видеодиски для библиотек и архивов;
- запоминающие диски для управления и канцелярского дела;
- аудиодиски с высококачественным воспроизведением звука.

Оптическая цифровая запись информации в магнитных слоях. В качестве носителя информации используется тонкий магнитооптический слой (преимущество: повторная запись данных).

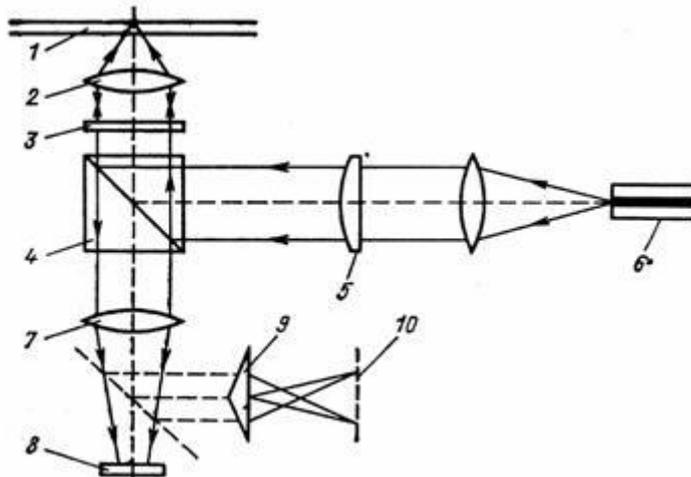


Рис. 5. Схема записывающей и считывающей головки для однократной записи:

1 - лазерный диск; 2 - считывающий объектив; 3 - пластинка 1/4; 4 - зависящий от поляризации делитель пучка; 5 - цилиндрическая линза; 6 - полупроводниковый

лазер; 7 - оптическая система; 8 - приемник для радиального контроля дорожки; 9 - призма Френеля; 10 - приемник для получения сигнала и контроля положения фокального пятна.

Принцип действия. Запись информации происходит благодаря тому, что маленькие области магнитного слоя нагреваются с помощью сфокусированного лазерного луча, причем одновременно накладывается магнитное поле, напряженность которого меньше, чем коэрцитивная сила. В нагретых таким образом при наложенном магнитном поле областях исчезает намагниченность (запись точки Кюри). Считывание осуществляется таким же лазером при уменьшенной мощности, причем плоскость поляризации отраженного от диска света в зависимости от направления намагничивания маленьких областей поворачивается на величину $0,5 - 8$ град (в зависимости от магнитооптического слоя) (магнитооптический эффект Керра).

Оптическое устройство записывающей и считывающей головки аналогично системам, используемым в описанных выше устройствах считывания и записи информации.

Дополнительно следует обратить внимание на рис. 6.

Свет, отраженный от маленьких перемагниченных областей, является эллиптически поляризованным и с помощью соответствующей фазовой пластинки преобразуется в линейно поляризованный. Линейно поляризованный свет разделяется на две составляющие, которые могут регистрироваться отдельно. Оба принятых сигнала подаются на дифференциальный усилитель и усиливаются. Усиленный сигнал прямо пропорционален поляризационному эффекту Керра.

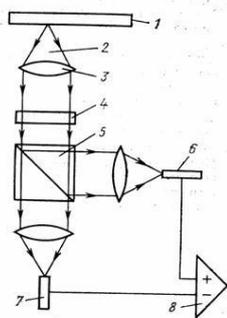


Рис. 6. Схема получения сигнала с помощью поляризационного эффекта Керра

1 - магнитный диск; 2 - отраженный свет; 3 - микрообъектив; 4 - фазовая пластинка; 5 - делитель пучка; 6 - приемник Nr2; 7 - приемник Nr1; 8 - дифференциальный усилитель.

Магнитооптическая запись позволяет в настоящее время иметь:

- емкость памяти запоминающего устройства 10^5 бит/см²;
- число циклов (запись, считывание, стирание) 10^6 ;
- свободно выбираемые времена доступа 150 мс;
- применение в качестве оперативной памяти в ЭВМ.

Оптический цифровой метод записи требует максимальной оптической и механической точности, а также:

- предельно малого ограниченного дифракцией считывающего объектива;
- считывающего объектива (микрообъектива) очень малой массы (0,6 г и меньше)
- радиальных отклонений считывающего объектива с точностью ± 1 мкм;
- ширины распределения интенсивности считывающего пятна по половине интенсивности примерно 1 мкм.

Цифровое оптическое запоминающее устройство позволяет производить неразрушающее считывание накопленной информации.

Вывод

Не смотря на то, что технологии оптической записи и считывания уже более 30 лет, она до сих пор актуальна и в наши дни. С годами она совершенствуется: появляются диски все большей емкости, скорость записи растет. К тому же хранение данных на оптических носителях не является затратным, что дает дополнительное преимущество использования данной технологии.

Список литературы

1. <http://www.radioland.net.ua>
2. <http://habrahabr.ru/>

В своих экспериментах по записи первых осевых голограмм Д. Габор (1948) использовал излучение ртутной лампы с низкой степенью когерентности. После изобретения в 1960 году Т. Мейманом рубинового лазера началась эпоха лазерных высококогерентных источников света. В 1961 году А. Джаваном с сотрудниками был создан He-Ne лазер, Е. Снитцером первый неодимовый лазер, а в 1964 году Е. Гордон с сотрудниками получили непрерывную генерацию в аргоне. Изобретение лазеров дало новый импульс исследованиям в области голографии. В 1962 году Е. Лейт и Дж. Упатниекс получили голограммы с внеосевым взаимным расположением интерферирующих пучков. В том же 1962 году Ю.Н. Денисюк предложил способ записи голограмм во встречных пучках. Современные лазеры для голографии генерируют излучение практически во всем видимом спектральном диапазоне с высокой степенью временной и пространственной когерентности. Лазеры могут быть непрерывного действия и импульсными ($t_{\text{импульса}} < 10^{-4}$ сек). Использование для записи голограмм импульсных лазеров позволяет осуществлять запись подвижных объектов или быстропротекающих процессов, характерные времена которых меньше длительности лазерного импульса. Кроме этого, при записи импульсных голограмм нет необходимости обеспечивать высокую степень виброустойчивости оптической схемы записи и отдельных её элементов, что характерно в случае использования лазеров непрерывного действия. Во время записи голограммы происходит «запоминание» регистрирующей средой взаимного расположения «стоячих» волн, образованных при интерференции так называемой опорной волны с объектной, т.е. рассеянной объектом волной. Пример такой картины интерференции плоской опорной и сферической объектной волны приведен на рис.1, где указаны также характерные положения регистрирующей среды при записи голограмм различного типа. Период интерференционной картины (d) зависит от угла между опорным и объектным пучками (2Θ), а также длины волны (λ) излучения при записи и определяется по формуле Брэгга (1912): $2 d \sin\Theta = \lambda$

Дифракционная эффективность голограммы характеризует её способность восстановить объектную волну (I_1) при освещении голограммы опорной волной (I_R) и определяется как свойствами регистрирующего материала, так и оптической схемой записи голограммы.

В таблице 1 приведены расчетные и экспериментальные значения максимальной дифракционной эффективности для голограмм различного типа при амплитудной или фазовой модуляции.

Таблица 1. Максимальная дифракционная эффективность голограмм различного типа

Тип голограммы	“Тонкая”		“Толстая” (или объёмная)			
	Пропускающая				Отражательная	
	Ампл.	Фазовая	Ампл.	Фазовая	Ампл.	Фазовая
$ДЭ_{\text{макс}}(\text{теор.})$	6.25	33.9	3.7	100	7.2	100
$ДЭ_{\text{макс}}(\text{эксп.})$	6.0	32.6	3.0	90	3.8	99.9

1,2,3 – голограммы в попутных пучках. 4 – голограммы во встречных пучках («отражательные»), 1 – осевые голограммы, 2 – внеосевые «тонкие» голограммы, 3 – внеосевые «толстые» голограммы.

Голографический метод записи получил широкое распространение для изготовления оптических элементов (голографических оптических элементов, ГОЭ) самого различного назначения: фокусирующих систем, проекционных экранов, переключателей и т.д. Голографические методы используют в интерферометрии, для хранения информации (оптическая память), изготовления защитных наклеек. К настоящему времени осуществлена запись полноцветных голографических изображений, реализована идея голографического кинематографа. В настоящее время продолжают работы по внедрению голографических методов в тех областях науки и техники где предъявляются повышенные требования к информативной ёмкости полученного изображения.

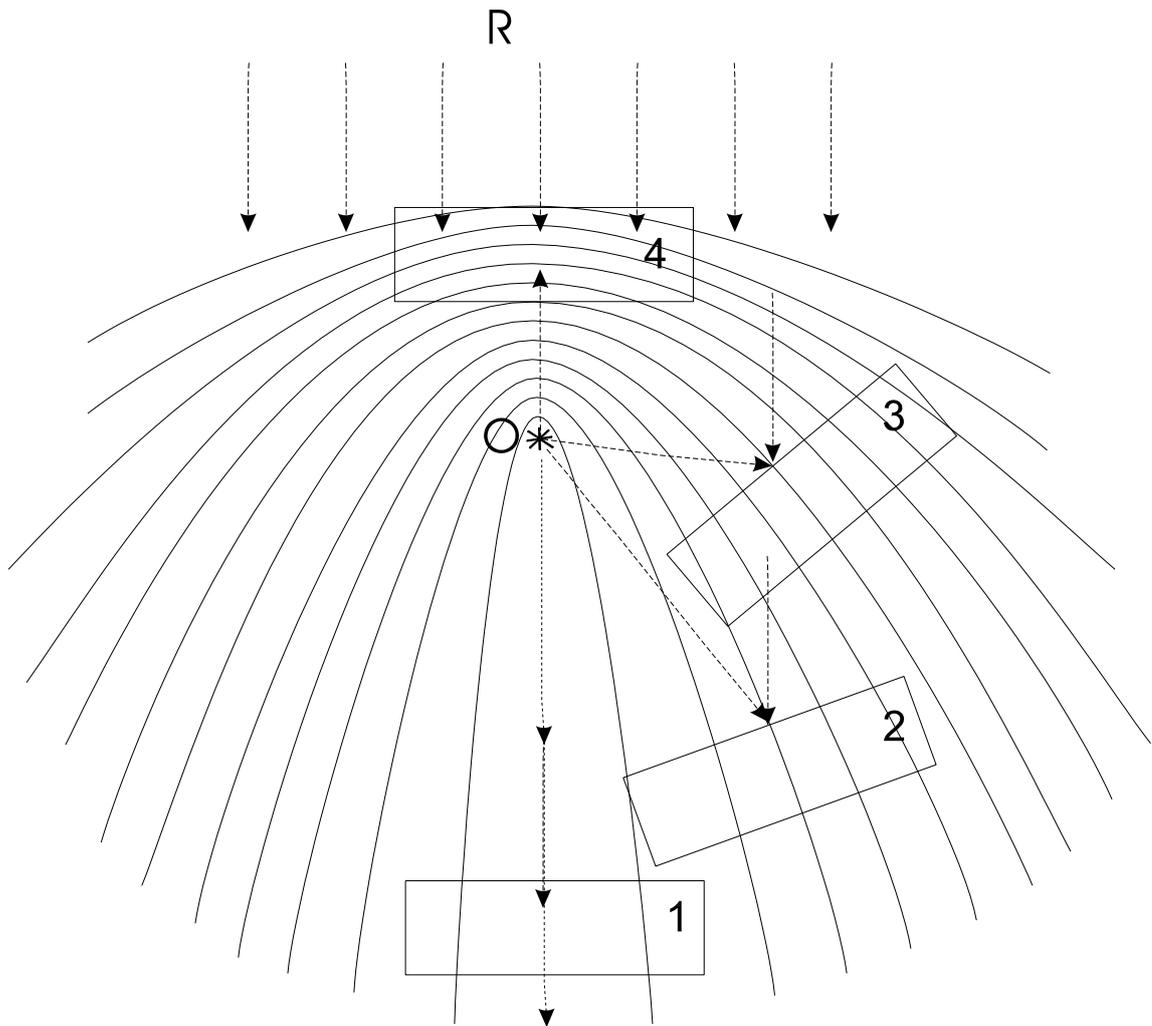


Рис. 1. Поперечное сечение максимумов интенсивности стоячих волн, образованных при интерференции плоской R и сферической волны O . 1,2,3,4 – характерные положения регистрирующей среды, соответствующие различным схемам записи голограмм.

Оптическая схема импульсной лазерной системы приведена на рис. 2.

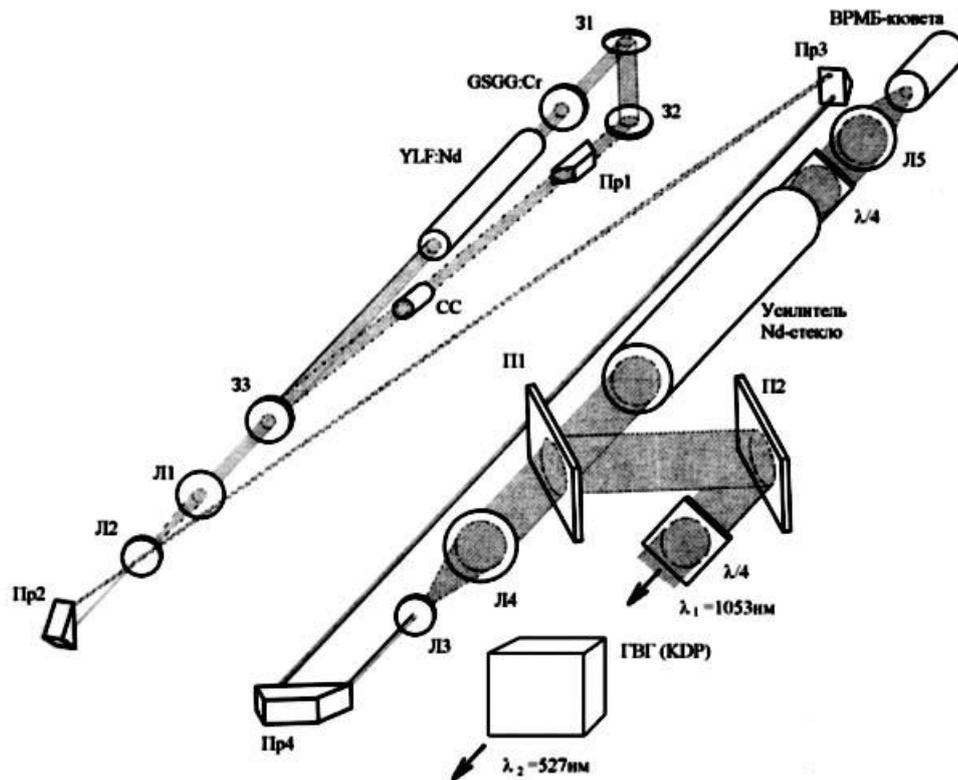


Рис. 2. Оптическая схема импульсной лазерной системы Green Star.

Зеркала 31, 32 и 33 образуют кольцевой резонатор задающего генератора импульсной лазерной системы, генерирующего на длине волны 1053 нм. Внутри резонатора находятся активный элемент – кристалл YLF:Nd, пассивный модулятор добротности на основе кристалла галлий-скандиевого граната, активированного ионами хрома (GSGG:Cr), и элементы СС и Пр1, обеспечивающие однонаправленный режим генерации в кольцевом резонаторе. С помощью линз Л1 – Л4, и призм Пр2 – Пр4 лазерный пучок расширяется и попадает на вход усилителя на основе неодимового стекла. После прохода по усилителю, лазерное излучение отражается от нелинейно-оптического элемента обращающего волновой фронт – ВРМБ-зеркала (пластинка $\lambda/4$, Л5 и кювета ВРМБ) и проходит второй раз по усилителю. На втором проходе по усилителю помимо усиления происходит коррекция aberrаций лазерного стержня усилителя, вследствие обращения волнового фронта. На выходе из усилителя, лазерное излучение выделяется поляризационным зеркалом П1 и после зеркала П2 и пластинки $\lambda/4$ поступает на кристалл калий-титанил фосфата (КТР), где происходит удвоение частоты лазерного излучения.

С использованием лазерного излучения с длиной волны второй гармоники 527 нм осуществляется запись голограмм в попутных (рис. 3) и во встречных (рис. 4) пучках.

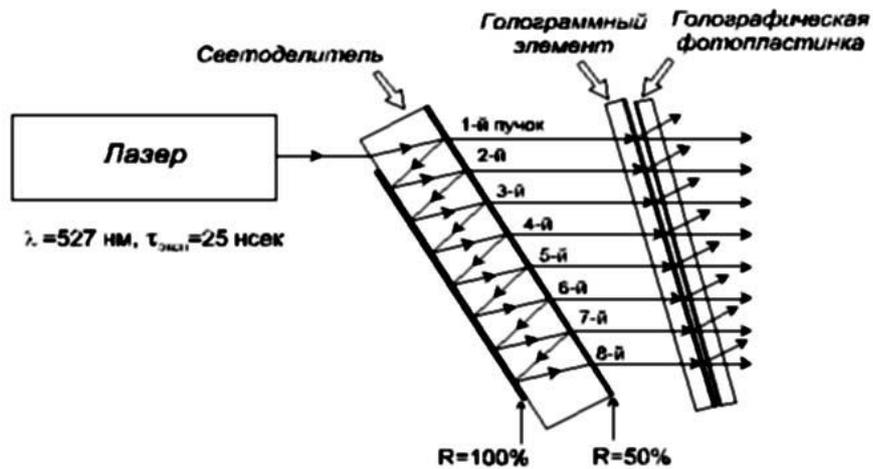


Рис. 3. Оптическая схема записи голограмм в попутных пучках.

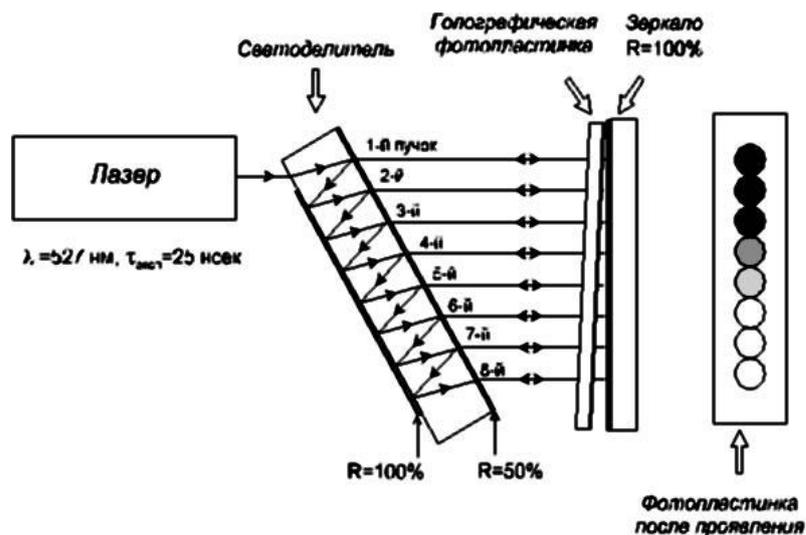


Рис. 4. Оптическая схема записи голограмм во встречных пучках.

Схема измерений характеристик голограмм приведена на рис. 5.

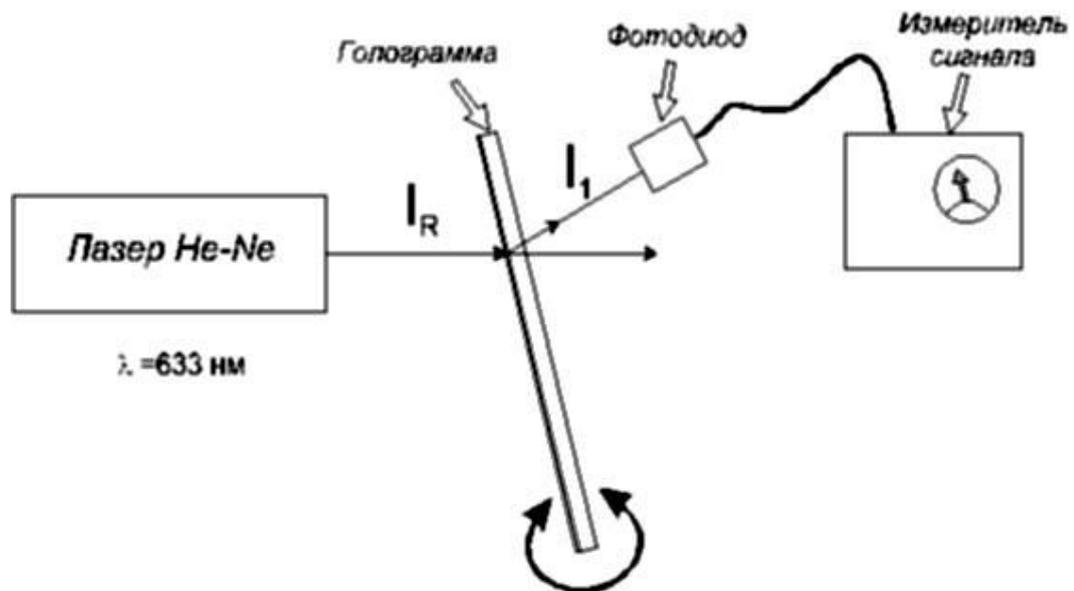


Рис. 5. Измерение дифракционной эффективности и контраста голограмм.

Дифракционная эффективность голограммы характеризует её способность восстановить объектную волну (I_1) при освещении голограммы опорной волной (I_R) и может быть измерена экспериментально и рассчитана по формуле:

$$ДЭ = \frac{I_1}{I_R} \times 100\%$$

В голографии для оценки качества регистрирующего материала часто используют зависимость ДЭ от экспозиции в опорном пучке H_R при заданном соотношении объектного и опорного пучков m ($m = H_0/H_R$). Рассмотрение экспериментальной зависимости $ДЭ = f(H_R)$ позволяет выбрать оптимальные значения экспозиции H_R , а также оценить величину максимального фазового набега для данного типа регистрирующего материала, которая может быть вычислена с использованием следующих выражений:

а) для случая зеркального («плоского») объектного пучка

$$ДЭ_{\text{зерк}} \approx T \times \varphi_0^2 \times \varepsilon^2 \times m \times \exp(-2\varepsilon)$$

б) для случая диффузного («рассеивающего») объектного пучка

$$ДЭ_{\text{диф}}(\varepsilon) = \frac{T \times \varphi_0^2 \times \varepsilon^2 \times m \times \exp(-2\varepsilon)}{(1 + 2\varepsilon \times m)^2}$$

где

T - амплитудное пропускание голограммы

φ_0 – максимальный фазовый набег

m – соотношение пучков при записи

ε - нормированная экспозиция ($\varepsilon = 1$ соответствует значению экспозиции H_R при которой достигается максимальное значение дифракционной эффективности $D\mathcal{E} = D\mathcal{E}_{\text{макс}}$)

Знание величины максимального фазового набега φ_0 для данного регистрирующего материала позволяет рассчитать значение $D\mathcal{E}(\varepsilon)$ с использованием вышеприведенных формул при любом другом соотношении пучков как для зеркальных, так и диффузных объектов.

Геометрическое изменение масштаба восстановленного голограммой изображения $M_{\text{геом}}$ в общем случае можно оценить, используя выражение:

$$M_{\text{геом}} = \left[1 \pm \left(\frac{z_1 \lambda_1}{z_3 \lambda_2} \right) - \left(\frac{z_1}{z_2} \right) \right]^{-1}$$

Знак “+” соответствует мнимому изображению, а знак “-” – действительному изображению.

В выражении для M используются следующие обозначения:

λ_1 – длина волны света при записи голограммы

λ_2 – длина волны света при восстановлении голограммы

z_1 – расстояние от голограммы до объекта при записи голограммы

z_2 – расстояние от голограммы до опорного пучка при записи голограммы

z_3 – расстояние от голограммы до опорного пучка при восстановлении голограммы

Изменение углового размера объекта будет определяться только изменением длины волны при восстановлении голограммы и составит:

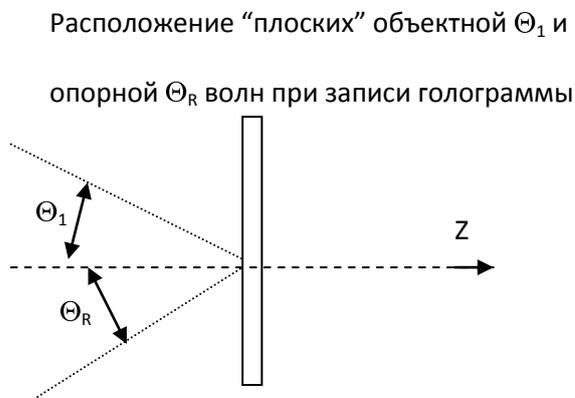
$$M_{\text{угл.}} = \frac{\lambda_2}{\lambda_1}$$

В общем случае, когда плоские опорная и восстанавливающая волны идут под углом к оси (Θ_R и Θ_C соответственно), углы дифракции для мнимого Θ_V и действительного Θ_T изображений можно определить по формулам:

$$\Theta_V = \mu\Theta_1 + \Theta_C - \mu\Theta_R$$

$$\Theta_T = -\mu\Theta_1 + \Theta_C + \mu\Theta_R$$

где $\mu = \lambda_2/\lambda_1$, Θ_1 – положение объектной волны.



Для случая записи голограммы диффузного объекта дополнительно измеряют контраст K восстановленного голограммой изображения. Для этого, как правило, экспериментально измеряют яркость максимально «светлых» ($I_{\text{св}}$) и «тёмных» ($I_{\text{т}}$) участков в фокусе восстановленного голограммой изображения и вычисляют K по формуле:

$$K = (I_{\text{св}} - I_0) / (I_{\text{т}} - I_0),$$

где I_0 – значение яркости в отсутствие лазерного освещения, определяемое величиной «темнового» тока измерительной аппаратуры.

КОНТРОЛЬНЫЕ ВОПРОСЫ

- 1) Принцип действия и основные элементы лазера?
- 2) Основные параметры лазерного излучения?
- 3) Типы голограмм и их отличительные свойства?
- 4) Регистрирующие среды для голографии?

5) Параметры, характеризующие свойства голограмм?

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Р. Кольер, К. Беркхарт, Л. Лин, «Оптическая голография», Мир, 1973.
2. «Оптическая голография», в 2-х томах, под редакцией Г. Колфилда, Мир, 1982.
3. В.Г. Комар, О.Б. Серов, «Изобразительная голография и голографический кинематограф», Искусство, 1987.
4. Ф. Качмарек, «Введение в физику лазеров», Мир, 1981.
5. Н. I. Bjelkhagen “Silver halide recording materials for holography and their processing”, Springer, Berlin, 1993.

В динамической голографии используются среды, которые совсем не требуют какой-либо промежуточной обработки и изменяют свои оптические характеристики непосредственно под действием падающего на них излучения. Строго говоря, любая фоточувствительная среда, галогидосеребряная или несеребряная, в той или иной степени меняет свои параметры под действием записывающих пучков, однако эти изменения обычно так малы, что практически не влияют на амплитуды и фазы волн, распространяющихся в объеме голограммы. Если же это влияние становится заметным, то такую среду называют нелинейной, а голограмму – динамической. При действии световых пучков на нелинейную среду в ее объеме возникает пространственная модуляция коэффициента поглощения или показателя преломления среды, повторяющая картину интерференции пучков – голографическая динамическая решетка, которая изменяет амплитуды и фазы, распространяющихся в голограмме пучков – пучки начинают взаимодействовать друг с другом. Это в свою очередь изменяет картину их интерференции и т. д. В результате стадии формирования и считывания динамических голограмм оказываются неразрывно связанными и происходят одновременно, а сам процесс запись – считывание является одноступенчатым.

Одним из наиболее важных применений динамической голографии является коррекция волновых фронтов лазерного излучения, особенно с использованием явления обращения волнового фронта. Другое направление практических приложений связано с обработкой и усилением оптических изображений. Третье направление связано с изучением сред и быстропротекающих процессов в них. В настоящее время для записи динамических голограмм используются самые разнообразные диэлектрики

и полупроводниковые кристаллы с широкими и узкими запрещенными зонами; неселективно поглощающие жидкости и газы; кристаллы, жидкости и газы, активные к молекулярным, акустическим или ориентационным колебаниям, резонансно поглощающие пары металлов; люминесцирующие растворы красителей; электрооптические, фоторефрактивные и жидкие кристаллы; полимеры и кристаллы с резонансным поглощением, охлаждаемые до низких температур и т. д.

В последнее время был предложен существенно новый метод трансформации пространственной структуры волновых полей, основанный на записи голограммы в нелинейном материале, обладающем электронной нелинейностью второго порядка (квадратичной нелинейностью), которая определяется величиной соответствующей восприимчивости среды $\chi^{(2)}$. Поскольку постоянная времени записи в таком материале предельно мала, требование взаимной когерентности записываемых волн в этом случае становится необязательным. В отличие от обычной голограммы изображение в данном случае формируется излучением второй гармоники, что приводит к изменению масштаба восстановленного изображения по глубине. Однако основным отличительным свойством такой голограммы является ее исключительное быстродействие, обусловленное чисто электронным механизмом взаимодействия света с веществом. В результате интерферирующие волны взаимодействуют в каждый момент времени с созданной ими в этот же момент времени голограммой. В следующий момент времени предыдущая голограмма исчезает, и волны взаимодействуют с новой голограммой, которая соответствует параметрам волн на данный момент.

Естественно возникает мысль, что голограмма, обладающая такими свойствами, может быть записана не только при некотором нарушении взаимной когерентности интерферирующих волн, но также и в том случае, когда длины волн объектной и референтной волн совершенно различны. Безинерционное взаимодействие световых пучков в средах с электронной квадратичной нелинейностью позволило реализовать в известных работах по нелинейной оптике процессы смешения излучения с различными длинами волн, приводящие к генерации излучения с суммарными и разностными частотами, а также использовать эти процессы для визуализации ИК-изображений.

В данной работе будут теоретически и экспериментально изучены процессы формирования и трансформации изображений, восстановленных динамическими голограммами, записанными по внеосевой схеме в квадратичной безинерционной нелинейной среде в условиях, когда частоты объектной и референтной волн различны,

причем разность частот составляет десятые доли от максимального значения частоты. Теория выполнена в скалярном приближении для случая параксиального хода лучей в предположении, что толщина голограммы мала. В дальнейшем для краткости будем называть исследуемые в данной работе безинерционные динамические голограммы с записью в квадратичных нелинейных средах просто динамическими $\chi^{(2)}$ -голограммами.

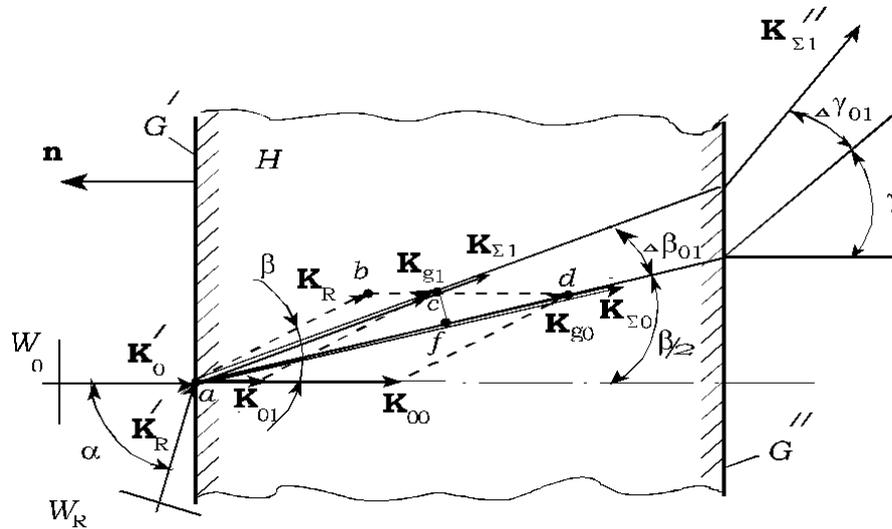


Рис. 1. Схема взаимодействия волн с волновыми векторами \mathbf{k}_i при записи многочастотных динамических голограмм в квадратичных нелинейных средах.

Согласно теории записи динамических голограмм пучками различных частот в квадратичных нелинейных средах, представленной в приложении и на рис.1, при неколлинеарном взаимодействии таких пучков на выходе голограммы помимо проходящих исходных пучков возникают новые пучки, отличающиеся от них как по частоте, так и по направлению, которые распространяются вблизи биссектрисы угла схождения исходных пучков.

Выражение для углов между новыми пучками $\Delta\beta'$, восстановленными такой голограммой, имеет вид:

$$\Delta\beta' = \frac{1}{2} \alpha \frac{n_{2\omega}}{n_{\omega}} \frac{\Delta\nu_{OR}}{2\nu_R - \Delta\nu_{OR}}, \quad (*)$$

где α - угол между пучками, падающими на голограмму, n_{ω} и $n_{2\omega}$ - показатели преломления кристалла на основной и удвоенной частотах записывающего голограмму излучения, ν_R - частота опорного пучка и $\Delta\nu_{OR}$ - разность частот объектного и

опорного пучков. Согласно этой формуле величина $\Delta\beta'$ определяется главным образом углом схождения падающих на голограмму пучков на голограмме α и их частотным сдвигом $\Delta\nu_{OR}$.

Эффективность преобразования исходного излучения в формируемые динамическими голограммами новые пучки существенно зависит от условий эксперимента, прежде всего от нелинейной квадратичной восприимчивости кристалла $\chi^{(2)}$ и плотности мощности падающего излучения, и может достигать десятков процентов. В проводимых экспериментах она составит около одного процента.

Цель части работы состоит в демонстрации возможности одновременного формирования в средах с квадратичной нелинейностью ряда голограмм с различными длинами волн объектного и референтного пучков, а также в измерении углов между направлениями распространения генерируемых голограммой волн с последующим сопоставлением результатов эксперимента и теории. Общая схема эксперимента представлена на рис.2.

В качестве источника излучения используется одномодовый Nd:YAG лазер «Picochrom», генерирующий световые импульсы длительностью 300 пс на длине волны $\lambda = 1064$ нм, с выходной энергией до 15 мДж и частотой следования 1 Гц.

На рис.3 представлена оптическая схема лазера «Picochrom» с временной компрессией импульса на основе явления вынужденного рассеяния Мандельштама-Бриллюэна (ВРМБ), генерирующего импульсы излучения с длиной волны 528 нм и длительностью 300 пс. В лазерном модуле формируется излучение на длине волны 1.064 мкм с длительностью импульса ~ 5 нс.

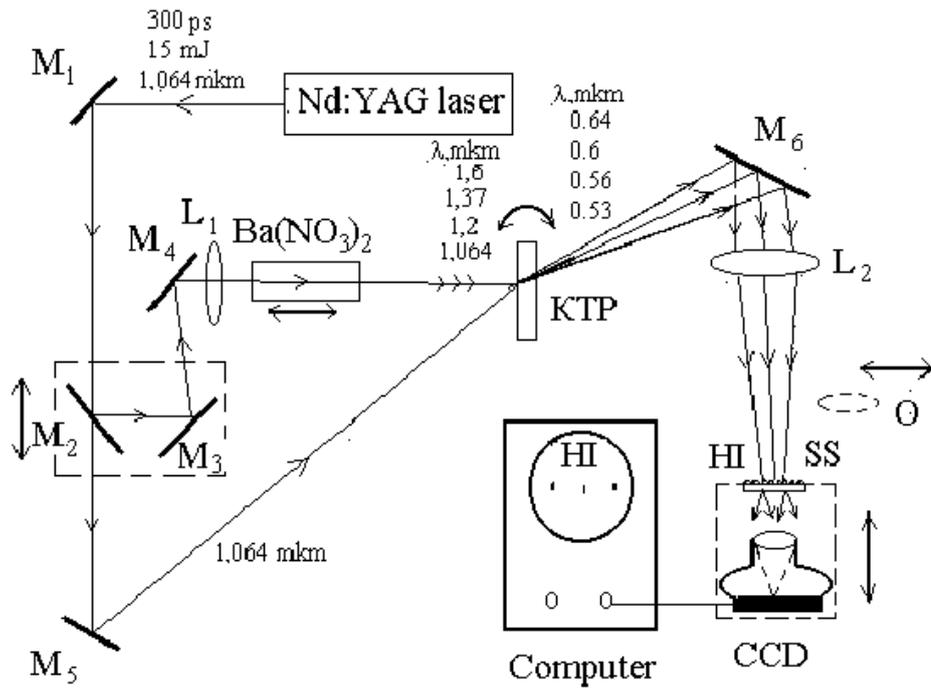


Рис. 2. Оптическая схема установки для записи многочастотных голограмм в квадратичных нелинейных средах. $M_1 - M_6$ – зеркала; L_1, L_2 – линзы; $Ba(NO_3)_2$ – комбинационно активный кристалл нитрата бария; КТР – кристалл, обладающий значительной нелинейностью второго порядка, который использовался в качестве светочувствительной среды для записи динамических голограмм; SS – диффузно рассеивающий экран.

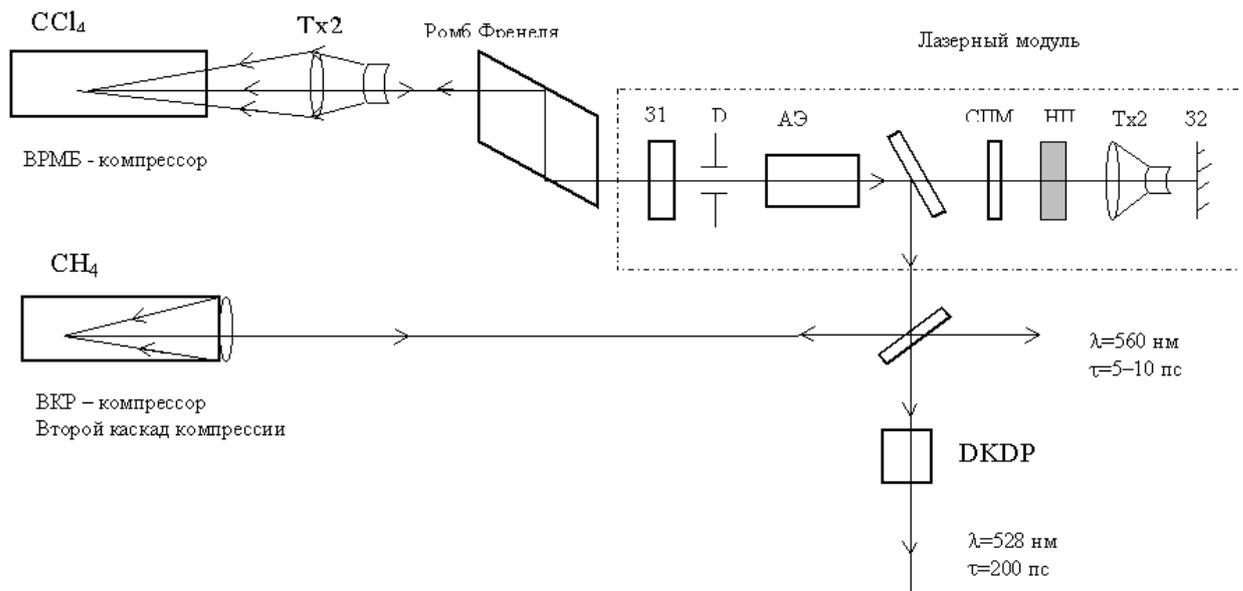


Рис. 3. Оптическая схема лазер «Picochrom» с ВРМБ и ВКР компрессией.

Зеркала 31 и 32 образуют резонатор лазера длиной 45 см, в центре резонатора расположен активный элемент (АЭ) - кристалл иттрий алюминиевого граната, активированный Nd^{3+} . Между активным элементом и задним глухим зеркалом расположены следующие элементы: СПМ – селектор поперечных мод, формирующий одночастотное излучение, НП - насыщающийся поглотитель, служащий для получения одиночных импульсов, $T \times 2$ - двукратный телескоп, позволяющий увеличить эффективную длину резонатора и совместно с диафрагмой Д формирующий одномодовое по поперечным индексам излучение. На выходе из лазерного модуля излучение, пройдя ромб Френеля, преобразует свою поляризацию из линейной в круговую и попадает на двукратный телескоп, расширяющий пучок излучения, для получения более узкой области фокусировки излучения в ВРМБ кювете, заполненной CCl_4 . От ВРМБ компрессора отражаются импульсы излучения длительностью 300 пс. Пройдя ромб Френеля, они приобретают ортогональную с первоначальной поляризацию, дополнительно усиливаются в АЭ и направляются интерференционным зеркалом и последующими элементами на дополнительный временной компрессор, на основе вынужденного комбинационного рассеяния (ВКР) состоящий из кюветы, наполненной CH_4 под давлением 15...25 атм. От ВКР компрессора отражаются импульсы излучения с длительностью 30 пс на длине волны 560 нм. Излучение также может попадать на кристалл DKDP, удваивающий частоту излучения 300 пс импульсов ($\lambda = 528$ нм) и использоваться для записи голограмм.

Другим существенным элементом экспериментальной установки являлся кристалл КТР толщиной 2 мм, обладающий значительной нелинейностью второго порядка, который использовался в качестве светочувствительной среды для записи динамических голограмм. Благодаря близкому к 90° -синхронизму при возбуждении генерации второй гармоники на длине волны выбранного источника излучения он позволяет достигнуть достаточной для уверенной регистрации изображений эффективности преобразования в широкой области углов падения лучей на кристалл, что необходимо при использовании неколлинеарных схем преобразования.

Третьим ключевым элементом экспериментальной установки является комбинационно активный кристалл нитрата бария $\text{Ba}(\text{NO}_3)_2$ с высоким (11 см/ГВт) инкрементом стационарного ВКР-усиления и длиной, равной 80 мм, который дискретно смещал частоту падающего на него излучения с длиной волны 1,064 мкм на интервалы, кратные его стоксову сдвигу $\Delta \nu_S = 1047 \text{ см}^{-1}$. Эксперимент показал, что

данный кристалл был способен осуществить тройной стоксов сдвиг излучения, в результате чего на выходе из кристалла можно было наблюдать одновременно 4 монохроматических волны: волну с основной частотой, соответствующей $\lambda = 1,064$ мкм, а также 3 волны с $\lambda = 1,2; 1,37$ и $1,6$ мкм, претерпевших стоксовы сдвиги на $\Delta\nu_1 = 1047 \text{ см}^{-1}$; $\Delta\nu_2 = 2 \times 1047 \text{ см}^{-1}$; $\Delta\nu_3 = 3 \times 1047 \text{ см}^{-1}$ соответственно.

Система зеркал М1-М5 формирует из излучения основной частоты два равных по интенсивности пучка с вертикальной линейной поляризацией, углом схождения на голограмме в горизонтальной плоскости, равным 14.5° и разностью хода не более 5 мм, что обеспечивало практически полное временное перекрытие этих пучков на голограмме. Тщательное совмещение на частоте основного излучения этих пучков в плоскости голограммы позволяет достичь и пространственного их перекрытия. Линза L_1 с фокусным расстоянием $f = 1$ м повышала плотность излучения на кристалле нитрата бария (до 10 раз), что необходимо для увеличения в объектном пучке доли излучения со смещенными частотами.

Перемещение кристалла нитрата бария вдоль оси излучения позволяет регулировать эффективность преобразования основного излучения в стоксовы компоненты, а также число этих компонентов, вплоть до трех при максимальной плотности мощности падающего излучения. Картина углового распределения формируемых голограммой пучков проецируется линзой L_2 ($f = 250$ мм) с однократным увеличением на диффузно рассеивающий экран SS (scattering screen), установленный в фокальной плоскости линзы на общем подвижном основании с цветной цифровой фотографической CCD камерой Fujifilm FinePix 4900 Zoom с числом пикселей 1200×2000 . Угловое расстояние между пучками излучения (в радианах) связано с измеренным линейным расстоянием между ними Δx в фокальной плоскости линзы следующим соотношением:

$$\Delta\beta' = (\Delta x)/f, \quad (**)$$

где f – фокусное расстояние линзы, выраженное как и Δx , в мм.

На практике свойство динамической $\chi^{(2)}$ - голограммы преобразовывать в реальном времени длину волны падающего излучения и изменять его направление, а также плоскость фокусировки может быть использовано в стекловолоконных линиях

связи, когда в процессе прохождения сигнала через систему необходимо изменить его длину волны и (или) направление, чтобы направить его по другому пути (маршруту, адресу). Это необходимо для сверхбыстрой коммутации потоков информации в оптических вычислителях и линиях связи.

КОНТРОЛЬНЫЕ ВОПРОСЫ:

1. Что такое динамическая голография?
2. Что является динамической голограммой в проводимой работе?
3. Для чего необходим кристалл нитрата бария?
4. Основные свойства кристалла КТР.

Линза как квадратичный фазовый фильтр

Преобразование пространственно-когерентных оптических сигналов. Описание дифракционных явлений интегралом суперпозиции указывает на возможность применения элементов теории линейных систем в оптике. Это следует из общего вида волнового уравнения, описывающего распространение света в среде. Использование интеграла суперпозиции выявляет важнейшее свойство линейной системы, состоящее в том, что она полностью определяется импульсным откликом $h(x_1, y_1; x_0, y_0)$:

$$s(x_1, y_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) h(x_1, y_1; x_0, y_0) dx_0 dy_0. \quad (2.43)$$

Импульсный отклик описывает выходной сигнал в точке x_1, y_1 , когда входным сигналом является δ -функция в точке x_0, y_0 . В оптике импульсный отклик часто называют функцией рассеяния. В решениях дифференциальных уравнений импульсному отклику соответствует функция Грина.

Из линейных оптических систем особо выделяются пространственно-инвариантные (изопланарные) системы, для которых импульсный отклик зависит только от разности координат:

$$h(x_1, y_1; x_0, y_0) = h(x_1 - x_0, y_1 - y_0).$$

Пространственно-инвариантные системы преобразуют входной сигнал вида $s(x_0 - \xi, y_0 - \eta)$ в выходной $s(x_1 - \xi, y_1 - \eta)$ для всех $s(x_0, y_0)$. Например, изображение источника, создаваемого изопланарной системой, при

Линза как квадратичный фазовый фильтр

Преобразование пространственно-когерентных оптических сигналов. Описание дифракционных явлений интегралом суперпозиции указывает на возможность применения элементов теории линейных систем в оптике. Это следует из общего вида волнового уравнения, описывающего распространение света в среде. Использование интеграла суперпозиции выявляет важнейшее свойство линейной системы, состоящее в том, что она полностью определяется импульсным откликом $h(x_1, y_1; x_0, y_0)$:

$$s(x_1, y_1) = \int \int_{-\infty}^{\infty} s(x_0, y_0) h(x_1, y_1; x_0, y_0) dx_0 dy_0. \quad (2.43)$$

Импульсный отклик описывает выходной сигнал в точке x_1, y_1 , когда входным сигналом является δ -функция в точке x_0, y_0 . В оптике импульсный отклик часто называют функцией рассеяния. В решениях дифференциальных уравнений импульсному отклику соответствует функция Грина.

Из линейных оптических систем особо выделяются пространственно-инвариантные (изопланарные) системы, для которых импульсный отклик зависит только от разности координат:

$$h(x_1, y_1; x_0, y_0) = h(x_1 - x_0, y_1 - y_0).$$

Пространственно-инвариантные системы преобразуют входной сигнал вида $s(x_0 - \xi, y_0 - \eta)$ в выходной $s(x_1 - \xi, y_1 - \eta)$ для всех $s(x_0, y_0)$. Например, изображение источника, создаваемого изопланарной системой, при неизменной форме меняет только свое положение, если источник в предметной плоскости перемещается. Условие пространственной инвариантности редко выполняется для реальных оптических систем полностью по входной и выходной плоскостям. Однако пространственная инвариантность имеет место в пределах малых областей, называемых изопланарными участками. Для полного описания системы достаточно разбить предметную плоскость на изопланарные участки, определив для каждого из них свой импульсный отклик. Интеграл суперпозиции (2.43) в случае пространственно-инвариантных систем принимает форму свертки входного сигнала и импульсного отклика:

$$s(x_1, y_1) = \int \int_{-\infty}^{\infty} s(x_0, y_0) h(x_1 - x_0, y_1 - y_0) dx_0 dy_0. \quad (2.44)$$

Анализ пространственно-инвариантных систем значительно упрощается с введением понятия когерентной передаточной функции. Основным свойством таких систем является то, что свертке сигнала с импульсным откликом в координатном пространстве соответствует произведение фурье-образа входного сигнала и передаточной функции в пространственно-частотной области. При этом предполагается, что масштабы пространственных частот входного и выходного сигналов согласованы:

$$F(K_x, K_y) = F_0(K_x, K_y)H(K_x, K_y). \quad (2.45)$$

Здесь $F(K_x, K_y)$ – фурье-образ выходной функции $s(x_1, y_1)$; $F_0(K_x, K_y)$ – фурье-образ функции $s(x_0, y_0)$; $H(K_x, K_y)$ – фурье-образ импульсного отклика; K_x, K_y – пространственная частота. Функция $H(K_x, K_y)$ называется когерентной передаточной функцией системы. По аналогии с (2.45) легко показать, что фурье-преобразование

свертки в координатном пространстве соответствует произведению фурье-образа одной функции на комплексно-сопряженный фурье-образ другой:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp -j K_x \xi + K_y \eta d \xi d \eta \times \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1(x_0, y_0) s_2^*(\xi - x_0, \eta - y_0) dx_0 dy_0 = s_1(K_x, K_y) s_2^*(K_x, K_y) .$$

(2.46)

В соответствии с (2.46) для автосвертки сигнала можно записать

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp -j K_x \xi + K_y \eta d \xi d \eta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) s^*(x_0 - x_0, \eta - y_0) dx_0 dy_0 = \\ = |s(K_x, K_y)|^2 .$$

(2.47)

Преобразование пространственно-некогерентных оптических сигналов. До сих пор мы говорили об оптических системах, линейно отображающих комплексную амплитуду входного сигнала в комплексную амплитуду выходного. Это относится к случаю, когда оптические сигналы пространственно-когерентны, поэтому и система называется когерентной. Если же входной сигнал не является пространственно-когерентным, импульсные отклики в выходной плоскости для пространственно-разнесенных точечных источников статистически независимы. В этом случае оптическая система, линейно преобразующая интенсивность входного сигнала, называется некогерентной. Импульсный отклик h_{NK} некогерентной системы пропорционален квадрату модуля когерентного импульсного отклика:

$$h_{NK} \sim |h(x_1 - x_0, y_1 - y_0)|^2 .$$

(2.48)

Соответственно интенсивность выходного сигнала определяется сверткой интенсивности входного сигнала с некогерентным импульсным откликом:

$$I(x_1, y_1) \sim \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(x_1 - x_0, y_1 - y_0)|^2 I(x_0, y_0) dx_0 dy_0 .$$

(2.49)

Для пространственно-инвариантной системы можно определить передаточную функцию

$$W(K_x, K_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{NK} \exp -j [K_x (x_1 - x_0) + K_y (y_1 - y_0)] \times \\ \times d(x_1 - x_0) d(y_1 - y_0) ,$$

(2.50)

которая называется оптической передаточной функцией (ОПФ). Она является фурье-образом импульсного отклика некогерентной изопланарной системы. По аналогии с (2.45) можно получить, что фурье-образ интенсивности выходного сигнала $G(K_x, K_y)$ равен произведению фурье-образа интенсивности входного сигнала $G_0(K_x, K_y)$

на ОПФ:

$$G(K_x, K_y) = G_0(K_x, K_y) W(K_x, K_y) .$$

(2.51)

Связь между ОПФ и КПФ находится из выражения, определяющего фурье-преобразование автосвертки:

$$W(K_x, K_y) = \int \int_{-\infty}^{\infty} H(\xi, \eta) H^*(K_x - \xi, K_y - \eta) d\xi d\eta. \quad (2.52)$$

Найдем импульсные отклики некоторых оптических элементов.

На рис. 2.8 изображена оптическая система, линейно преобразующая сигнал, заданный в плоскости $x_0 y_0$, в выходной сигнал, формируемый в плоскости $x_1 y_1$. В этой системе сигнал проходит через оптический элемент, расположенный в плоскости $x y$. В дальнейшем будем широко пользоваться функциями вида

$$\psi(x, y; p) = \exp\left\{jk/2 p (x^2 + y^2)\right\}, \quad (2.53)$$

где p – действительная или комплексная величина. Очевидно, что $\psi^*(x, y; p) = \psi(x, y; -p)$. Поместим во входную плоскость точечный источник, координаты которого x_0, y_0 . Поле точечного источника в плоскости, локализованной непосредственно перед оптическим элементом, в приближении Френеля описывается импульсным откликом свободного пространства (2.34):

$$h(x, y; x_0, y_0) = \mathcal{O}_0 / j\lambda \psi(x - x_0; y - y_0; D_0). \quad (2.54)$$

Здесь и далее используется обозначение функции (2.53), где $D_0 = 1/d_0$. Пусть функция $\varphi(x, y)$ определяет амплитудное и фазовое пропускание исследуемого оптического элемента. Поле непосредственно за элементом можно описать выражением

$$E(x, y) = \mathcal{O}_0 / j\lambda \psi(x - x_0, y - y_0; D_0) \varphi(x, y).$$

Импульсный отклик пространства между оптическим элементом и плоскостью анализа $x_1 y_1$ представим с учетом (2.53) и (2.54) в виде

$$h(x_1, y_1; x, y) = \mathcal{O}_1 / j\lambda \psi(x_1 - x, y_1 - y; D_1),$$

где $D_1 = 1/d_1$. Поле в выходной плоскости определяется интегралом суперпозиции

$$h(x_1, y_1; x_0, y_0) = \iint_{-\infty}^{\infty} E(x, y) h(x_1, y_1; x, y) dx dy.$$

После преобразований с учетом (2.53) получаем

$$h(x_1, y_1; x_0, y_0) = -\left(\frac{D_0 D_1}{\lambda}\right) \psi(x_0, y_0; D_0) \psi(x_1, y_1; D_1) \iint_{-\infty}^{\infty} \varphi(x, y) \times \\ \times \psi(x, y; D_0 + D_1) \exp\left\{-jkD_0 \left[\left(x_0 + \frac{x_1}{\Gamma}\right)x + \left(y_0 + \frac{y_1}{\Gamma}\right)y\right]\right\} dx dy, \quad (2.55)$$

где $\Gamma = D_0/D_1$. Для установления явного вида функции $\varphi(x, y)$ необходима дополнительная информация о свойствах исследуемого оптического элемента. Найдем условия, при которых рассматриваемая система является пространственно-инвариантной. Например, выберем в качестве выходной плоскости системы плоскость изображения входного сигнала. Пусть оптический элемент с неограниченной апертурой пространственно-инвариантно отображает с точностью до фазовых множителей $\psi(x_0, y_0; D_0)$ и $\psi(x_1, y_1; D_1)$ точку, заданную в предметной плоскости, в инвертированную

относительно оптической оси точку на плоскости изображения. В этом случае можно записать

$$\psi^* \left(x_0, y_0; D_0 \right) \xrightarrow{D_0} \psi^* \left(x_1, y_1; D_1 \right) \xrightarrow{D_1} h \left(x_1, y_1; x_0, y_0 \right) \xrightarrow{D_1} -\Gamma^{-1} \delta \left(x_0 + \frac{x_1}{\Gamma}, y_0 + \frac{y_1}{\Gamma} \right). \quad (2.56)$$

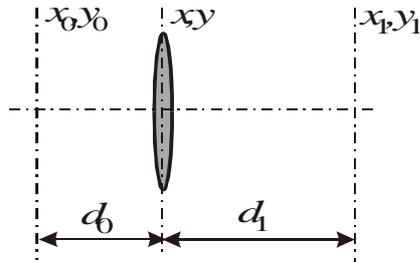


Рис. 2.8

Уравнение (2.56) удовлетворяется при $\varphi \left(x, y \right) \xrightarrow{D_0} \psi \left(x, y; D_0 + D_1 \right) \xrightarrow{D_1} 1$. Отсюда находим

$$\varphi \left(x, y \right) \xrightarrow{D_0} \psi \left(x, y; -D_0 + D_1 \right).$$

Следовательно, оптический элемент в рассматриваемой системе выполняет фазовое преобразование

$$\varphi \left(x, y \right) = \exp \left\{ -\frac{jk}{2} \Phi \left(x^2 + y^2 \right) \right\}, \quad (2.57)$$

где Φ удовлетворяет условию

$$D_0 + D_1 - \Phi = 0. \quad (2.58)$$

Видно, что (2.58) является известным условием формирования изображения тонкой линзой, оптическая сила которой $\Phi = 1/f$, где f – фокусное расстояние. Тогда Γ принимает смысл линейного коэффициента передачи (увеличения) оптической системы. Следовательно, тонкая линза осуществляет квадратичное фазовое преобразование сигнала. Исследуем свойства линзы как оптического элемента, линейно преобразующего комплексную амплитуду сигнала. Подставляя (2.57) в (2.55), получаем выражение для импульсного отклика тонкой линзы с неограниченной апертурой

$$h \left(x_1, y_1; x_0, y_0 \right) \xrightarrow{D_1} - \left(\frac{D_0 D_1}{\lambda^2} \right) \psi \left(x_0, y_0; D_0 \right) \xrightarrow{D_0} \psi \left(x_1, y_1; D_1 \right) \xrightarrow{D_1} \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi \left(x, y; Q \right) \exp \left[-jk \left[D_0 x_0 + D_1 x_1 \left(x + D_0 y_0 + D_1 y_1 \right) \right] \right] dx dy, \quad (2.59)$$

где $Q = D_0 + D_1 - \Phi$. При $Q = 0$ входная и выходная плоскости являются соответственно предметной и плоскостью изображения. Импульсный отклик в этом случае запишем так:

$$h(x_1, y_1; x_0, y_0) = -\Gamma^{-1} \psi(x_0, y_0; D_0) \psi(x_1, y_1; D_1) \delta\left(x_0 + \frac{x_1}{\Gamma}; y_0 + \frac{y_1}{\Gamma}\right). \quad (2.60)$$

Пространственная инвариантность функции рассеяния (2.60) нарушается присутствием квадратичных фазовых множителей $\psi(x_0, y_0; D_0)$ и $\psi(x_1, y_1; D_1)$.

Как видно из (2.59), импульсный отклик тонкой линзы с неограниченной апертурой определяется через фурье-преобразование гауссовой функции $\psi(x, y; Q) = \exp\{-jk/2 Q(x^2 + y^2)\}$. Известно, что фурье-образ гауссовой функции имеет также гауссову форму:

$$\int_{-\infty}^{\infty} \exp\{-\beta\xi^2\} \exp\{-jk\xi\} d\xi = (\pi/\beta)^{1/2} \exp\{-K^2/4\beta\}. \quad (2.61)$$

Учитывая это, для импульсного отклика (2.59) получаем

$$h(x_1, y_1; x_0, y_0) = \frac{kD_0D_1}{2\pi jQ} \psi\left(x_0, y_0; D_0 - \frac{D_0^2}{Q}\right) \psi\left(x_1, y_1; D_1 - \frac{D_1^2}{Q}\right) \times \exp\left\{-jk \frac{D_0D_1}{Q} (x_0x_1 + y_0y_1)\right\}. \quad (2.62)$$

Когерентный импульсный отклик линзы в общем случае не может быть представлен в пространственно-инвариантной форме из-за присутствия квадратичных фазовых множителей. Если на входе системы сформирован сигнал $s(x_0, y_0)$, выходной сигнал, согласно (2.44), равен

$$s(x_1, y_1) = \int_{-\infty}^{\infty} s(x_0, y_0) h(x_1, y_1; x_0, y_0) dx_0 dy_0. \quad (2.63)$$

Обратимся снова к выражению (2.59), описывающему импульсный отклик, и исследуем, при каких условиях тонкая линза является пространственно-инвариантной системой. При $Q = 0$, когда выходная плоскость совпадает с плоскостью изображения входного сигнала, пространственная инвариантность когерентной системы нарушается квадратичными фазовыми множителями $\psi(x_0, y_0; D_0)$ и $\psi(x_1, y_1; D_1)$. Компенсация этих фазовых множителей может быть осуществлена в оптической схеме (рис. 2.9), где пространственная инвариантность достигается за счет помещения в плоскостях x_0, y_0 и x_1, y_1 компенсационных линз (1 и 3), фокусные расстояния которых выбираются соответственно равными $f_1 = 1/D_0$ и $f_3 = 1/D_1$. При этом импульсный отклик системы приводится к пространственно-инвариантному виду

$$h(x_1, y_1) = -\Gamma^{-1} \delta\left(x_0 + \frac{x_1}{\Gamma}, y_0 + \frac{y_1}{\Gamma}\right). \quad (2.64)$$

Рассмотрим влияние конечных размеров апертуры линзы на ее свойства. Пусть функция $P(x, y)$ описывает апертуру линзы. Найдем импульсный отклик для системы, изображенной на рис. 2.9:

$$h(x_1, y_1; x_0, y_0) = -\frac{D_0 D_1}{\lambda^2} \times \\ \times \int_{-\infty}^{\infty} \int P(x, y) \exp \left\{ -jkD_0 \left[\left(x_0 + \frac{x_1}{\Gamma} \right) x + \left(y_0 + \frac{y_1}{\Gamma} \right) y \right] \right\} dx dy.$$

Для цилиндрически симметричной системы $P(x, y) = \text{circ}(2r/d)$, где d – диаметр линзы.

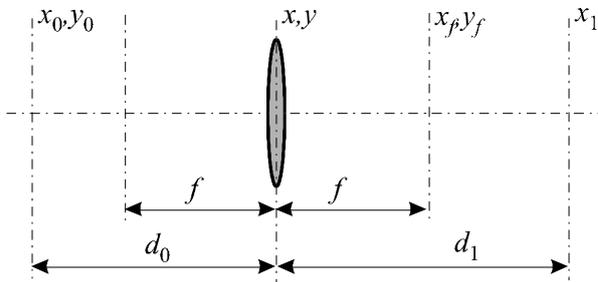


Рис. 2.10

Преобразование Фурье такой апертурной функции имеет также круговую симметрию:

$$h(\rho) = \frac{dD_1}{2\lambda} \frac{J_1(\rho/2)}{\rho}.$$

Здесь $\rho = kD_0 \left[\left(x_0 + \frac{x_1}{\Gamma} \right)^2 + \left(y_0 + \frac{y_1}{\Gamma} \right)^2 \right]^{1/2}$;

$J_1\left(\frac{d\rho}{2}\right)$ – функция Бесселя. Импульсному отклику (2.65) соответствует КПФ

$$H(\rho) = \text{circ}(\rho/d). \quad (2.66)$$

Из выражения (2.65) для импульсного отклика следует, что когерентная оптическая система, показанная на рис. 2.8, сохраняет пространственную инвариантность и при конечной апертуре линзы. Разрешение этой оптической системы, согласно (2.65), определяется с точностью до постоянного множителя функцией

$$h(r_1) = \frac{1}{r_1} J_1\left(\frac{K_m r_1}{\Gamma}\right), \quad (2.67)$$

где $r_1 = (x_1^2 + y_1^2)^{1/2}$; $K_m = kD_0 d/2$ – максимальная пространственная частота, пропускаемая системой. Поскольку выходной сигнал пространственно-инвариантной системы описывается сверткой входного сигнала с импульсным откликом, каждый точечный источник во входной плоскости трансформируется на выходе системы в пятно, размер которого принимается равным расстоянию между нулями функции (2.67) ($\Delta = 1,22\lambda/D_1 d$).

Для рассмотрения преобразования интенсивности светового сигнала системой, показанной на рис. 2.8, следует найти ее некогерентный импульсный отклик и ОПФ по формулам (2.48) и (2.50), учитывая (2.63) и (2.64). Эта оптическая система сохраняет пространственную инвариантность при отображении интенсивности некогерентных сигналов.

Таким образом, линза в общем случае преобразует оптический сигнал как квадратичный фазовый фильтр.

Оптическое преобразование Фурье

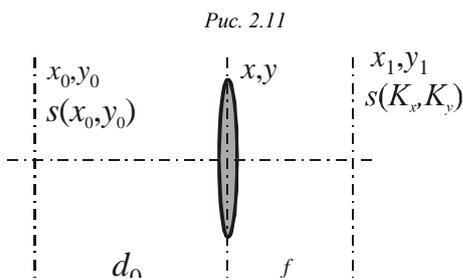
В зависимости от расположения плоскости, в которой локализован входной сигнал, и плоскости его регистрации линза может выполнять различные функции (рис. 2.10). При условии $D_1 = \Phi$ импульсный отклик (2.62) принимает вид

$$h(x_f, y_f; x_0, y_0) = \frac{k\Phi}{2\pi j} \psi\left(x_f, y_f; \Phi - \frac{\Phi^2}{D_0}\right) \exp[-j(K_x x_0 + K_y y_0)], \quad (2.68)$$

где $K_x = (k/f)x_f$, $K_y = (k/f)y_f$ – круговые пространственные частоты.

Таким образом, если выходной сигнал регистрируется в задней фокальной плоскости (рис. 2.11), то он с точностью до квадратичного фазового множителя представляет собой фурье-образ входного сигнала. Из (2.63) с учетом (2.62) и (2.68) имеем

$$\begin{aligned} s(x_f, y_f) &= \frac{k\Phi}{2\pi j} \psi\left(x_f, y_f; \Phi - \frac{\Phi^2}{D_0}\right) \times \\ &\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) \exp[-j(K_x x_0 + K_y y_0)] dx_0 dy_0 = \\ &= \frac{k\Phi}{2\pi j} \psi\left(x_f, y_f; \Phi - \frac{\Phi^2}{D_0}\right) s(K_x, K_y). \end{aligned} \quad (2.69)$$



Следовательно, в задней фокальной плоскости линзы всегда формируется распределение комплексных амплитуд выходного сигнала, которое имеет вид фурье-образа входного сигнала. С точностью до квадратичного фазового множителя это распределение не зависит от положения плоскости, в которой задается входной сигнал. Квадратичный фазовый множитель $\psi(x_f, y_f; \Phi - \Phi^2/D_0)$ обращается в

единицу, если входной сигнал задан в передней фокальной плоскости линзы ($\Phi_0 = \Phi$). В этом случае выходной сигнал является фурье-образом входного сигнала с точностью до постоянного множителя $k\Phi/2\pi j$:

$$\begin{aligned} s(x_f, y_f) &\approx \frac{k\Phi}{2\pi j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) \exp[-j(K_x x_0 + K_y y_0)] dx_0 dy_0 = \\ &= \frac{k\Phi}{2\pi j} s(K_x, K_y). \end{aligned} \quad (2.70)$$

При этом импульсный отклик имеет вид

$$h(x_f, y_f; x_0, y_0) \approx \frac{k\Phi}{2\pi j} \exp[-j(K_x x_0 + K_y y_0)]. \quad (2.71)$$

Так как сигналы в передней и задней фокальных плоскостях линзы оказываются фурье-сопряженными, эти плоскости называются фурье-плоскостями. Если входной сигнал сформирован в произвольной плоскости, для получения точного фурье-преобразования в задней фокальной плоскости линзы следует поместить другую линзу, компенсирующую в (2.69) квадратичный фазовый множитель $\psi(x_f, y_f; \Phi - \Phi^2/D_0)$. Фокусное расстояние компенсирующей линзы выбирается из условия

$\psi(x_f, y_f; \Phi - \Phi^2/D_0) \psi^*(x_f, y_f; 1/f_K) = 1$, откуда $1/f_K = \Phi(-\Phi/D_0)$. В частности, если входной сигнал задан в плоскости, к основной линзе $D_0 \rightarrow \infty$, линза, выполняющая фурье-идентичны (рис. 2.12).

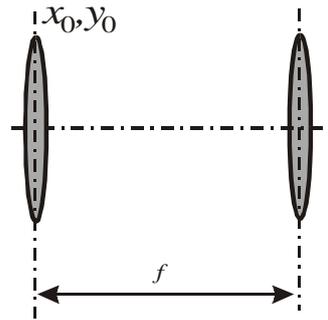


Рис. 2.12

Преобразование Фурье, системой, связывает частотную $K_x K_y$ плоскости.

разложение

структуры сложного

множество плоских волн, направляющие косинусы которых соответствуют пространственным частотам. Координатная и частотная плоскости в рассматриваемой оптической системе являются фурье-сопряженными, а преобразование Фурье – двумерным. Рассмотрим ситуацию, когда условие $D_1 = \Phi$ не выполняется. В этом случае импульсный отклик (2.62) описывает ядро интегрального преобразования Френеля. Выходной сигнал (2.63) с точностью до квадратичного множителя и постоянного комплексного коэффициента представляет собой френелевский образ входного сигнала, а импульсный отклик (2.62) с учетом (2.53) принимает вид

реализуемое оптической координатную x_0, y_0 и

Оно выполняется как

пространственной

светового сигнала на

$$h(x_1, y_1; x_0, y_0) = \frac{D_0 D_1}{j\lambda Q} \exp \left\{ -j \frac{k D_1 \Phi}{2(D_1 - \Phi)} (x_1^2 + y_1^2) \right\} \times \\ \times \exp \left[j\gamma \left(x_0 - \xi \right)^2 + \left(y_0 - \eta \right)^2 \right],$$

$$\text{где } \gamma = \frac{k}{2} \left(D_0 - \frac{D_0^2}{Q} \right); \quad \xi = \frac{D_1 x_1}{D_1 - \Phi}; \quad \eta = \frac{D_1 y_1}{D_1 - \Phi}.$$

Квадратичный фазовый множитель можно скомпенсировать, поместив в плоскость x_1, y_1 тонкую линзу, выполняющую фазовое преобразование $\psi(x_1, y_1; D_1 \Phi / (D_1 - \Phi))$. Выходной сигнал в такой системе является френелевским образом входного сигнала с точностью до постоянного комплексного множителя,

$$s(\xi, \eta) = \frac{D_0 D_1}{j\lambda Q} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) \exp \left[j\gamma \left(x_0 - \xi \right)^2 + \left(y_0 - \eta \right)^2 \right] dx_0 dy_0.$$

Отсюда видно, что система с компенсирующей линзой пространственно-инвариантна.

Следовательно, фурье-образ выходного сигнала определяется произведением фурье-образа входного сигнала на когерентную передаточную функцию, являющуюся фурье-образом функции Френеля. Согласно (2.61),

$$H(K_x, K_y) = \frac{D_0 D_1}{j\lambda Q} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left[j\gamma \left(x_0 - \xi \right)^2 + \left(y_0 - \eta \right)^2 \right] \times \\ \times \exp \left[-j \left[K_x \left(x_0 - \xi \right) + K_y \left(y_0 - \eta \right) \right] \right] dx_0 - \xi dy_0 - \eta = \\ = \frac{\pi D_0 D_1}{\gamma \lambda Q} \exp \left\{ -j \frac{1}{4\gamma} \left(K_x^2 + K_y^2 \right) \right\}.$$

Преобразование Френеля обратимо, а его повторное применение с комплексно-сопряженным ядром восстанавливает исходную функцию. Преобразование Френеля основано на явлении дифракции Френеля и имеет в оптике такое же фундаментальное значение, как и преобразование Фурье.

Рассмотрим особенности *обратного преобразования Фурье в оптике*. На рис. 2.13 представлены последовательно расположенные вдоль оптической оси две идентичные линзы с оптически сопряженными фурье-плоскостями. Входной сигнал $s(x_0, y_0)$ задан в передней фокальной плоскости первой линзы. Согласно (2.70), в задней фокальной плоскости формируется фурье-спектр сигнала

$$s(x_f, y_f) = \frac{k\Phi}{2\pi j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_0, y_0) \exp -j [K_x x_0 + K_y y_0] dx_0 dy_0 =$$

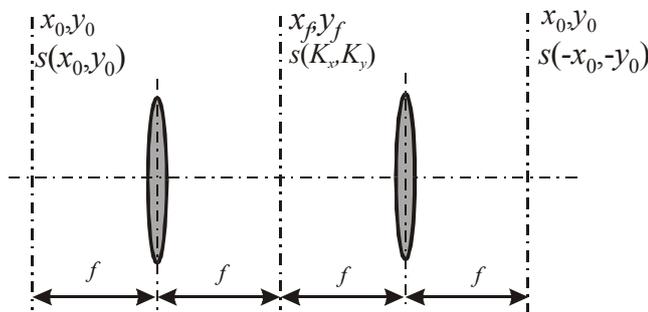


Рис. 2.13

$$= \frac{k\Phi}{2\pi j} s(K_x, K_y).$$

Вторая линза выполняет обратное фурье-преобразование этого сигнала, сформированного в ее передней фурье-плоскости:

$$\begin{aligned} & \frac{k\Phi}{2\pi j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x_f, y_f) \exp -j [K_x x_0 + K_y y_0] dx_f dy_f = \\ & = \frac{k\Phi}{2\pi j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(K_x, K_y) \exp -j [K_x x_0 + K_y y_0] dx_f dy_f = \\ & = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(K_x, K_y) \exp -j [K_x x_0 + K_y y_0] dK_x dK_y = \\ & = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(K_x, K_y) \exp j [K_x (-x_0) + K_y (-y_0)] dK_x dK_y = \\ & = -s(-x_0, -y_0). \end{aligned}$$

Отсюда следует, что обратное преобразование Фурье в оптике выполняется с инверсией координатных осей. В случае, когда фокусные расстояния линз, выполняющих последовательно прямое и обратное преобразования Фурье, не одинаковы, исходный сигнал восстанавливается в инвертированных осях с линейным коэффициентом увеличения Γ , равным отношению фокусных расстояний:

$$\frac{1}{2\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(K_x, K_y) \exp\left\{j\left[K_x\left(-\frac{x_0}{\Gamma}\right) + K_y\left(-\frac{y_0}{\Gamma}\right)\right]\right\} dK_x dK_y =$$

$$= s\left(-\frac{x_0}{\Gamma}, -\frac{y_0}{\Gamma}\right), \quad (2.72)$$

где $\Gamma = f_1/f_2$.

В разделе 1.4 обсуждались свойства одномерного преобразования Фурье, поэтому, не повторяясь в доказательствах, приведем основные свойства оптического преобразования Фурье.

Преобразование масштабов. Сжатие сигнала ведет к расширению его спектральной плотности в частотном пространстве. При этом амплитуда спектральной плотности уменьшается обратно пропорционально коэффициенту сжатия. Растяжение сигнала вызывает сжатие спектральной плотности в частотном пространстве и соответствующее увеличение ее амплитуды.

$$s(x_0, y_0) \xrightarrow{\text{сжатие}} |ab|^{-1} s\left(\frac{K_x}{a}, \frac{K_y}{b}\right). \quad (2.73)$$

Сжатие (растяжение) масштаба в координатной плоскости вызывает соответственно растяжение (сжатие) фурье-сопряженной (частотной) плоскости.

Теорема смещения. Смещение сигнала в координатной плоскости соответствует умножению его фурье-образа на фазовый множитель, являющийся линейной функцией частоты:

$$s(x_0 - a, y_0 - b) \leftrightarrow s(K_x, K_y) \exp -j aK_x + bK_y, \quad (2.74)$$

а смещение фурье-спектра – умножению сигнала в координатной плоскости на фазовый множитель

$$s(K_x - a, K_y - b) \leftrightarrow s(x_0, y_0) \exp -j ax_0 + by_0. \quad (2.75)$$

Изменение знака в фазовом множителе связано с тем, что при обратном фурье-преобразовании в оптической системе происходит инверсия координатных осей. Инверсия координат в плоскости, где задан сигнал, соответствует смене знака пространственной частоты для фурье-образа сигнала

$$s(x_0, -y_0) \leftrightarrow s(K_x, -K_y). \quad (2.76)$$

Операция комплексного сопряжения входного сигнала в координатной области отображается в частотной плоскости как комплексное сопряжение фурье-образа сигнала с одновременной инверсией пространственных частот:

$$s(x_0, y_0) \leftrightarrow s^*(K_x, -K_y). \quad (2.77)$$

Отсюда непосредственно следует, что действительной и четной функции в координатном пространстве соответствует действительный и четный фурье-образ, а

свертке сигналов в координатной плоскости – произведение фурье-спектров в частотном пространстве:

$$s_1(x_0, y_0) \otimes s_2(x_0, y_0) \leftrightarrow s_1(K_x, K_y) \cdot s_2(K_x, K_y). \quad (2.78)$$

Дифференцирование сигнала соответствует умножению его фурье-спектра на комплексный множитель $(jK)^n$, где n – порядок производной:

$$\left\{ \frac{d^n}{dx_0^n} \frac{d^m}{dy_0^m} s(x_0, y_0) \right\} \leftrightarrow (jK_x)^n (jK_y)^m s(K_x, K_y). \quad (2.79)$$

Интегрирование оптического сигнала соответствует делению его фурье-спектра на комплексный множитель jK :

$$\left\{ \int_0^{x_0} \int_0^{y_0} s(\xi, \eta) d\xi d\eta \right\} \leftrightarrow \frac{1}{jK_x} \frac{1}{jK_y} s(K_x, K_y). \quad (2.80)$$

При дифференцировании подавляются спектральные компоненты на низких пространственных частотах и подчеркиваются высокочастотные компоненты, что ведет к оконтуриванию сигнала. Операция интегрирования соответствует подавлению высокочастотных спектральных компонент, в результате чего происходит сглаживание сигнала.

Контрольные вопросы:

1. Локальная и распределенная запись информации.
2. Оптические дисковые системы записи и хранения информации.
3. Магнитооптические технологии, голографические технологии, регистрирующие среды и механизмы записи.
4. Быстродействие, считывание информации в реальном времени - динамическая голография, ассоциативная голографическая память.
5. Аналоговые оптические вычисления, Фурье-голография, голографическая коммутация, мультиплексирование и демультимплексирование сигналов.
6. Оптическая би- и мультстабильность, цифровая оптическая обработка сигналов.

5. Перспективы использования и ограничения квантовой криптографии, квантовых вычислений

Когерентно-оптические системы распознавания образов, оптические нейронные сети, оптические системы нечеткой и нейро-нечеткой логики

Голографическая парадигма в искусственном интеллекте, реализация принципов информатики мозга методами оптоинформатики

Разные люди понимают под шифрованием разные вещи. Дети играют в игрушечные шифры и секретные языки. Это, однако, не имеет ничего общего с настоящей криптографией. Настоящая криптография должна обеспечивать такой уровень секретности, чтобы можно было надежно защитить критическую информацию от расшифровки крупными организациями - такими как мафия, транснациональные корпорации и крупные государства. Настоящая криптография в прошлом использовалась лишь в военных целях. Однако сейчас, со становлением информационного общества, она становится центральным инструментом для обеспечения конфиденциальности.

По мере образования информационного общества, крупным государствам становятся доступны технологические средства тотального надзора за миллионами людей. Поэтому криптография становится одним из основных инструментов обеспечивающих конфиденциальность, доверие, авторизацию, электронные платежи, корпоративную безопасность и бесчисленное множество других важных вещей.

Криптография не является более придумкой военных, с которой не стоит связываться. Настала пора снять с криптографии покровы таинственности и использовать все ее возможности на пользу современному обществу. Широкое распространение криптографии является одним из немногих способов защитить человека от ситуации, когда он вдруг обнаруживает, что живет в тоталитарном государстве, которое может контролировать каждый его шаг.

Криптография – это одна из старейших наук, ее история насчитывает несколько тысяч лет, наука о методах обеспечения конфиденциальности (невозможности прочтения информации посторонним) и аутентичности (целостности и подлинности авторства, а также невозможности отказа от авторства) информации.

Изначально криптография изучала методы шифрования информации - обратимого преобразования открытого (исходного) текста на основе секретного алгоритма или ключа в зашифрованный текст (шифротекст). Традиционная криптография образует раздел симметричных криптосистем, в которых зашифрование и

расшифрование проводится с использованием одного и того же секретного ключа. Помимо этого раздела современная криптография включает в себя системы электронной цифровой подписи (ЭЦП), хеш-функции, управление ключами, получение скрытой информации, квантовую криптографию.

История криптографии насчитывает около 4 тысяч лет. В качестве основного критерия периодизации криптографии, возможно, использовать технологические характеристики используемых методов шифрования.

Первый период (приблизительно с 3-го тысячелетия до н. э.) характеризуется господством моноалфавитных шифров (основной принцип - замена алфавита исходного текста другим алфавитом через замену букв другими буквами или символами).

Второй период (хронологические рамки - с IX века на Ближнем Востоке и с XV века в Европе - до начала XX века) ознаменовался введением в обиход полиалфавитных шифров.

Третий период (с начала и до середины XX века) характеризуется внедрением электромеханических устройств в работу шифровальщиков. При этом продолжалось использование полиалфавитных шифров.

Четвертый период (с середины до 70-х годов XX века), период перехода к математической криптографии. Появляются строгие математические определения количества информации, передачи данных, энтропии, функций шифрования. До 1975 года криптография оставалась «классической», или же, более корректно, криптографией с секретным ключом.

Современный период развития криптографии (с конца 1970-х годов по настоящее время) отличается зарождением и развитием нового направления - криптография с открытым ключом. Её появление знаменуется не только новыми техническими возможностями, но и сравнительно широким распространением криптографии для использования частными лицами (в предыдущие эпохи использование криптографии было исключительной прерогативой государства). Правовое регулирование использования криптографии частными лицами в разных странах сильно различается - от разрешения до полного запрета.

Современная криптография образует отдельное научное направление на стыке математики и информатики - работы в этой области публикуются в научных журналах, организуются регулярные конференции. Практическое применение криптографии стало неотъемлемой частью жизни современного общества - её используют в таких отраслях

как электронная коммерция, электронный документооборот (включая цифровые подписи), телекоммуникации и других.

Открытый текст - в криптографии исходный текст, подлежащий шифрованию, либо получившийся в результате расшифрования. Может быть прочитан без дополнительной обработки. Открытый текст часто является читабельным текстом на одном из человеческих языков, в связи, с чем перед шифрованием (с использованием современной вычислительной техники) его сжимают с целями увеличения производительности и уменьшения избыточности, свойственной человеческим языкам.

Открытым текстом будет называться информация даже в том случае, если она сохранена в нетекстовом виде - например, музыка или звук. Главное, чтобы для использования данной информации не требовалось производить дешифрование.

Шифротекст - результат операции шифрования. Часто также используется вместо термина «криптограмма», хотя последний подчёркивает сам факт передачи сообщения, а не шифрования. Процесс применения операции шифрования к шифротексту называется перешифровкой.

Ключ - секретная информация, используемая криптографическим алгоритмом при шифровке или расшифровке сообщений, постановке и проверке цифровой подписи, вычислении кодов аутентичности. При использовании одного и того же алгоритма результат шифрования зависит от ключа. Для современных алгоритмов сильной криптографии утрата ключа приводит к практической невозможности расшифровать информацию.

Шифр - какая-либо система преобразования текста (код) для обеспечения секретности передаваемой информации. Шифры применяются для тайной переписки дипломатических представителей со своими правительствами, а также в вооруженных силах для передачи текста секретных документов по техническим средствам связи.

Шифр может представлять собой совокупность условных знаков (условная азбука из цифр или букв) либо алгоритм кодирования с использованием обычных цифр и букв. Процесс засекречивания сообщения с помощью шифра называется шифрованием.

Код - совокупность алгоритмов криптографических преобразований (шифрования), отображающих множество возможных открытых данных на множество возможных зашифрованных данных, и обратных им преобразований. Важным параметром любого шифра является ключ - параметр криптографического алгоритма, обеспечивающий выбор одного преобразования из совокупности преобразований, возможных для этого алгоритма.

Шифрование - способ преобразования открытой информации в закрытую и обратно. Применяется для хранения важной информации в ненадёжных источниках или передачи её по незащищённым каналам связи.

Криптоанализ - наука о методах получения исходного значения зашифрованной информации, не имея доступа к секретной информации (ключу), необходимой для этого. В большинстве случаев под этим подразумевается нахождение ключа. В нетехнических терминах, криптоанализ есть взлом шифра (кода). Термин был введён американским криптографом Уильямом Фридманом в 1920 году.

Под термином «криптоанализ» также понимается попытка найти уязвимость в криптографическом алгоритме или протоколе. Хотя основная цель осталась неизменной с течением времени, методы криптоанализа претерпели значительные изменения, эволюционировав от использования лишь ручки и бумаги до широкого применения вычислительных мощностей специализированных криптоаналитических компьютеров в наши дни. Если раньше криптоаналитиками были большей частью лингвисты, то в наше время это удел «чистых» математиков.

Криптоаналитик - специалист по криптоанализу. Одним из первых криптоаналитиков был Аристотель, криптографически вскрывший скиталу - одно из первых известных криптографических устройств. Современный криптоаналитик должен иметь познания в области математики и теории алгоритмов. Часто также является криптографом - специалистом по созданию и использованию средств шифрования.

Дешифрование (дешифровка) - процесс извлечения открытого текста без знания криптографического ключа на основе известного зашифрованного. Термин дешифрование обычно применяют по отношению к процессу криптоанализа шифротекста (криптоанализ сам по себе, может заключаться и в анализе шифросистемы, а не только зашифрованного ею открытого сообщения).

Криптографическая стойкость (или криптостойкость) - способность криптографического алгоритма противостоять возможным атакам на него или оценка алгоритма, способного взломать шифр. Атакующие криптографический алгоритм используют методы криптоанализа. Стойким считается алгоритм, который для успешной атаки требует от противника недостижимых вычислительных ресурсов, недостижимого объёма перехваченных открытых и зашифрованных сообщений или же такого времени раскрытия, что по его истечению защищенная информация будет уже не актуальна, и т. д.

Имитозащита - защита системы шифровальной связи или другой криптосистемы от навязывания ложных данных. Реализуется с помощью добавления к сообщению дополнительного кода, имитовставки, зависящей от содержания сообщения и секретного элемента, известного только отправителю и получателю (ключа). Закладка избыточности позволяет обнаружить внесённые в сообщение несанкционированные изменения.

В Российской Федерации коммерческая деятельность, связанная с использованием криптографических средств, подлежит обязательному лицензированию. С 22 января 2008 года действует Постановление Правительства РФ от 29 декабря 2007 N 957 «Об утверждении положений о лицензировании отдельных видов деятельности, связанных с шифровальными (криптографическими) средствами», которым приняты Положения о лицензировании деятельности по:

1. Распространению шифровальных (криптографических) средств;
2. Техническому обслуживанию шифровальных (криптографических) средств;
3. Предоставлению услуг в области шифрования информации;
4. Разработке, производству шифровальных (криптографических) средств, защищенных с использованием шифровальных (криптографических) средств информационных и телекоммуникационных систем.

Следует отметить, что приложения к данному Постановлению содержат жёсткие требования к лицу-соискателю лицензии, включая его образование, квалификацию, стаж, требования к помещению, охране, информационной и эксплуатационной безопасности при разработке и реализации средств. К примеру, требуется «наличие в штате у следующего квалифицированного персонала: руководитель или лицо, уполномоченное руководить работами по лицензируемой деятельности, имеющие высшее профессиональное образование или профессиональную подготовку в области информационной безопасности, а также стаж работы в этой области не менее 5 лет; инженерно-технические работники, имеющие высшее профессиональное образование или прошедшие переподготовку в области информационной безопасности с получением специализации, необходимой для работы с шифровальными (криптографическими) средствами».

В настоящее время действует также Приказ ФСБ России от 9 февраля 2005 г. N 66 «Об утверждении положения о разработке, производстве, реализации и эксплуатации шифровальных (криптографических) средств защиты информации».

(положение пкз-2005)», который определяет порядок разработки и эксплуатации криптографических средств.

В частности, согласно приказу, средства криптографии реализуются «юридическим лицом или индивидуальным предпринимателем, имеющим право на осуществление данного вида деятельности, связанного с шифровальными (криптографическими) средствами вместе с правилами пользования ими, согласованными с ФСБ России».

Ранее был издан Указ Президента РФ от 3 апреля 1995 N 334 «О мерах по соблюдению законности в области разработки, производства, реализации и эксплуатации шифровальных средств, а также предоставления услуг в области шифрования информации», постановивший «Запретить использование государственными организациями и предприятиями в информационно-телекоммуникационных системах шифровальных средств, включая криптографические средства обеспечения подлинности информации (электронная подпись), и защищенных технических средств хранения, обработки и передачи информации, не имеющих сертификата Федерального агентства правительственной связи и информации при Президенте Российской Федерации, а также размещение государственных заказов на предприятиях, в организациях, использующих указанные технические и шифровальные средства, не имеющие сертификата Федерального агентства правительственной связи и информации при Президенте Российской Федерации».

Криптографический протокол - это абстрактный или конкретный протокол, включающий набор криптографических алгоритмов. В основе протокола лежит набор правил, регламентирующих использование криптографических преобразований и алгоритмов в информационных процессах.

Функции криптографических протоколов

- Аутентификация источника данных
- Аутентификация сторон
- Конфиденциальность данных
- Невозможность отказа
- Невозможность отказа с доказательством получения
- Невозможность отказа с доказательством источника
- Целостность данных
- Обеспечение целостности соединения без восстановления
- Обеспечение целостности соединения с восстановлением
- Разграничение доступа

Классификация

Протоколы шифрования и расшифрования. В основе протокола этого класса содержится некоторый симметричный или асимметричный алгоритм шифрования и расшифрования. Алгоритм шифрования выполняется на передаче отправителем сообщения, в результате чего сообщение преобразуется из открытой формы в зашифрованную. Алгоритм расшифрования выполняется на приеме получателем, в результате чего сообщение преобразуется из зашифрованной формы в открытую. Так обеспечивается свойство конфиденциальности.

Для обеспечения свойства целостности передаваемых сообщений симметричные алгоритмы шифрования и расшифрования, обычно, совмещаются с алгоритмами вычисления имитозащитной вставки (ИЗВ) на передаче и проверки ИЗВ на приеме, для чего используется ключ шифрования. При использовании асимметричных алгоритмов шифрования и расшифрования свойство целостности обеспечивается отдельно путем вычисления электронной цифровой подписи (ЭЦП) на передаче и проверки ЭЦП на приеме, чем обеспечиваются также свойства безотказности и аутентичности принятого сообщения.

Протоколы электронной цифровой подписи (ЭЦП). В основе протокола этого класса содержится некоторый алгоритм вычисления ЭЦП на передаче с помощью секретного ключа отправителя и проверки ЭЦП на приеме с помощью соответствующего открытого ключа, извлекаемого из открытого справочника, но защищенного от модификаций. В случае положительного результата проверки протокол, обычно, завершается операцией архивирования принятого сообщения, его ЭЦП и соответствующего открытого ключа. Операция архивирования может не выполняться, если ЭЦП используется только для обеспечения свойств целостности и аутентичности принятого сообщения, но не безотказности. В этом случае, после проверки, ЭЦП может быть уничтожена сразу или по прошествии ограниченного промежутка времени ожидания.

Протоколы идентификации и аутентификации. В основе протокола идентификации содержится некоторый алгоритм проверки того факта, что идентифицируемый объект (пользователь, устройство, процесс, и т.д.), предъявивший некоторое имя (идентификатор), знает секретную информацию, известную только заявленному объекту, причем метод проверки является, конечно, косвенным, то есть без предъявления этой секретной информации.

Обычно с каждым именем (идентификатором) объекта связывается перечень его прав и полномочий в системе, записанный в защищенной базе данных. В этом случае

протокол идентификации может быть расширен до протокола аутентификации, в котором идентифицированный объект проверяется на правомочность заказываемой услуги.

Если в протоколе идентификации используется ЭЦП, то роль секретной информации играет секретный ключ ЭЦП, а проверка ЭЦП осуществляется с помощью открытого ключа ЭЦП, знание которого не позволяет определить соответствующий секретный ключ, но позволяет убедиться в том, что он известен автору ЭЦП.

Протоколы аутентифицированного распределения ключей. Протоколы этого класса совмещают аутентификацию пользователей с протоколом генерации и распределения ключей по каналу связи. Протокол имеет двух или трёх участников; третьим участником является центр генерации и распределения ключей (ЦГРК), называемый для краткости сервером S. Протокол состоит из трёх этапов, имеющих названия: генерация, регистрация и коммуникация. На этапе генерации сервер S генерирует числовые значения параметров системы, в том числе, свой секретный и открытый ключ. На этапе регистрации сервер S идентифицирует пользователей по документам (при личной явке или через уполномоченных лиц), для каждого объекта генерирует ключевую или идентификационную информацию и формирует маркер безопасности, содержащий необходимые системные константы и открытый ключ сервера S (при необходимости). На этапе коммуникации реализуется собственно протокол аутентифицированного ключевого обмена, который завершается формированием общего сеансового ключа.

Задачи

- Обеспечение различных режимов аутентификации
- Генерация, распределение и согласование криптографических ключей
- Защита взаимодействий участников
- Разделение ответственности между участниками

Разновидности атак на протоколы

- Атаки, направленные против криптографических алгоритмов
- Атаки против криптографических методов, применяемых для реализации

протоколов

- Атаки против самих протоколов (активные или пассивные)

6. Требования к безопасности протокола

1. Аутентификация (нешироковещательная):

- аутентификация субъекта
 - аутентификация сообщения
 - защита от повтора
2. Аутентификация при рассылке по многим адресам или при подключении к службе подписки/уведомления:
 - неявная (скрытая) аутентификация получателя
 - аутентификация источника
 3. Авторизация (доверенной третьей стороной)
 4. Свойства совместной генерации ключа:
 - аутентификация ключа
 - подтверждение правильности ключа
 - защищенность от чтения назад
 - формирование новых ключей
 - защищенная возможность договориться о параметрах безопасности
 5. Конфиденциальность
 6. Анонимность:
 - защита идентификаторов от прослушивания (несвязываемость)
 - защита идентификаторов от других участников
 7. Ограниченная защищенность от атак типа «отказ в обслуживании»
 8. Инвариантность отправителя
 9. Невозможность отказа от ранее совершенных действий:
 - подотчетность
 - доказательство источника
 - доказательство получателя
 10. Безопасное временное свойство

7. Современная криптография

Для современной криптографии характерно использование открытых алгоритмов шифрования, предполагающих использование вычислительных средств. Известно более десятка проверенных алгоритмов шифрования, которые при использовании ключа достаточной длины и корректной реализации алгоритма криптографически стойки. Распространенные алгоритмы:

- симметричные DES, AES, ГОСТ 28147-89, Camellia, Twofish, Blowfish, IDEA, RC4 и др.;
- асимметричные RSA и Elgamal (Эль - Гамаль);
- хэш-функций MD4, MD5, MD6, SHA-1, SHA-2, ГОСТ Р 34.11-94.

Во многих странах приняты национальные стандарты шифрования. В 2001 году в США принят стандарт симметричного шифрования AES на основе алгоритма Rijndael с длиной ключа 128, 192 и 256 бит. Алгоритм AES пришёл на смену прежнему алгоритму DES, который теперь рекомендовано использовать только в режиме Triple DES. В Российской Федерации действует стандарт ГОСТ 28147-89, описывающий алгоритм блочного шифрования с длиной ключа 256 бит, а также алгоритм цифровой подписи ГОСТ Р 34.10-2001.

8. Криптоанализ

Криптоанализ - наука о методах получения исходного значения зашифрованной информации, не имея доступа к секретной информации (ключу), необходимой для этого. В большинстве случаев под этим подразумевается нахождение ключа. В нетехнических терминах, криптоанализ есть взлом шифра (кода). Термин был введён американским криптографом Уильямом Ф. Фридманом в 1920 году.

Под термином «криптоанализ» также понимается попытка найти уязвимость в криптографическом алгоритме или протоколе. Хотя основная цель осталась неизменной с течением времени, методы криптоанализа претерпели значительные изменения, эволюционировав от использования лишь ручки и бумаги до широкого применения вычислительных мощностей специализированных криптоаналитических компьютеров в наши дни. Если раньше криптоаналитиками были большей частью лингвисты, то в наше время это удел «чистых» математиков.

Результаты криптоанализа конкретного шифра называют криптографической атакой на этот шифр. Успешную криптографическую атаку, дискредитирующую атакуемый шифр, называют взломом или вскрытием.

Классический криптоанализ. Хотя понятие криптоанализ было введено сравнительно недавно, некоторые методы взлома были изобретены десятки веков назад. Первым известным письменным упоминанием о криптоанализе является «Манускрипт о дешифровке криптографических сообщений», написанный арабским учёным Ал-Кинди ещё в 9 веке. В этом научном труде содержится описание метода частотного анализа.

Частотный анализ - основной инструмент для взлома большинства классических шифров перестановки или замены. Данный метод основывается на предположении о существовании нетривиального статистического распределения символов, а также их последовательностей одновременно и в открытом тексте, и в шифротексте. Причём данное распределение будет сохраняться с точностью до замены символов как в процессе шифрования, так и в процессе дешифрования. Стоит отметить, что при условии достаточно большой длины шифрованного сообщения моноалфавитные шифры легко поддаются частотному анализу: если частота появления буквы в языке и частота появления некоторого присутствующего в шифротексте символа приблизительно равны, то в этом случае с большой долей вероятности можно предположить, что данный символ и будет этой самой буквой. Самым простым примером частотного анализа может служить банальный подсчёт количества каждого из встречающихся символов, затем следуют процедуры деления полученного числа символов на количество всех символов в тексте и умножение результата на сто, чтобы представить окончательный ответ в процентах. Далее полученные процентные значения сравниваются с таблицей вероятностного распределения букв для предполагаемого языка оригинала.

Современный криптоанализ. По мере развития новых методов шифрования математика становилась всё более и более значимой. Так, например, при частотном анализе криптоаналитик должен обладать знаниями и в лингвистике, и в статистике. В то время как теоретические работы по криптоанализу Энигмы выполнялись преимущественно математиками, например, Аланом Матисоном Тьюрингом. Тем не менее благодаря всё той же математике криптография достигла такого развития, что количество необходимых для взлома элементарных математических операций стало достигать астрономических значений. Современная криптография стала гораздо более устойчивой к криптоанализу, чем некогда используемые, устаревшие методики, для взлома которых было достаточно ручки и листа бумаги. Может показаться, что чистый теоретический криптоанализ не способен более эффективно взламывать современные шифры.

Бандитский криптоанализ. Криптоаналитик может использовать так называемый «человеческий фактор», т.е. пытаться с помощью шантажа, подкупа, пыток или иных способов получить информацию о системе шифрования или даже сам ключ шифрования. Например, дача взятки, как одна из разновидностей бандитского криптоанализа, может носить название «Вскрытие с покупкой ключа». Таким образом

методика вскрытия построена на слабости людей как составной части системы защиты информации.

Бандитский криптоанализ считается очень мощным способом взлома системы, а зачастую и наилучшим путём вскрытия шифров.

Атака на основе шифротекста. Допустим, криптоаналитик обладает некоторым числом шифротекстов, полученных в результате использования одного и того же алгоритма шифрования. В этом случае криптоаналитик может совершить только атаку на основе шифротекста. Целью криптографической атаки в этом случае является нахождение как можно большего числа открытых текстов, соответствующих имеющимся шифротекстам, или, что ещё лучше, нахождение используемого при шифровании ключа.

Входные данные для подобного типа атак криптоаналитик может получить в результате простого перехвата зашифрованных сообщений. Если передача осуществляется по открытому каналу, то реализация задачи по сбору данных сравнительно легка и тривиальна. Атаки на основе шифротекста являются самыми слабыми и неудобными.

Атака на основе адаптивно подобранного шифротекста. Атака такого типа является более удобным частным случаем атаки на основе подобранного открытого текста. Удобство атаки на основе адаптивно подобранного шифротекста состоит в том, что помимо возможности выбирать шифруемый текст, криптоаналитик может принять решение о шифровании того или иного открытого текста на основе уже полученных результатов операций шифрования. Другими словами, при осуществлении атаки на основе подобранного открытого текста криптоаналитик выбирает всего один большой блок открытого текста для последующего шифрования, а потом на основе этих данных начинает взламывать систему. В случае организации адаптивной атаки криптоаналитик может получать результаты шифрования любых блоков открытого текста, чтобы собрать интересующие его данные, которые будут учтены при выборе следующих отправляемых на шифрование блоков открытого текста и так далее. Наличие обратной связи даёт атаке на основе адаптивно подобранного шифротекста преимущество перед всеми вышеперечисленными типами атак.

Атака на основе подобранного ключа. Вопреки своему названию атака на основе подобранного ключа не подразумевает под собой того, что криптоаналитик занимается простым перебором ключей в надежде найти нужный. Атака такого типа строится на том, что криптоаналитик может наблюдать за работой алгоритма шифрования, в котором используются несколько ключей. Криптоаналитик изначально ничего не знает

о точном значении ключей, зато ему известно некоторое математическое отношение, связывающее между собой ключи. Примером тому может служить ситуация, когда криптоаналитик выяснил, что последние 80 битов у всех ключей одинаковы, хотя сами значения битов могут быть неизвестными.

Атака на основе подобранного открытого текста. Для осуществления такого типа атаки криптоаналитику необходимо иметь не только какое-то количество открытых текстов и полученных на их основе шифротекстов. Помимо прочего в данном случае криптоаналитик должен обладать возможностью подобрать несколько открытых текстов и получить результат их шифрования. Задачи криптоаналитика повторяют задачи для атаки на основе открытого текста, то есть получить ключ шифрования, либо создать дешифрующий алгоритм для данного ключа.

Получить входные данные для такого вида атаки можно, например, следующим образом:

1. Создать и отправить поддельное не зашифрованное сообщение якобы от одного из пользователей, которые обычно пользуются шифрованием.

2. В некоторых случаях можно получить ответ, в котором будет содержаться зашифрованный текст, цитирующий содержание поддельного сообщения.

При осуществлении атаки подобного типа криптоаналитик имеет возможность подбирать блоки открытого текста, что при определённых условиях может позволить получить больше информации о ключе шифрования.

Атака на основе открытых текстов и соответствующих шифротекстов. Пусть в распоряжении криптоаналитика есть не только шифротексты, но и соответствующие им открытые тексты.

Тогда существуют два варианта постановки задачи:

1. Найти ключ, использованный для преобразования открытого текста в шифротекст

2. Создать алгоритм, способный дешифровать любое сообщение, закодированное с помощью этого ключа

Получение открытых текстов играет решающую роль в осуществлении этой атаки. Открытые тексты извлекают из самых различных источников. Так, например, можно догадаться о содержимом файла по его расширению.

В случае взлома переписки можно сделать предположение, что письмо имеет структуру типа:

- «Приветствие»
- «Основной текст»

- «Заключительная форма вежливости»
- «Подпись»

Следовательно, атака может быть организована путём подбора различных видов «Приветствия» (например, «Здравствуйте!», «Добрый день» и т. д.) или «Заключительной формы вежливости» (таких как «С уважением», «Искренне Ваш» и т. п.). Легко заметить, что данная атака сильнее атаки на основе одного лишь шифротекста.

Атака на основе подобранного шифротекста. Допустим, что у криптоаналитика имеется временный доступ к дешифрующему средству или устройству. В таком случае за ограниченный промежуток времени криптоаналитик может получить из известных ему шифротекстов соответствующие им открытые тексты, после чего криптоаналитику нужно будет приступить к взлому системы. При осуществлении подобного типа атаки цель взлома - получить ключ шифрования.

Криптографические генераторы случайных чисел производят случайные числа, которые используются в криптографических приложениях, например - для генерации ключей. Обычные генераторы случайных чисел, имеющиеся во многих языках программирования и программных средах, не подходят для нужд криптографии (они создавались с целью получить статистически случайное распределение, криптоаналитики могут предсказать поведение таких случайных генераторов).

В идеале случайные числа должны основываться на настоящем физическом источнике случайной информации, которую невозможно предсказать. Примеры таких источников включают шумящие полупроводниковые приборы, младшие биты оцифрованного звука, интервалы между прерываниями устройств или нажатиями клавиш. Полученный от физического источника шум затем "дистиллируется" криптографической хэш-функцией так, чтобы каждый бит зависел от каждого бита. Достаточно часто для хранения случайной информации используется довольно большой пул (несколько тысяч бит) и каждый бит пула делается зависимым от каждого бита шумовой информации и каждого другого бита пула криптографически надежным (strong) способом.

Когда нет настоящего физического источника шума, приходится пользоваться псевдослучайными числами. Такая ситуация нежелательна, но часто возникает на компьютерах общего назначения. Всегда желательно получить некий шум окружения -- скажем от величины задержек в устройствах, цифры статистики использования ресурсов, сетевой статистики, прерываний от клавиатуры или чего-то иного. Задачей является получить данные, непредсказуемые для внешнего наблюдателя. Для

достижения этого случайный пул должен содержать как минимум 128 бит настоящей энтропии.

Контрольные вопросы:

1. Перспективы использования и ограничения квантовой криптографии,
2. Физические основы квантовых вычислений.
3. Когерентно-оптические системы распознавания образов.
4. Оптические нейронные сети, оптические системы нечеткой и нейро-нечеткой логики.
5. Голографическая парадигма в искусственном интеллекте, реализация принципов информатики мозга методами оптоинформатики.