

**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ**

Утверждаю:
Зав. каф. РЗИ
_____ Задорин А.С.

**А.С.Задорин
СРЕДСТВА И СИСТЕМЫ ТЕХНИЧЕСКОГО ОБЕСПЕЧЕНИЯ ОБРАБОТ-
КИ, ХРАНЕНИЯ И ПЕРЕДАЧИ ИНФОРМАЦИИ**

КУРС ЛЕКЦИЙ
для студентов, обучающихся по специальности 090103
«Организация и технология защиты информации»

ТОМСК 2006

Оглавление

Введение.....	5
1. Сигналы и способы их представления	6
1.1. Основные понятия и терминология.....	6
1.2. Количество информации в дискретном сообщении	7
1.3. Энтропия.....	9
1.4. Спектр сигнала.....	11
1.4.1. Некоторые общие соотношения между сигналом и его спектром	12
1.4.2. Частотный отклик линейной стационарной системы.....	14
1.5. Импульсный отклик.....	14
1.5.1. Отклик системы на сигнал произвольного вида.....	15
1.6. Детерминированные сигналы	16
1.7. Случайные сигналы	19
1.7.1. Характеристики случайных сигналов	21
1.7.2. Корреляционные характеристики случайных сигналов	23
1.7.3. Спектральные характеристики случайных сигналов	24
1.8. Представление непрерывных сигналов отсчетами.....	25
2. Способы модуляции сигналов	28
2.1. Амплитудная модуляция (АМ).....	29
2.2. Частотная модуляция (ЧМ).....	29
2.3. Фазовая модуляция (ФМ).....	30
3. Классификация электрических цепей	32
3.1. Линейные и нелинейные цепи	32
3.2. Характеристики пассивных элементов электрических цепей	34
3.3. Нелинейные искажения сигналов в электрических цепях	34
4. Генерация электрических сигналов.....	35
4.1. Принципы построения схем генераторов электрических сигналов	35
4.2. Генераторы гармонических колебаний.....	37
4.3. Стабилизация частоты генераторов	38
4.4. Схемы кварцевых генераторов	41
4.5. Релаксационные генераторы.....	41
5. Акусто-электрические конверторы сигналов	42
5.1. Распространение акустических волн.....	42
5.2. Возбуждение и прием акустических волн	47
5.3. Другие способы возбуждения и приема звуковых волн.....	49
5.4. Основные характеристики преобразователей акустических волн	50
5.5. Электрическое и акустическое согласование преобразователей.....	51
6. Элементы оптоэлектроники и инфракрасной техники.....	54
6.1. Источники оптического излучения для систем передачи	55
6.2. Светодиодающие диоды.....	56
6.2.1. Конструкции светодиодов для оптической связи	56
6.2.2. Принцип действия светодиода.....	57
6.2.3. Основные характеристики светодиодов.....	57
6.3. Лазеры для систем связи	59
6.4. Фотоприёмники оптических систем связи	62
6.4.1. Характеристики фотоприемников	62
6.4.2. Принцип работы PIN-фотодиода	64
6.4.3. Лавинные фотодиоды	66
6.5. Оптические волокна	68

6.5.1. Конструкция оптических волокон	68
6.5.2. Изготовление оптических волокон	69
6.5.3. Оптические потери в световолокне	70
6.5.4. Волноводные моды в оптическом волокне	71
6.5.5. Дисперсия групповых скоростей	75
6.5.6. Типы оптических волокон	77
7. Системы приема и передачи информации	80
7.1. Обобщенная схема аппаратуры передачи цифровой информации	80
7.2. Параметры и характеристики каналов связи	83
7.3. Классификация систем радиосвязи по частотному диапазону	85
7.4. Иерархический принцип построения цифровых телекоммуникационных систем	87
7.5. Системы плезихронных цифровых иерархий	91
7.5.1. Общие особенности плезихронных цифровых иерархий	91
7.5.2. Недостатки плезихронной цифровой иерархии	91
7.6. Синхронная цифровая иерархия – SDH	92
7.6.1. Схема мультиплексирования PDH сигналов в технологии SDH	93
7.6.2. Структура цикла SDH	94
7.6.3. Сети SDH	95
7.7. Преобразование сигналов в ЦСП	98
7.7.1. Принципы цифрового преобразования аналоговых сигналов	98
7.7.2. Оценка отношения уровня сигнала к шуму квантования	103
7.8. Волоконно-оптические системы передачи (ВОСП)	107
7.8.1. История развития ВОСП	107
7.8.2. Структура цифровых волоконно-оптических линейных трактов [2]	111
7.8.3. Помехоустойчивость цифровых ВОСП	114
7.8.4. Оценка коэффициента битовых ошибок	115
7.8.5. Помехоустойчивость ВОСП в условиях гауссовых шумов	115
7.9. Оптические системы передачи со спектральным уплотнением	116
8. Компьютерные сети	119
8.1. Принцип построения компьютерных сетей	119
8.2. Локальные вычислительные сети (ЛВС)	121
8.2.1. Основные типы конфигурации ЛВС	121
8.2.2. Методы доступа к моноканалу	124
8.2.3. Адаптеры и приемопередатчики ЛВС	128
8.2.4. Управление передачей кадров	130
8.2.5. Стандарты ЛВС	132
8.2.6. Технология Ethernet	133
8.3. Глобальные компьютерные сети	138
8.3.1. Характеристика глобальных компьютерных сетей	138
8.2.2. Сеть Internet	139
8.2.3. Услуги Internet	140
8.2.4. Возможности сети Internet	140
8.2.4. Доступ к информационным ресурсам	141
8.2.5. Адресация и протоколы в Интернет	141
8.2.6. Протокол Frame Relay (FR)	142
9. Антенны	143
9.1. Назначение и классификация антенн	143
9.1.1. Назначение антенны	143
9.1.2. Классификация антенн	143
9.2. Первичные параметры и характеристики передающих антенн	146
9.2.1. Сопротивление излучения	147
9.2.2. Поле излучения диполя Герца	147
9.2.3. Входное сопротивление антенны	148
9.2.4. Характеристика направленности антенны	149
9.2.5. Диаграмма направленности (ДН) антенны	149
9.2.6. Предельно пропускаемая мощность	152

10. Оптическая голография	152
10.1. Физические принципы голографии	152
10.1.1. Интерференция	153
10.1.2. Дифракция	155
10.2. Формирование голограмм	156
10.3. Восстановление волнового фронта	157
10.4. Основное уравнение голографии	157
10.5. Основные типы голограмм	158
10.5.1. Двумерная и объемная голограммы	158
11. Список литературы	161

ВВЕДЕНИЕ

Дисциплина «Средства и системы технического обеспечения обработки, хранения и передачи информации» предусмотрена действующим государственным образовательным стандартом (ГОС) в качестве общепрофессиональной для студентов дневной формы обучения специальности № 090103 - Организация и технология защиты информации. Предметом ее изучения являются средства и системы технического обеспечения обработки, хранения и передачи информации.

Цель дисциплины – дать систематический обзор существующих средств и систем технического обеспечения обработки, хранения и передачи информации; изучить и освоить принципы их построения, рассмотреть перспективные направления развития существующих систем.

В результате изучения данной дисциплины студент должен знать:

- основные типы электронных средств генерации и преобразования сигналов;
- преобразователи спектра сигналов.;
- акусто-электрические и электроакустические конверторы энергии сигналов;
- элементы оптоэлектроники и инфракрасной техники;
- методы и средства записи, хранения и воспроизведения информации на магнитных носителях;
- голографические носители информации;
- основные виды антенно-фидерных устройств, радиопередатчиков и радиоприемников;
- системы передачи и приема видеoinформации, звуковой и цифровой информации.;
- организацию связи с помощью ЭВМ, телекоммуникационные сети;
- методы и средства измерения сигналов в электрических цепях.;
- основы измерений информативных характеристик электромагнитных полей.

и уметь:

на основе электронных средств осуществлять генерацию сигналов; измерять сигналы в электрических цепях;

- преобразовывать спектры сигналов;
- осуществлять запись, хранение и воспроизведение информации;
- обеспечивать организацию связи с помощью ЭВМ;
- создавать телекоммуникационные сети.

В пособии принята нумерация подразделов в пределах отдельного подраздела. При ссылке в каком-либо из подразделов на формулу или рисунок другого подраздела перед их номером пишутся разделенные точкой номера подраздела. Перед ссылкой на формулу или ри-

сунк того же подраздела указывается только номер формулы

1. СИГНАЛЫ И СПОСОБЫ ИХ ПРЕДСТАВЛЕНИЯ

1.1. Основные понятия и терминология

Информация - это совокупность сведений или данных о каких-либо событиях, явлениях или предметах, то есть это совокупность знаний об окружающем нас мире. В отличие от материального и энергетического ресурсов, информационный ресурс не уменьшается при потреблении, а накапливается со временем, и при помощи технических средств может обрабатываться, храниться и передаваться на значительные расстояния. *Передача и хранение информации* осуществляется с помощью различных знаков (символов), которые позволяют представить её в некоторой форме.

Сообщение - это совокупность знаков, отображающих ту или иную информацию. Передача сообщений (а следовательно, и информации) на расстояние осуществляется с помощью какого-либо материального носителя, например, бумаги или магнитной ленты или физического процесса, например, звуковых или электромагнитных волн, тока и т.д.

Модуляция - это процесс изменения параметров носителя информации.

Сигнал - это физический процесс, отображающий (несущий) передаваемое сообщение. В качестве сигналов в настоящее время в основном используются электрические и оптические сигналы. Сигнал передаёт (развёртывает) сообщение во времени, то есть всегда является функцией времени.

Дискретный или дискретный по уровню (амплитуде) сигнал - это сигнал, принимающий по величине (амплитуде) только определённые дискретные значения.

Непрерывный или аналоговый сигнал - это сигнал, который может принимать любые уровни значений в некотором интервале величин.

Дискретный по времени сигнал - это сигнал, заданный только в определённые моменты времени.

Непрерывный по времени сигнал - это сигнал, заданный на всей оси времени.

Цифровой сигнал - это дискретный сигнал по уровню и времени, причём число дискретных значений уровней у него конечно. Так как в этом случае уровни дискретного сигнала можно пронумеровать числами с конечным числом разрядов, то такой дискретный сигнал и называется цифровым.

Ниже будут рассматриваться системы связи, предназначенные для передачи только *цифровых сообщений*. Информация, содержащаяся в таком сообщении, передаётся с помощью цифровых сигналов от источника сообщения (ИС) к получателю сообщения (ПС) по каналу передачи цифровых сообщений (КПЦС), который является трактом передачи информации и обеспечивает связь между источником и получателем сообщений.

Поскольку объектом, переносящим сообщение по таким системам связи, является цифровой сигнал, то техника такой связи по существу является техникой транспортирования (передачи) цифровых сигналов по каналам связи.

Основными параметрами сигнала, наиболее существенными с точки зрения его передачи по каналам связи, являются следующие:

- **Длительность сигнала T** - определяет интервал времени, в пределах которого сигнал существует, а так как любой сигнал рассматривается как временной процесс то он имеет начало и конец.
- **Динамический диапазон D** - это отношение наибольшей мгновенной мощности сигнала к той наименьшей мощности, которую необходимо отличать от нуля при заданном качестве передачи.
- **Ширина спектра сигнала P** - это диапазон частот в пределах которого сосредоточена его основная энергия. Она определяет скорость изменения сигнала внутри интервала его существования.

В технике связи спектр передаваемого сигнала часто сознательно сокращают, для экономии средств на аппаратуру линии связи. Возможную величину сокращения спектра сигнала при этом определяют исходя из допустимого искажения сигнала в данной системе связи. Например, при телефонной связи требуется, чтобы речь была настолько разборчива, чтобы абоненты могли узнать друг друга по голосу. Для этого достаточно передать речевой сигнал в полосе частот 300-3400 Гц. Передача более широкого спектра речи в этом случае нецелесообразна, так как ведёт к техническим осложнениям и увеличению затрат на канал связи. Спектр модулированного сигнала обычно шире спектра передаваемого сообщения и зависит от вида модуляции.

С помощью приведённых выше параметров можно ввести общую и наглядную характеристику передаваемых по каналам связи сигналов - **объём сигнала**:

$$V_c = T_c \cdot F_c \cdot D_c \quad (1)$$

Объём сигнала V_c даёт общее представление о возможностях данного множества сигналов как переносчиков сообщений. Чем больше объём сигнала, тем больше информации можно «вложить» в этот объём и тем труднее передать такой сигнал по каналу связи с требуемым качеством.

1.2. Количество информации в дискретном сообщении

Чтобы иметь возможность сравнивать между собой различные источники сообщений, каналы и другие элементы системы связи, необходимо ввести количественную меру, которая позволяла бы оценивать, содержащуюся в сообщениях, сигналах информацию.

Строгие методы такой оценки предложены в 1948 г. К. Шенноном, что послужило началом построения теории информации.

Дискретный источник выдает сообщение a , принадлежащее некоторому конечному ансамблю $A(a \in A)$. Определим количество информации, содержащееся в этом сообщении, используя три исходных очевидных требования:

- 1) количество информации должно быть аддитивной величиной, т. е. в двух независимых сообщениях количество информации определяется как сумма количеств информации в каждом из них;
- 2) количество информации в сообщении о достоверном событии равно 0;
- 3) количество информации не должно зависеть от качественного содержания сообщения (степени важности, возможных последствий его передачи, эмоциональной окраски и т. п.).

В общем случае сообщение a из ансамбля A характеризуется вероятностью $p(a)$, что источник формирует или посылает это сообщение, т. е. количество информации $i(a)$, содержащейся в сообщении a должно быть функцией от вероятности $p(a)$.

Вспользуемся далее первым требованием аддитивности. Пусть a_1 и a_2 — два независимых сообщения. Вероятность $p(a_1, a_2)$ того, что источник выдаст одно за другим эти два сообщения

$$p(a_1, a_2) = p(a_1)p(a_2), \quad (1)$$

где $p(a_1)$, $p(a_2)$ - вероятности формирования сообщения a_1 и a_2 соответственно.

Общее количество информации $i(a_1, a_2)$, содержащейся в этих двух сообщениях, согласно условию аддитивности определяется как сумма количеств информации в каждом из них:

$$i(a_1, a_2) = i(a_1) + i(a_2). \quad (2)$$

Таким образом, надо найти функцию от вероятности p такую, чтобы при перемножении двух аргументов значения функции складывались. Этому условию удовлетворяет только логарифмическая функция

$$i(a) = k \log[p(a)], \quad (3)$$

где k — произвольный коэффициент.

Логарифм, вообще говоря, может быть взят по любому основанию. Эта формула может быть использована для определения количества информации, содержащейся в сообщении a . Эта формула удовлетворяет и требованию 2): в случае достоверного события вероятность сообщения $p(a) = 1$. Тогда количество информации согласно полученной формуле:

$$i(a) = k \log 1 = k \cdot 1 = 0, \quad (4)$$

Поскольку $p(a) \leq 1$, и следовательно, $\log p(a) \leq 0$, то, чтобы измерять количество информации неотрицательными числами, выбираем значение коэффициента $k=-1$:

$$i(a) = -\log p(a) = \log \frac{1}{p(a)}. \quad (5)$$

Основание логарифма чаще всего в формуле для определения количества информации выбирают равным двум. Получаемая при этом единица информации носит название *двоичная единица*, или *бит*. Она равна количеству информации, содержащейся в сообщении о событии, происходящем с вероятностью $p(a) = 0.5$:

$$i(a) = \log_2 \frac{1}{0,5} = \log_2 2 = 1 \text{ бит}. \quad (6)$$

Такая единица наиболее удобна потому, что в современной вычислительной технике, технике связи широко используются двоичные коды, двоичные дискретные устройства.

В дальнейшем, записывая обозначение \log , будем подразумевать, что речь идет о двоичном логарифме:

$$i(a) = \log_2 \frac{1}{p(a)}. \quad (7)$$

Количество информации в сообщении тем больше, чем оно менее вероятно или чем оно более неожиданно.

1.3. Энтропия

Количество информации является случайной величиной, поскольку сами сообщения случайны. Очевидно, что для количества информации существует свое распределение вероятностей, которое зависит от распределения вероятностей сообщения в ансамбле. Поэтому удобнее для характеристики всего ансамбля (источника сообщения) использовать математическое ожидание количества информации, которое называют *энтропией*:

$$H(A) = M[i(a)] = M \left[\log \frac{1}{p(a)} \right], \quad (1)$$

где $H(A) = M \left[\log \frac{1}{p(a)} \right]$ — энтропия дискретного источника сообщения.

Усреднение производится по всему ансамблю, с учетом всех вероятностных связей между сообщениями. Чем больше энтропия источника, тем больше степень неожиданности передаваемых им сообщений в среднем, т. е. тем больше неопределенность ожидаемого сообщения в среднем. Поэтому энтропию часто называют *мерой неопределенности сообщений*. При этом под неопределенностью понимают неопределенность, существующую до того как сообщение принято. После приема всякая неопределенность устраняется. Количество информации

можно расценивать как меру уменьшения неопределенности.

Энтропия — основная информационная характеристика источника сообщений. Чем больше энтропия, тем труднее запомнить сообщение и передать его по каналу связи. Как правило, чем больше энтропия, тем больше энергетические затраты на передачу сообщения.

Основные свойства энтропии следующие:

энтропия неотрицательна $H(A) \geq 0$ и равна 0 только для "вырожденного" ансамбля, когда одно сообщение передается с вероятностью $p(a) = 1$, а другие с нулевой вероятностью;

энтропия аддитивна; это приводит к тому, что если рассматривать последовательность n сообщений как одно укрупненное, то энтропия источника таких укрупненных сообщений будет в n раз больше по сравнению с энтропией исходного источника;

если ансамбль содержит k различных сообщений (k — объем алфавита источника), то $H(A) \leq \log k$ В частном случае двоичного источника без памяти ($k=2$; сообщения передаются

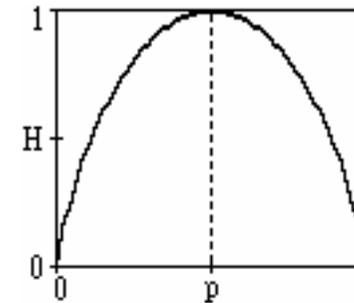


Рис.1. Зависимость энтропии от вероятности сообщения

статически независимо друг от друга) энтропия максимальна в том случае, когда $p(a_1) = p(a_2) = 0,5$, при этом $H(A) = \log k = \log 2 = 1$. В этом случае график зависимости энтропии от вероятности появления того или иного сообщения примет вид, представленный на рис. 1. Для данного источника без памяти с объемом алфавита k

$$H(A) = M \left[\log \frac{1}{p(a)} \right] = \sum_{k=1}^k p(a_k) \log \frac{1}{p(a_k)}. \quad (2)$$

Пусть, например, $k=32$ (модель источника-алфавита русской или английской речи). Тогда при условии равновероятности выбора букв по последней формуле определим:

$$H(A) = \sum_1^{32} \frac{1}{32} \log_2 \frac{1}{1/32} = 32 \cdot \frac{1}{32} \cdot 5 = 5 \text{ бит/буква}. \quad (3)$$

Если буквы алфавита передаются не хаотически, то они передаются не равновероятно (сравним появление букв «е» и «ь») и зависимо, энтропия ансамбля уменьшается. Например, для текстов художественной прозы энтропия составляет 1,5 бит/буква, для ансамбля поэтических произведений — примерно 1 бит/буква (рифма и ритм вносят дополнительные вероятностные связи), для ансамблей текстов телеграмм — менее 0,8 бит/буква (однообразие текстов).

Введем понятие избыточности источника с объемом алфавита k :

$$\chi = \frac{H - H(A)}{H} = \frac{\log k - H(A)}{\log k}, \quad (4)$$

которая определяет, какая доля максимально возможной для данного алфавита энтропии не используется источником. Для рассмотренного примера для поэтических произведений $\chi = (5 - 1)/5 = 0,8$.

1.4. Спектр сигнала.

Гармонические функции обладают фундаментальным свойством: они образуют полный базис в пространстве функций, суммируемых по модулю, и любой физический сигнал $s(t)$ можно разложить в бесконечный ряд гармонических колебаний

$$s(t) = \int_{-\infty}^{+\infty} S(f) \cdot e^{i2\pi f t} \cdot df \quad (1)$$

Функция $S(f)$ в данном выражении называется **спектром** (или спектральной плотностью) сигнала, который связан с $s(t)$ преобразованием Фурье:

$$S(f) = \int_{-\infty}^{+\infty} e^{-i2\pi f t} \cdot s(t) \cdot dt \quad (2)$$

Спектральная плотность вообще говоря является комплексной функцией,

$$\dot{S}(\omega) = S(\omega) e^{i\varphi(\omega)},$$

где $S(\omega)$ — амплитудно-частотная характеристика (АЧХ) сигнала; $\varphi(\omega)$ — его фазо-частотная характеристика (ФЧХ).

Для некоторого физического сигнала $s(t)$, являющегося действительной функцией времени, справедливо следующее соотношение:

$$S(-f) = \int_{-\infty}^{+\infty} e^{i2\pi f t} \cdot s(t) dt = S^*(f)$$

Действительная часть спектра является четной функцией:

$$\operatorname{Re} \cdot S(-f) = \operatorname{Re} \cdot S(f),$$

а мнимая — нечетной (рис. 1.16):

$$\operatorname{Im} \cdot S(-f) = -\operatorname{Im} \cdot S(f)$$

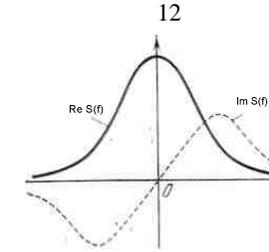


Рис. 1.16. Спектр действительного сигнала.

Таким образом, для восстановления действительного сигнала достаточно знать только половину спектра (область положительных частот)

$$s(f) = 2 \operatorname{Re} \int_0^{\infty} e^{i2\pi f t} \cdot df.$$

Если мы теперь будем считать, что сигнал $s(t)$ состоит из четной $s_+(t)$ и нечетной $s_-(t)$ частей $s(t) = s_+(t) + s_-(t)$, то его спектр можно записать следующим образом:

$$S(f) = 2 \int_{-\infty}^{+\infty} s_+(t) \cdot \cos(2\pi f t) \cdot dt - 2i \int_{-\infty}^{+\infty} s_-(t) \cdot \sin(2\pi f t) \cdot dt$$

Отсюда следует, что спектр является действительным и четным, если сигнал четный $s_-(t) = 0$, и мнимым и нечетным, если сигнал нечетный $s_+(t) = 0$.

1.4.1. Некоторые общие соотношения между сигналом и его спектром

Примеры спектров некоторых распространенных в технике сигналов представлены на рис. 1. Перечислим ряд общих соотношений. Связывающих между собой сигнал $s(t)$ с его спектром $S(f)$.

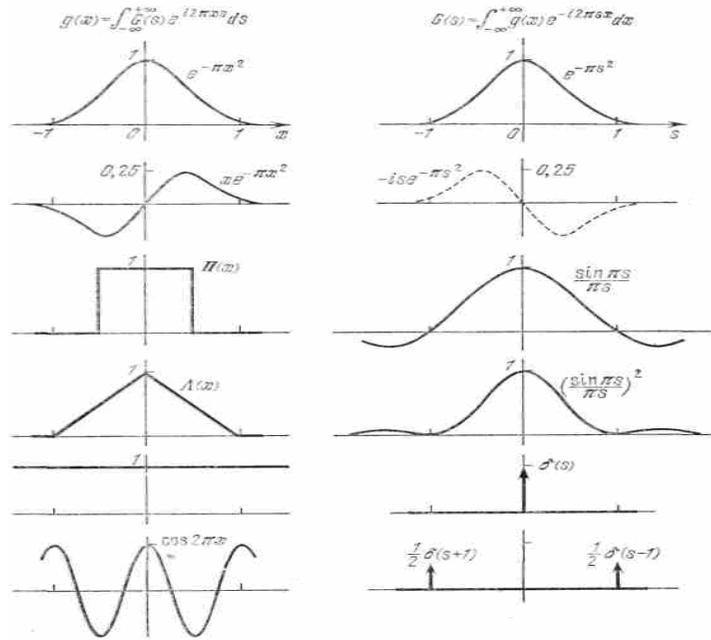


Рис. 1. Спектры некоторых распространенных сигналов.

Преобразование подобия. Сжатие сигнала во времени приводит к соответствующему расширению его спектра и наоборот. Таким образом, если $S(f)$ есть спектр сигнала $s(t)$, то спектр сигнала $s(at)$ имеет вид $\frac{1}{|a|} S\left(\frac{f}{a}\right)$.

Сдвиг во времени. Сдвиг сигнала во времени вызывает фазовый сдвиг спектра, пропорциональный частоте, т.е. спектр сигнала $s(t + \tau)$ имеет вид $e^{i2\pi f \tau} S(f)$

Дифференцирование. Дифференцирование сигнала усиливает в его спектре высокие частоты, ослабляет низкие частоты и полностью подавляет нулевую частоту. Иначе данное свойство можно сформулировать так: если $S(f)$ является спектром сигнала $s(t)$, то спектр производной ds/dt этого сигнала имеет вид $2\pi f S(f)$.

Амплитудная модуляция. Спектр $S(f)$ амплитудно модулированного сигнала $s(t) = e(t) \cdot \cos(2\pi f_0 t)$ получается путем сдвига исходного спектра $E(f)$ огибающей $e(t)$,

уменьшенного по амплитуде в два раза, влево и вправо на величину, равную частоте несущей (рис. 2):

$$S(f) = \int_{-\infty}^{+\infty} e^{-i2\pi f t} \cdot e(t) \cdot \cos(2\pi f_0 t) \cdot dt = \frac{1}{2} E(f - f_0) + \frac{1}{2} E(f + f_0)$$

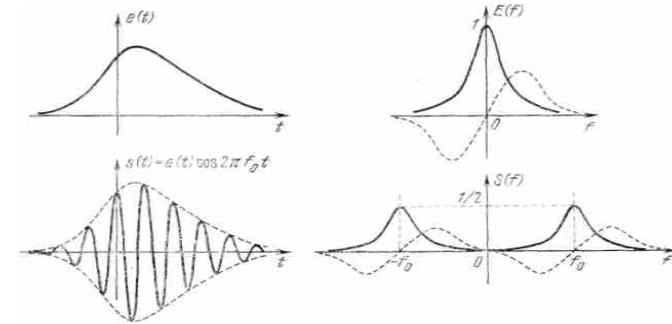


Рис. 2. Спектр синусоидального сигнала, модулированного по амплитуде.

1.4.2. Частотный отклик линейной стационарной системы.

Рассмотрим отклик $s_2(t)$ линейной стационарной системы на гармонический сигнал $s_1(t) = e^{i\omega t}$ с круговой частотой ω , представляющий собой гармоническую функцию той же частоты:

$$s_2(t) = H(\omega) e^{i\omega t} \tag{1}$$

Последнее соотношение означает, что действие системы сводится к линейному преобразованию, причем каждая гармоническое воздействие $s_1(t) = e^{i\omega t}$ при этом является собственной функцией с собственным значением $H(f)$. Величина $H(f)$ называется **передаточной (частотной) характеристикой** системы.

1. 5. Импульсный отклик.

Импульсный отклик $h(t)$ системы, называемый также **импульсной характеристикой** системы, является откликом на сигнал в виде дельта-функции Дирака $\delta(t)$.

Если входной сигнал представить в виде следующей суперпозиции дельта-функций:

$$s_1(t) = \int_{-\infty}^{+\infty} \delta(t - \tau) \cdot s_1(\tau) \cdot d\tau \tag{1}$$

то, учитывая линейность и стационарность системы, сигнал на выходе будет выражаться через $h(t)$ как:

$$s_2(t) = \int_{-\infty}^{+\infty} h(t-\tau) \cdot s_1(\tau) \cdot d\tau \quad (2)$$

Полученное соотношение между сигналом $s_1(t)$ и его импульсной характеристикой $h(t)$ называется **сверткой**. Операция свертки в сокращенном обозначении имеет вид:

$$s_2(t) = s_1(t) * h(t).$$

Согласно соотношению (1, 2), отклик системы на входной гармонический сигнал, т.е. ее передаточная и импульсная характеристики связаны преобразованием Фурье:

$$H(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} h(t) \cdot dt \quad (3)$$

Другими словами частотная (передаточная) характеристика линейной стационарной системы является спектром ее импульсного отклика. Если обратить соотношение (1.4.2), то можно убедиться, в том, что частотная характеристика системы является обратным преобразованием Фурье от частотной характеристики, т.е.

$$h(t) = \int_{-\infty}^{+\infty} H(f) \cdot e^{-i2\pi f t} df \quad (4)$$

1.5.1. Отклик системы на сигнал произвольного вида.

Как было показано выше имеются два метода нахождения отклика $s_2(t)$ линейной стационарной системы на произвольное воздействие.

1. Временной анализ процессов, определяющий значение $s_2(t)$ через свертку импульсного отклика с входным сигналом:

$$s_2(t) = h(t) * s_1(t). \quad (1)$$

2. Гармонический анализ, основанный на представлении входного сигнала в виде суммы гармонических функций:

$$s_1(t) = \int_{-\infty}^{+\infty} S_1(f) e^{-i2\pi f t} df$$

что в силу линейности системы приводит к выражению:

$$s_2(t) = \int_{-\infty}^{+\infty} H(f) S_1(f) e^{-i2\pi f t} df.$$

Спектр выходного сигнала является произведением спектра входного сигнала на частотную характеристику системы:

$$S_2(f) = H(f) S_1(f) \quad (2)$$

Прохождение сигнала через линейную стационарную систему сопровождается сдвигом фаз $\Phi(\omega)$, равным аргументу частотной характеристики. Так как этот сдвиг не пропорционален частоте, система обладает дисперсией. **Групповое временное запаздывание** (ГВЗ) τ_g спектральной группы составляющих сигнала с частотами вблизи центральной частоты ω_0 определяется как:

$$\tau_g = -\left(\frac{d\Phi}{d\omega}\right)_{\omega_0} \quad (3)$$

Взаимосвязь различных характеристик линейной стационарной системы представлена в табл. 1. Здесь пары функций, связанных преобразованием Фурье, соединены стрелками.

Таблица 1.1. Связь отклика линейной стационарной системы с ее характеристиками

Вход	→	ЛИНЕЙНАЯ СТАЦИОНАРНАЯ СИСТЕМА	→	Выход
Гармонический сигнал $e^{i2\pi f t}$		$H(f)$		Гармонический сигнал $H(f)e^{i2\pi f t}$
Дельта- функция Дирака $\delta(t)$				Импульсный отклик $h(t)$
сигнал $s(t)$		свертка с $h(t)$		сигнал $s_2(t) = s_1(t) * h(t)$
спектр $S(f)$		умножение на $H(f)$		спектр $S_2(f) = H(f)S_1(f)$

1.6. Детерминированные сигналы

Детерминированными называются сигналы, заданные вполне определенной функцией времени, например:

$$S(t) = U_m \sin(\omega_0 t + \varphi_0)$$

где амплитуда U_m , частота ω_0 , фаза φ_0 — известные заданные величины.

Случайный (стохастический) сигнал изменяется во времени случайным образом, т.е. представляет собой случайный процесс. Квазидетерминированный (по сути случайный) процесс представляется реализациями, которые описываются функциями времени определенного вида, содержащими несколько случайных параметров, например $S(t) = U_m \sin[\omega_0 t + \varphi(t)]$. При известных U_m и φ_0 , $\varphi(t)$ — случайная изменяющаяся во времени фаза. Такие сигналы описывают большое число сигналов-"переносчиков", у которых один из параметров (в данном случае фаза) изменяется по закону сообщения (модулируется сообщением).

Периодический сигнал — это сигнал, любое значение которого повторяется через пе-

риод T , т. е. $S(t) = S(t+T)$, $-\infty < t < \infty$ (последнее приводит к тому, что ни один реальный сигнал не может быть периодическим).

Непериодическим называется детерминированный сигнал, не удовлетворяющий этому условию.

Энергия и мощность детерминированного сигнала. Требуется, чтобы информация передавалась с заданным качеством при минимальном значении энергии. Если детерминированный сигнал $S(t)$ представляет собой напряжение или ток, то его *мгновенная мощность*, выделяемая на сопротивлении 1 Ом, $P(t) = S^2(t)$.

Энергия этого сигнала в интервале времени (t_a, t_b)

$$E = \int_{t_a}^{t_b} P(t) dt = \int_{t_a}^{t_b} S^2(t) dt. \quad (1)$$

Средняя мощность этого сигнала в этом же интервале

$$P_{cp} = P = \frac{1}{t_b - t_a} \int_{t_a}^{t_b} P(t) dt = \frac{E}{T}, \quad (2)$$

где $t_b - t_a = T$.

Например, в одном периоде $T_0 = 2\pi/\omega_0$ гармонического сигнала $S(t) = \cos(\omega_0 T)$ содержится энергия $E = \int_0^{T_0} \cos^2 \omega_0 t dt = T_0/2$, а средняя мощность $P_{cp} = E/T_0 = 1/2$.

Корреляционные характеристики детерминированных сигналов. Часто возникает необходимость выяснения степени сходства различных сигналов, например сигнала и его оценки.

Взаимная корреляционная функция — это количественная мера, например в единицах мощности, степени сходства сигналов $S_1(t)$ и $S_2(t)$:

$$R_{S_1, S_2}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} S_1(t) S_2(t - \tau) dt, \quad (3)$$

где τ — временной сдвиг сигналов относительно друг друга.

Взаимная корреляционная функция может принимать положительные значения, когда с ростом одного сигнала другой также увеличивается, отрицательные значения, когда с уменьшением одного сигнала другой увеличивается, и значение, равное нулю. По мере изменения τ характер корреляции может меняться.

Автокорреляционная функция характеризует связь между сигналом и смещенной на интервал τ его копией:

$$R_S(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} S(t) S(t - \tau) dt. \quad (4)$$

По мере увеличения τ корреляционные связи у всех сигналов (кроме периодических) ослабевают, уменьшаются и в пределе при $\tau \rightarrow \infty$ $R_S(\tau)$ стремится к нулю.

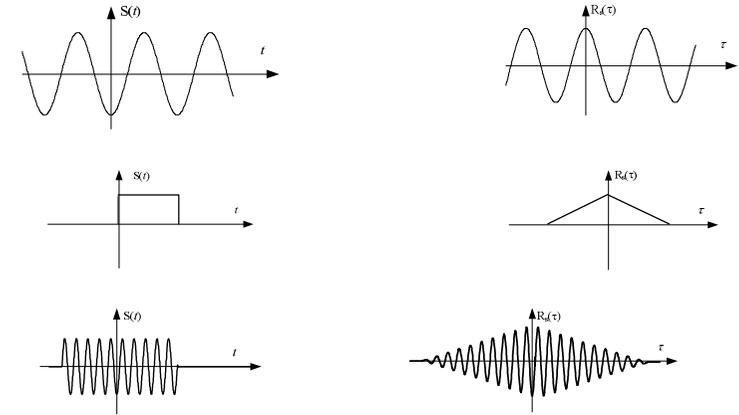


Рис. 1. Автокорреляционные функции некоторых типов детерминированных сигналов.

Корреляционная функция четная, т. е. $R_S(\tau) = R_S(-\tau)$. При $\tau = 0$ автокорреляционная функция достигает максимального значения, равного средней мощности сигнала:

$$R_S(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} S^2(t) dt. \quad (5)$$

По приведенным формулам вычисляют корреляционные функции сигналов с бесконечно большой энергией и ограниченной мощностью, например периодических сигналов. Корреляционная функция сигнала с ограниченной энергией

$$R_S(\tau) = \int_{-\infty}^{\infty} S(t) S(t - \tau) dt. \quad (6)$$

В этом случае максимальное значение $R_S(0)$ корреляционной функции равно энергии сигнала. По этой формуле вычисляется корреляционная функция детерминированного сигнала конечной длительности.

При определении корреляционной функции периодических сигналов усреднение выполняется на интервале, равном периоду. Например, для сигнала

$$S(t) = U_m \cos(\omega_0 t + \varphi_0)$$

корреляционная функция

$$R_s(\tau) = \frac{U_m}{T_0} \int_{-T/2}^{T/2} \cos(\omega_0 t + \varphi_0) \cos[\omega_0(t - \tau) + \varphi] dt = 0,5U_m^2 \cos \omega_0 \tau.$$

1.7. Случайные сигналы

Детерминированное, т. е. заранее известное сообщение, не содержит информации, поэтому источник реальных сообщений следует рассматривать как устройство, которое осуществляет выбор из некоторого множества возможных сообщений. При этом конкретно реализованное сообщение выбирается из множества с определенной вероятностью. Точно так же из множества сигналов и помех выбирается реализация сигнала или помехи.

Ансамбль — это множество, на котором задана вероятностная мера. Ансамбли сообщений и сигналов могут быть конечны (в дискретном случае) или бесконечны.

Ансамбль $X(t)$ — функция времени $x_i(t)$ — является случайным процессом. Каждая входящая в этот ансамбль функция $x_i(t)$ называется *выборочной* функцией или *реализацией процесса*. Именно наличие различных реализаций позволяет сигналу переносить информацию. Чтобы использовать эту возможность, необходимо установить вполне определенное соответствие между каждой реализацией сообщения и каждой реализацией сигнала. Поскольку помехи также являются случайными процессами, то по реализации принятого сигнала можно лишь с некоторой вероятностью судить о том, какое именно сообщение было передано.

В общем случае случайный процесс $X(t)$ считается полностью заданным, если для любого набора моментов времени t_1, t_2, \dots, t_n и любых значений сигнала x_1, x_2, \dots, x_n можно вычислить вероятность того, что в указанные моменты времени сигнал $X(t)$ принимает значения, не превышающие x_1, x_2, \dots, x_n т. е.

$$P\{x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n\} = P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\}, \quad (1)$$

где $P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\}$ — совместная вероятность событий, записанных в скобках. $X(t_k)$ — это случайная величина, так называемое *сечение* случайного процесса в момент времени t_k .

Функция $F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ — n -мерная распределения плотности вероятностей процесса, x_1, x_2, \dots, x_n — ее аргументы, t_1, t_2, \dots, t_n — параметры. Случайный процесс полностью задан, когда для любого n и любых значений t_1, t_2, \dots, t_n можно найти $F(\bullet)$.

Если существуют частные производные функции распределения по всем x_k то можно определить n -мерную *плотность распределения вероятностей* (ПРВ) дифференцированием

ем по всем аргументам:

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (2)$$

Очевидно, зная плотность, можно определить функцию интегрированием:

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2 \dots dx_n \quad (3)$$

В частном случае одномерного распределения ($n = 1$) получаем: $w(x, t) \geq 0$ при всех значениях x ,

$$\int_{-\infty}^{\infty} w(x, t) dx = 1; \quad 0 \leq F(x, t) \leq 1;$$

$$F(\infty, t) = 1; \quad F(-\infty, t) = 0; \quad \int_{-\infty}^x w(\alpha, t) d\alpha = F(x, t).$$

Стационарный процесс — это случайный процесс, у которого определенная группа вероятностных характеристик не изменяется при сдвиге начала отсчета, т. е. при замене аргумента t на $(t + \tau)$, где τ — произвольный интервал времени. Различают стационарность в узком (строгом) и широком смысле. Процесс стационарен в узком смысле, когда n -мерные распределения вероятностей не изменяются во времени (не зависят от выбора момента начала отсчета), т. е. если говорить о плотности, то для любого n и τ справедливо равенство

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = w(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau). \quad (4)$$

Процесс называют стационарным в широком смысле, когда его среднее значение (математическое ожидание) и дисперсия вообще не зависят от времени, а корреляционная функция зависит лишь от $\tau = t_2 - t_1$.

Процесс стационарный в узком смысле стационарен и в широком, но не наоборот.

По виду закона распределения вероятностей случайные процессы делятся на две группы: гауссовские (нормальные), негауссовские. Нормальным называется процесс, у которого любая n -мерная плотность распределения вероятностей

$$w(x_1, x_2, \dots, x_n; t_n) = A_n \exp\left\{-\sum_{i=1}^n \sum_{j=1}^n C_{ij} [(x_i - a_i)(x_j - a_j)]\right\}, \quad (5)$$

где A_n ; C_{ij} ; a_i, a_j — постоянные, определяемые выбором сечений t_1, \dots, t_n и связанные между собой.

В частности, в одномерном случае для нестационарного процесса

$$w(x, t) = \frac{1}{\sqrt{2\pi}\sigma(t)} \exp\left\{-\frac{[x - a(t)]^2}{2\sigma^2}\right\}, \quad (6)$$

для стационарного процесса формула упрощается:

$$w(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{[x-a]^2}{2\sigma^2}\right\}. \quad (7)$$

Случайные сигналы могут быть непрерывными по состоянию и времени; дискретными по состоянию и непрерывными по времени, непрерывными по состоянию и дискретными по времени, дискретными по состоянию и времени.

1.7.1. Характеристики случайных сигналов

Среднее значение случайного процесса по ансамблю (математическое ожидание)

$$\overline{X(t)} \equiv m_x(t) = M[X(t)] = \int_{-\infty}^{\infty} xw(x,t)dx, \quad (1)$$

где $M[X(t)]$ - символ операции определения математического ожидания.

Математическое ожидание — детерминированная функция времени. Разность между случайным процессом и его математическим ожиданием называют *центрированным* процессом:

$$\overset{0}{X} = X(t) - \overline{X(t)}. \quad (2)$$

Значение $m_x(t)$ при каждом значении аргумента t равно математическому ожиданию значения случайной функции $X(t)$ при данном t .

Дисперсия случайного процесса (центральная моментная функция второго порядка) — математическое ожидание квадрата соответствующего центрированного процесса.

Дисперсия вычисляется по формуле

$$\sigma_x^2(t) \equiv [X(t)] = M\left[\overset{0}{X}^2(t)\right] = D[X(t)] = \int_{-\infty}^{\infty} [x - m_x(t)]^2 w(x,t)dx. \quad (3)$$

Дисперсия характеризует степень рассеяния возможных реализаций случайного процесса от его среднего значения и характеризует эту степень в квадратичной форме.

Часто используют такую характеристику как *среднее квадратическое отклонение*

$$\sigma_x = \sqrt{D_x}. \quad (4)$$

Средний квадрат — это моментная функция второго порядка, математическое ожидание квадрата самого (не *центрированного*) случайного процесса:

$$\eta_x^2(t) = \overline{X^2(t)} = \left[\overline{X(t) + m_x(t)}\right]^2 = \left[\overset{0}{X}(t)\right]^2 + 2\overline{X(t)m_x(t)} + \overline{m_x^2(t)} = \sigma_x^2(t) + m_x^2(t), \quad (5)$$

поскольку $\overset{0}{X}(t) = 0$, $\overline{m_x(t)} = m_x(t)$.

У стационарных случайных процессов, как было отмечено ранее,

$$m_x(t) = m_x = const, \quad \sigma_x(t) = \sigma_x = const, \quad \eta_x(t) = \eta_x = const.$$

Если стационарный эргодический процесс $X(t)$ является напряжением или током в нагрузке сопротивлением 1 Ом, то $\eta_x^2 = \overline{X^2(t)}$ является полной средней мощностью процесса,

$m_x^2 = [\overline{X(t)}]^2$ — мощностью постоянной составляющей, $\sigma_x^2 = \left[\overset{0}{X}(t)\right]^2$ — мощностью пере-

менной составляющей.

Важно правило "Трех сигм": с той или иной вероятностью p можно утверждать, что значения непрерывного случайного процесса $X(t)$ не выходят за пределы $m_x \pm 3\sigma_x$ (для нормального процесса $p = 0,997$). Согласно неравенству Чебышева, вероятность p не может быть меньше 8/9.

Кроме средних значений по ансамблю, можно определить среднее значение случайного процесса по времени. Для стационарного финитного (заданного на интервале от t_1 до t_2) процесса можно определить постоянную составляющую i -й реализации

$$\tilde{x}_i(t) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x_i(t)dt. \quad (6)$$

Для процесса, заданного на всей оси времени,

$$\tilde{x}_i(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)dt. \quad (7)$$

Стационарный процесс называют *эргодическим*, если для него усреднение по времени по одной достаточно продолжительной реализации приводит к тому же результату, что и статистическое усреднение, т. е. постоянная составляющая равна математическому ожиданию, мощность переменной составляющей равна дисперсии и т. п.

Стационарность — необходимое, но не достаточное условие эргодичности. Эргодичность требует того, чтобы результаты усреднения по времени различных реализаций совпадали друг с другом и с результатом усреднения по множеству реализаций.

Эргодичность существенно упрощает определение вероятностных характеристик (распределений вероятностей, числовых характеристик), поскольку они могут быть определены по одной реализации (не надо производить усреднение по множеству), например определение одномерной интегральной функции $F(t)$ сводится к измерению относительного времени пребыва-

ния реализации $x(t)$ ниже заданного уровня анализа.

1.7.2. Корреляционные характеристики случайных сигналов

Ковариационная функция (иногда ее называют корреляционной) — это характеристика нецентрированного случайного процесса. Она представляет собой математическое ожидание произведения двух его сечений:

$$B_x(t_1, t_2) = \overline{X(t_1)X(t_2)} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} x_1 x_2 w(x_1, x_2; t_1, t_2) dx_1 dx_2, \quad (1)$$

где $w(x_1, x_2; t_1, t_2)$ — двумерная (совместная) плотность распределения вероятностей процесса в этих сечениях.

Точно так же, но уже для центрированного процесса $\overset{0}{X}(t)$ определяется корреляционная функция случайного процесса $X(t)$:

$$R_x(t_1, t_2) = \overline{X(t_1)X(t_2)} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} [x_1 - \overline{X(t_1)}][x_2 - \overline{X(t_2)}] w(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (2)$$

Корреляционная функция (ее иногда называют автокорреляционной) характеризует связь между значениями одного и того же случайного процесса в два различных моменты времени.

Взаимная корреляционная функция случайных процессов $X(t), Y(t)$:

$$R_{XY}(t_1, t_2) = \overline{X(t_1)Y(t_2)} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} [x_1 - \overline{X(t_1)}][y_2 - \overline{Y(t_2)}] w(x_1, y_1; t_1, t_2) dx_1 dy_2 \quad (3)$$

где $w(x)$ — двумерная плотность распределения вероятностей сечений процесса $X(t)$ в момент времени t_1 и процесса $Y(t)$ в момент времени t_2 .

Для стационарного случайного процесса корреляционная функция $R_x(t_1, t_2)$ является лишь функцией разности $\tau = t_2 - t_1$ т. е. $R_x(t_1, t_2) = R_x(\tau)$ не зависит от размещения начала отсчета.

Корреляционную функцию не всегда удобно использовать при расчетах и инженерных оценках. Поэтому часто используют величину, определяющую корреляционные связи между сечениями случайного процесса, — **интервал корреляции**.

Интервал корреляции дает ориентировочное представление о том, на каких интервалах времени в среднем имеют место корреляции между значениями случайного процесса.

Существует множество способов определения интервала корреляции:

1) $\tau_k = \tau_1$, если при $\tau > \tau_1$ коэффициент корреляции $|r_x(\tau)|$ пренебрежимо мал, напри-

мер, 0,1, т. е. $|r_x(\tau_k)| = |r_x(\tau_1)| = 0,1$ (рис. 1);

2) $\tau_k = \tau_2$, где τ_2 — это основание прямоугольника с высотой равной 1, равновеликого фигуре, ограниченной модулем нормированной корреляционной функции и осью τ (рис. 1 б):

$$\tau_2 = \int_0^{\infty} |r_x(\tau)| d\tau \quad (4)$$

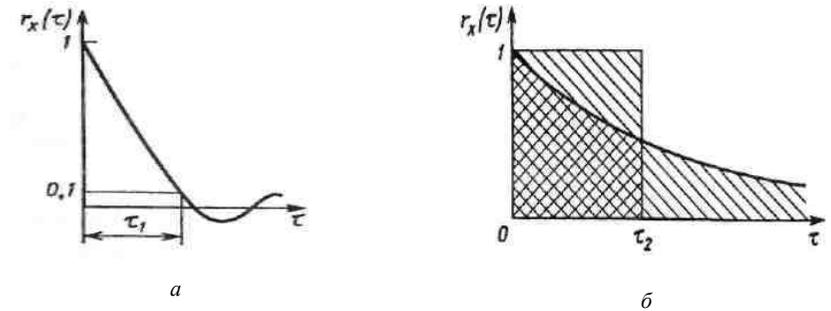


Рис.1. Определение интервала корреляции.

1.7.3. Спектральные характеристики случайных сигналов

Спектральная плотность мощности (**энергетический спектр**) $S_x(\omega)$ выражает зависимость от частоты средней мощности стационарного случайного сигнала $X(t)$, приходящейся на единицу полосы частот, и имеет размерность энергии (Вт/Гц).

Энергетический спектр стационарного случайного процесса связан с его корреляционной функцией парой преобразований Фурье:

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau, \quad R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} d\omega. \quad (1)$$

Приведенные соотношения представляют **теорему Винера—Хинчина**.

Для сравнения различных случайных процессов по ширине полосы их спектра используют способ, использованный выше при оценке интервала корреляции. Ширина полосы спектра сигнала может быть определена разными способами:

1) ширина спектра на каком-то уровне, часто на уровне 0,5 (рис. 1.9а);

2) эффективная шумовая ширина спектра. По правилу эквивалентного прямоугольника (рис. 1.9б):

$$\Delta f = \int_0^{\infty} g_x(f) df = \frac{1}{2\tau_3}. \quad (2)$$

Параметр Δf использован выше в аппроксимациях корреляционных функций и энерге-

тических спектров.

Случайный процесс называют *белым (белый шум)*, если его спектр равномерный, т. е. имеет постоянное значение $S_x(\omega) = const$ при всех значениях частот. В этом случае используется аналогия с белым светом, т. е. электромагнитными волнами, имеющими равномерный спектр в области видимых частот. Согласно равенству $r_x(\tau) = \delta(\tau)$, любые сколь угодно близкие сечения белого шума оказываются некоррелированными. Белый шум — это не реальный физический процесс, а математическая идеализация потому, что его мощность

Встречаются на практике процессы часто имеют равномерную спектральную плотность мощности в весьма широкой полосе частот (по крайней мере более широкой, чем полосы пропускания устройств или цепей, на которые они воздействуют). Типичный пример — *тепловой шум*, спектральная плотность которого равномерна на частотах

$$f < \frac{kT}{h}, \tag{3}$$

где k — постоянная Больцмана; h — постоянная Планка; T — температура, К.

При $T = 300$ К имеем $f < 6 \cdot 10^{12}$ Гц.

Замена такого шума идеализированным белым не приводит к ощутимым погрешностям.

1.8. Представление непрерывных сигналов отсчетами.

При дискретизации по времени в качестве координат сигнала используют текущие его значения, так называемые *отсчеты* (выборки) $S(t_1), S(t_2), \dots, S(t_n)$ в моменты времени (моменты дискретизации, точки опроса) t_1, t_2, \dots, t_n . Таким образом, при представлении отсчетами координаты $S_k = S(t_k)$ (рис. 1).

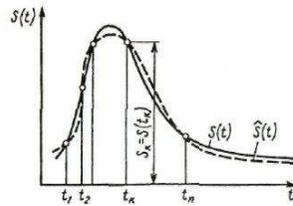


Рис.1. Представление сигнала отсчетами

По этим отсчетам в результате интерполяционной обработки осуществляется восстановление сигнала. Мгновенные отсчеты могут быть представлены в виде дельта-функции $\delta(t)$. Наиболее часто осуществляется периодическая дискретизация, когда при определенном

k :

$$t_{k+1} - t_k = T_0 = const. \tag{1}$$

Периодический опрос просто реализуется электронными ключевыми схемами. При периодической дискретизации для последовательных отсчетов можно записать:

$$S^n(t) = \sum_{k=-\infty}^{\infty} S(t) \Big|_{t=kT_0} = \sum_{k=-\infty}^{\infty} S(kT_0) \delta(t - kT_0) = S(t) \sum_{k=-\infty}^{\infty} \delta(t - kT_0). \tag{2}$$

В общем случае процесс интерполяции непрерывного сигнала $S(t)$ по его отсчетам $S(t) = S(t_k)$ может быть описан следующим образом:

$$\hat{S}(t) = \sum_{k=0}^{N_k} \hat{S}(t_k) W_k(t - t_k) = \sum_{k=0}^{N_k} \hat{S}(kT_0) W_H(t - kT_0), \tag{3}$$

где $\hat{S}(t)$ — оценка восстанавливаемого сигнала; N_{k+1} — число отсчетов сигнала, участвующих в интерполяции; $\hat{S}(t_k)$ — отсчет сигнала в k -й момент времени (с учетом его зашумленности в процессе передачи по каналу связи); $W_H(\bullet)$ — элементарная интерполяционная функция, импульсная характеристика интерполятора.

Сигнал на выходе интерполятора определяется как линейная сумма его реакций на отсчеты, т. е. для интерполятора справедливо свойство суперпозиции (рис. 2).

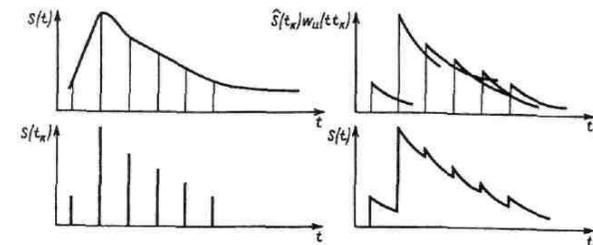


Рис.2. Представление сигнала отсчетами

Теорема Котельникова. Теорема Котельникова была доказана в 1933 г. в виде теоремы о функциях с финитным спектром, за рубежом — в виде теоремы отсчетов (приписывают Клоду Шеннону); отчасти предвосхищена Найквистом в 1924 г. В качестве интерполяции между равно отстоящими точками она была рассмотрена в 1915 г. Уиттикером. Теорема Котельникова — основа теории дискретного представления отсчетами.

Существо теоремы следующее. Функцию $S(t)$ с финитным (ограниченным) спектром можно точно восстановить (интерполировать) по ее отсчетам $S(kT_0)$, взятым через интервалы $T_0 \leq 1/2F$, где F — верхняя частота спектра функции. Это осуществляется с помощью ряда Котельникова

$$\hat{S}(t) = S(t) = \sum_{k=-\infty}^{\infty} S(kT_0) W_H(t - kT_0), \tag{4}$$

где в качестве интерполирующих функций $W(\bullet)$ используются функции отсчетов $W_{II}(t - kT_0) = \text{sin c}[2\pi F(t - kT_0)]$

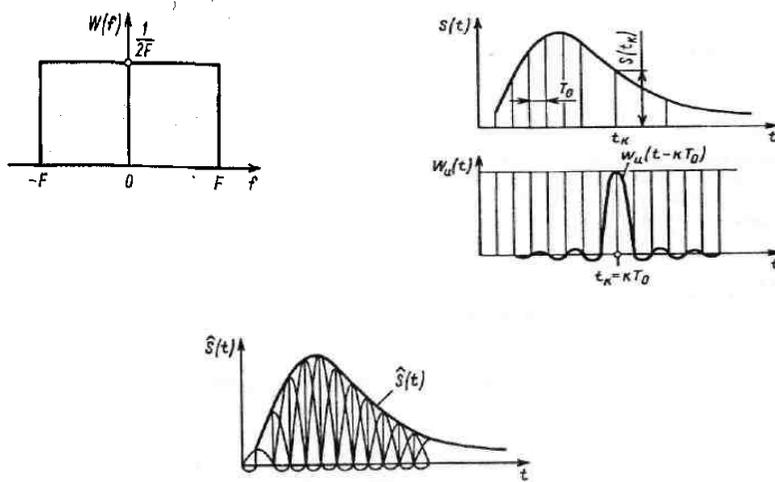


Рис. 3. К теореме Котельникова

Эти функции представляют собой весовую или импульсную характеристику идеального ФНЧ. Передаточная функция идеального ФНЧ (рис. 2.20)

$$W(f) = \begin{cases} 1/2F, & -F < f < F; \\ 0, & |f| > F. \end{cases} \quad (5)$$

Из теоремы Котельникова, в частности следует, что если сигнал $S(t)$ имеет конечную длительность T и ширину спектра F , то для его представления потребуется $M = 2FT$ независимых отсчетов.

Условие финитности $S(f) = 0$ при $|f| > F$ является весьма существенным для правильного восстановления сигнала по его отсчетам. Если оно не соблюдается, то возникают искажения сигнала, называемые искажением перехлестывания (рис. 4). Поскольку в ситуации, иллюстрируемой рис. 4, не выполняется условие теоремы Котельникова и $F > 1/2T_0$, где F — верхняя частота спектра, T_0 — период дискретизации, то точное восстановление сигнала невозможно (мощность шумов перехлестывания определяется площадями заштрихованных фигур на рис. 4).

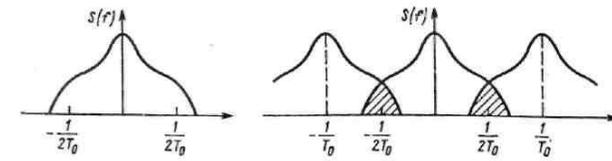


Рис.4. Спектр дискретизированных сигналов

2. СПОСОБЫ МОДУЛЯЦИИ СИГНАЛОВ

Существует три основных метода модуляции сигналов: амплитудная (АМ), частотная (ЧМ) и фазовая (ФМ), — каждый из которых имеет большое число вариантов. Все рассматриваемые методы модуляции можно представить ВЧ несущими, модулированными сигналами, передающими цифровую информацию. Для получения желаемых характеристик модулированного сигнала (малой ширины спектра, заданного распределения спектральных составляющих в заданной полосе и т. п.) перед подачей на модулятор с помощью специальных фильтров модулирующему сигналу (импульсам) придается определенная форма, они иногда перекодируются, преобразуются в многопозиционные сигналы и т. п.

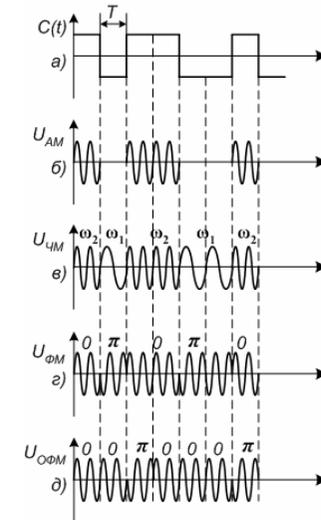


Рис 1. Временные диаграммы сигналов: исходный (а), АМ (б), ЧМ (в), ФМ (г), ОФМ (д)

В приемнике в результате детектирования, фильтрации и регенерации стремятся восстановить исходную информацию. Для когерентного детектирования требуется синусоидальный опорный сигнал, совпадающий с принимаемой несущей как по частоте, так и по фазе. Такой опорный сигнал можно получить либо с помощью специально передаваемого вспомога-

тельного сигнала, либо в результате обработки модулированного сигнала. Для некогерентного детектирования подстройка опорного сигнала под входящий с точностью до фазы не требуется. На выходе детектора устанавливается регенератор, имеющий решающее устройство (РУ) и устройство выделения тактовой частоты (ВТЧ) из принимаемого сигнала.

2.1. Амплитудная модуляция (АМ)

При АМ дискретный сигнал можно записать в виде

$$U_{AM}(t) = U_c [1 + c(t) \cdot \cos \omega_c t] \quad (1)$$

где $c(t)$ – модулирующий сигнал; U_c – амплитуда сигнала; ω_c – частота несущей. При $c(t) \in \{+1, -1\}$ получаем простейший вид цифровой АМ – бинарную АМ (рис. 1а, б). Демодуляция АМ сигналов осуществляется амплитудным детектором, и решающее устройство принимает решение о приеме символа по превышению входным сигналом порогового уровня, который обычно выбирается вблизи уровня $U_c/2$. Разновидностью дискретной АМ является квадратурная амплитудная модуляция (КАМ), при которой производится модуляция двух сигналов, сдвинутых по фазе на 90° с последующим их сложением:

$$U_{КАМ}(t) = U_c [c_1(t) \cos \omega_c t + c_2(t) \sin \omega_c t], \quad (2)$$

где $c_1(t)$; $c_2(t)$ — модулирующие сигналы в квадратурных каналах. При приеме сигналов с КАМ производится когерентное детектирование.

2.2. Частотная модуляция (ЧМ)

При модуляции ВЧ колебания по частоте оно может быть описано выражением

$$U_{ЧМ}(t) = U_c \cos \left[\omega_c t + \Delta \omega_d \int_0^t c(t) dt \right], \quad (1)$$

где $\Delta \omega_d$ — девиация частоты. В самом простейшем случае, когда $c(t) \in \{+1, -1\}$, модуляция осуществляется частотным сдвигом с использованием двух частот $\omega_2 = \omega_c + \Delta \omega$, соответствующей $c(t)=+1$, и $\omega_1 = \omega_c - \Delta \omega$, соответствующей $c(t)=-1$ (рис. 4.35, в), разнесенных на $\Delta f_p = f_2 - f_1$, девиация частоты $\Delta f = \Delta f_p / 2$, а индекс модуляции

$$m = \Delta f_p T, \quad (2)$$

где $T = 1/B$ — длительность символа; B — скорость передачи.

При приеме ЧМ сигналов чаще всего применяется некогерентное детектирование, которое осуществляется либо по огибающей, либо по мгновенной частоте. Для приема по огибающей используется фильтровая схема частотного детектора (ЧД) (рис. 1, а). Разделение сигналов с частотами f_1 и f_2 производится с помощью полосовых фильтров, настроенных на эти

частоты, после которых стоят амплитудные детекторы АД (детекторы огибающей) и решающее устройство.

Прием по мгновенной частоте осуществляется при помощи линейного частотного дискриминатора (рис. 1, б), в котором производится преобразование изменений частоты в изменения амплитуды. После детектирования и фильтрации посылки постоянного тока поступают на решающее устройство.

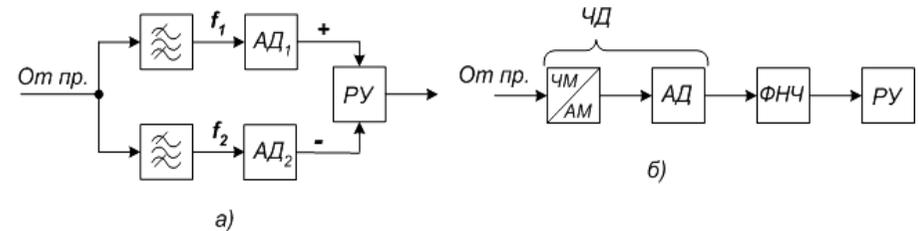


Рис. 1. Структурные схемы детекторов ЧМ сигналов

При точной настройке приемника, когда средняя частота ЧД f_0 равна

$$f_0 = (f_1 + f_2) / 2, \quad (3)$$

напряжение на его выходе представляет собой двухполярные посылки постоянного тока. При наличии нестабильности несущей ЧМ сигнала относительно частоты f_0 нарушается равенство (3) и возникают амплитудные преобладания посылок одного знака, что снижает помехоустойчивость приема.

Усовершенствованным видом частотной модуляции является ЧМ с непрерывной фазой (ЧМНФ), при которой в моменты переходов двоичных символов не происходит скачков фазы, что позволяет сократить полосу частот, занимаемую ЧМ сигналом. Плавный переход фазы реализуется на частотном модуляторе аналоговых РРЛ с линейной модуляционной характеристикой, на который поступает дискретный сигнал (рис. 4.35, а), пропущенный через фильтр и имеющий плавное изменение амплитуды (например, приподнятый косинусообразный импульс). При этом в результате модуляции частота изменяется по закону приподнятого косинуса, а фаза ЧМ сигнала будет непрерывной. Детектирование сигналов ЧМНФ может производиться ЧД на основе дискриминатора (рис 1, б).

2.3. Фазовая модуляция (ФМ)

В настоящее время разработано несколько вариантов двухпозиционной (бинарной) и многопозиционной ФМ. Наиболее простой является бинарная ФМ, когда фаза несущей сдвигается на 0 или 180° при изменении полярности двоичных символов (рис. 1):

$$U_{\text{ФМ}}(t) = U_c \cos \left\{ \omega_c t + \frac{\pi}{2} [1 - c(t)] \right\}. \quad (1)$$

Точный эталон фазы, требуемый для когерентного детектирования в фазовом детекторе (ФД), получают путем нелинейного преобразования входного фазомодулированного сигнала (рис. 1а). Удвоение частоты сигнала, модулированного по фазе на 180°, приводит в этой схеме к устранению модуляции. Это позволяет включить после умножителя узкополосный фильтр для ослабления помех. Затем путем деления частоты на два получаем исходное колебание без модуляции фазы и с меньшим уровнем помех.

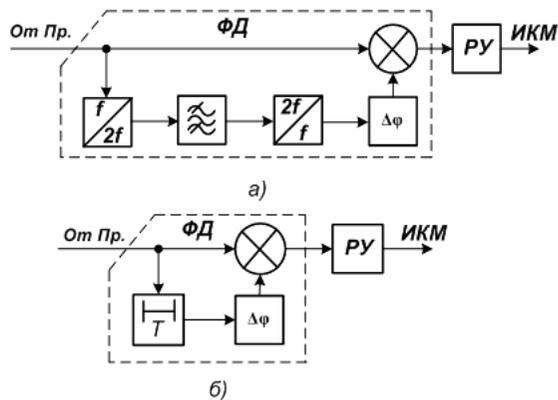


Рис. 1. Структурные схемы детекторов ФМ сигналов

Все возрастающие требования к сокращению полосы частот, занимаемой высокоскоростными цифровыми сигналами, привели к использованию многократных способов амплитудно-фазовой модуляции (АФМ), которая предполагает амплитудную модуляцию синфазной и квадратурной составляющих. Если в (1) модулирующие сигналы $c_1(t)$, $c_n(t)$ принимают значения ± 1 , то получаем квадратурную АМ. Если же для модуляции как в синфазном, так и в квадратурном каналах используются четырехуровневые сигналы $c(t) - +1; \pm 3$, то при этом получается 16-кратная КАФМ, которую можно записать

$$U_{\text{КАМФ-16}}(t) = \sum_{i=1}^{16} A_i \cos(\omega_c t + \theta_i) \quad (2)$$

и представить в фазово-амплитудном пространстве в виде рис.2, где точками показаны положения концов вектора сигнала A_j при различных значениях i .

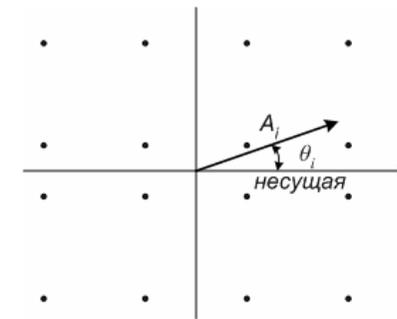


Рис. 2. Расположение сигнальных точек 16-кратной АФМ в фазово-амплитудном пространстве

3. КЛАССИФИКАЦИЯ ЭЛЕКТРИЧЕСКИХ ЦЕПЕЙ

3.1. Линейные и нелинейные цепи

При рассмотрении электрических цепей удобно выделить те характерные их признаки, на основании которых целесообразно все цепи разбить на определенные группы.

На рис. 1 представлена обобщенная электрическая цепь. На ее вход действует сигнал $x(t)$, который обычно называют *входным воздействием*. На выходе цепи формируется сигнал $y(t)$, являющийся *реакцией цепи* на данное входное воздействие.

Реакция цепи на входное воздействие $x(t)$ зависит как от вида $x(t)$, так и от свойств системы. На этом основании выходной сигнал $y(t)$ можно записать в форме причинно-следственной связи:

$$y(t) = f(x(t), a, b, c, \dots) \quad (1)$$

Здесь символами $a, b, c \dots$ обозначены параметры цепи.

Функциональная зависимость (1) определяется свойствами системы. Эти свойства у различных систем различны и являются удобными признаками для классификации цепей и систем. Первым признаком, по которому следует делить электрические цепи, является вопрос о соответствии цепи принципу суперпозиции. По этому признаку все цепи целесообразно делить на линейные и нелинейные.

Линейными называются цепи, для которых справедливы равенства:

$$\begin{aligned} f(x_1 + x_2) &= f(x_1) + f(x_2), \\ f(ax) &= af(x). \end{aligned} \quad (2)$$

Первое равенство выражает свойство аддитивности, второе - однородности. Свойство аддитивности отражает *принцип суперпозиции*, который формулируется так: реакция цепи на сумму воздействий равна сумме реакций на каждое воздействие в отдельности. Если пре-

образователь не обладает свойствами аддитивности и однородности, он относится к группе нелинейных.

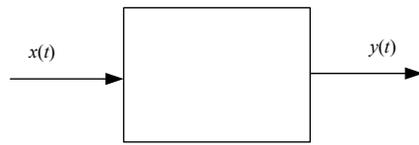


Рис.1. Обобщенная электрическая цепь

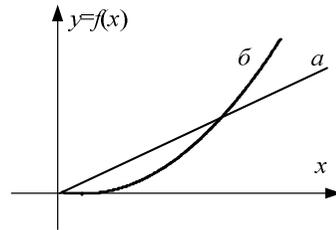


Рис.2. Графики зависимости $y=f(x)$
а) линейной и б) нелинейной цепей

Свойства аддитивности и однородности, определяющие линейность системы, позволяют сразу решать вопросы о пригодности тех или иных математических методов исследования. Например, методы изучения переходных процессов в линейных цепях (операционное исчисление, преобразование Фурье) основаны на использовании принципа суперпозиции, и для изучения нелинейных систем эти методы применять нельзя.

Вторым признаком, по которому удобно классифицировать цепи и системы, является вопрос о возникновении в результате преобразования сигнала цепью новых спектральных составляющих. В линейных системах с постоянными параметрами новых частот, т. е. частот, не содержащихся во входном сигнале, не возникает. В нелинейных и параметрических цепях имеет место преобразование спектра сигнала.

По указанным выше признакам все цепи целесообразно делить на группы:

1. **Линейные цепи.** Для таких цепей параметры a, b, c в формуле (1) являются постоянными величинами, не зависящими от входного сигнала $x(t)$. В таких цепях обогащение спектра сигнала $x(t)$ новыми спектральными составляющими невозможно.

2. **Нелинейные цепи.** Параметры $a, b, c \dots$ (один или несколько из них) зависят от электрического режима цепи, т. е. от тока или напряжения. Следствием этого является нелинейность функциональной зависимости (1), которая может иметь, например, вид

$$y(t) = f[x(t); a(x), b(x), c(x), \dots] \quad (3)$$

3. **Параметрические цепи.** Параметры a, b, c, \dots не зависят от электрического режима, но зависят (все или один из них) от текущего времени t :

$$y(t) = f[x(t); a(t), b(t), c(t), \dots] \quad (4)$$

Строго говоря, все цепи следует считать нелинейными. В условиях, однако, когда отклонения функциональной зависимости (1) от линейного закона относительно малы и параметры меняются во времени достаточно медленно, цепь следует рассматривать как линейную. Если при линейной функциональной зависимости (1) параметры заметно изменяются во времени, цепь нужно считать параметрической. Если же при относительно постоянных параметрах $a, b, c \dots$ зависимость (1) нелинейная, систему следует считать нелинейной.

3.2. Характеристики пассивных элементов электрических цепей

Как известно, пассивные электрические цепи состояются из элементов трех типов: резисторов, конденсаторов и катушек индуктивности. В зависимости от способности элементов рассеивать или запасать энергию их можно делить на диссипативные и консервативные. **Диссипативные элементы** не обладают способностью сохранять энергию. Энергия на них рассеивается, т. е. необратимо превращается в тепло. К этой группе относятся резисторы, которые часто называют активными сопротивлениями. **Консервативные элементы** сохраняют (консервируют) энергию. К ним относятся конденсаторы, обладающие способностью запасать электрическую энергию, и катушки индуктивности, запасующие магнитную энергию. Элементы этой группы часто называют реактивными, так как в цепях, находящихся под воздействием гармонических колебаний, они имеют реактивное сопротивление, не связанное с расходом мощности.

Все элементы электрической цепи характеризуются своими параметрами: сопротивлением R , емкостью C или индуктивностью L . Как указывалось выше, параметры элементов могут быть постоянными, и тогда элементы определяются как *линейные*, зависеть от электрического режима - *нелинейные* и меняться во времени - *параметрические*. Для каждого из этих элементов характерна своя связь между электрическими параметрами. Эти связи считаются известными из опыта как выражающие свойства элементов. Например, для конденсатора связь между током i и зарядом q имеет вид: $i = dq/dt$.

3.3. Нелинейные искажения сигналов в электрических цепях

При работе тракта передачи информации всегда имеются известные отличия передаваемого сигнала $S(t)$ от сигнала на его выходе $\tilde{S}(t)$, т. е.

$$\tilde{S}(t) = S(t) + \varepsilon(t) \quad (1)$$

где $\varepsilon(t)$ - сигнал ошибки.

Изменения передаваемого сигнала (1.2) называются искажениями сигнала, а сигнал $\tilde{S}(t)$ на выходе тракта - искаженным. Величина $\varepsilon(t)$ или связанные с ней параметры обычно нормиру-

ются. В зависимости от природы цепей, образующих тракт или отдельные его устройства, различают линейные и нелинейные искажения.

Нелинейные искажения (НИ) возникают в нелинейных и параметрических цепях. Эти цепи обладают общим свойством порождать новые колебания в спектрах проходящих через них сигналов и помех, что существенно усложняет вид сигнала $\varepsilon(t)$ и методику его оценки.

Одной из оценок нелинейных искажений сигнала в тракте является *коэффициент гармоник* K_z , определяемый как отношение среднеквадратической суммы заданных компонентов выходного сигнала, отсутствующих в спектре входного гармонического сигнала к среднеквадратической сумме заданных спектральных компонентов выходного сигнала.

$$K_z = \frac{\sqrt{A_2^2 + A_3^2 + \dots}}{\sqrt{A_1^2 + A_2^2 + A_3^2 + \dots}} * 100\% \quad (2)$$

Иногда используют коэффициенты гармоник по отдельной гармонике

$$K_{zn} = \frac{A_n}{A_1} * 100\% \quad (3)$$

4. ГЕНЕРАЦИЯ ЭЛЕКТРИЧЕСКИХ СИГНАЛОВ

4.1. Принципы построения схем генераторов электрических сигналов

Электронные схемы могут генерировать переменные сигналы различной формы и в широком диапазоне частот. Для этого активные электронные компоненты (транзисторы, лампы) с помощью которых создаются различные типы генераторов, служащие для получения электромагнитных колебаний.

Генератор представляет собой устройство, преобразующее мощность источника постоянного напряжения в мощность генерируемого сигнала с заданными характеристиками. Помимо частоты и формы генерируемых колебаний имеют значение также эффективность указанного преобразования и стабильность частоты генератора.

Генерация колебаний чаще всего обеспечивается за счет применения положительной обратной связи в усилителях.

Обратной связью принято называть связь между входными и выходными цепями усиленной системы. За счет обратной связи часть энергии усиливаемых колебаний поступает обратно с выхода на вход. При этом пришедшие обратно колебания складываются с начальными.

Напряжение обратной связи, подаваемое на вход усилителя цепью обратной связи, может быть в фазе со входным сигналом или в противофазе. В первом случае обратная связь называется *положительной*, во втором — *отрицательной*. В усилителях положительная обратная связь, как правило, приводит к возникновению генерации. Установим общие соотношения, ха-

рактерные для усилителей с обратной связью.

Между входными зажимами усилителя действует сумма двух напряжений (рис. 1)

$$U_g = U_{вх} + U_{oc} \quad (1)$$

где $U_{вх}$ - напряжение, развиваемое входным источником, U_{oc} - напряжение обратной связи. Напряжение обратной связи представляет некоторую β -ю часть выходного напряжения $U_{вых}$, т. е.

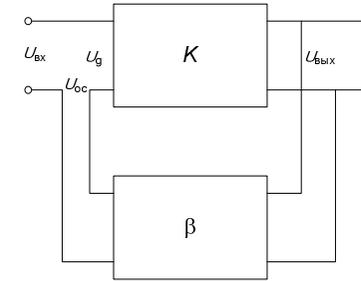


Рис.1. Усилитель с обратной связью по напряжению.

$$U_{oc} = \beta U_{вых}, \quad (2)$$

При положительной обратной связи U_{oc} находится в фазе с напряжением источника $U_{вх}$, поэтому здесь выполняется соотношение:

$$U_g > U_{вх} \quad (3)$$

При отрицательной обратной связи U_{oc} и $U_{вх}$ противофазны, поэтому,

$$U_g < U_{вх} \quad (4)$$

Уменьшение напряжения действующего на входе усилителя и является причиной снижения коэффициента усиления при отрицательной обратной связи. По принципу действия обратную связь делят на два вида: 1) обратная связь по напряжению - когда эффект влияния обратной связи зависит от величины выходного напряжения; 2) обратная связь по току - при которой эффект влияния обратной связи зависит от величины выходного тока. При этом напряжение обратной связи пропорционально выходному току через нагрузку.

Рассмотрим как изменится коэффициент усиления $K_{сб}$ охваченного петлей обратной связи. По определению:

$$K_{сб} = U_{вых} / U_{вх}. \quad (5)$$

С другой стороны,

$$U_{вых} = K U_g = K(U_{вх} + \beta U_{вых}). \quad (6)$$

Из полученных формул получаем,

$$K_{сб} = K / (1 - K\beta). \quad (7)$$

Как видим, под действием обратной связи коэффициент усиления изменился в $(1 - K\beta)$ раз. Найденный множитель обозначают как γ и называют глубиной обратной связи, $\gamma = 1 - K\beta$.

Отсюда следует, что в условиях положительной обратной связи, если $1 - K\beta = 0$, коэффи-

коэффициент усиления $K_{св}$ стремится к бесконечности. Физически это означает самовозбуждение усилителя, т.е. превращение его в автогенератор. Учитывая комплексный характер K и β найденное выше условие самовозбуждения может быть представлено в виде двух скалярных соотношений:

$$|K\beta| = 1, \quad (8)$$

$$\varphi_k + \varphi_\beta = 2\pi n, \quad (9)$$

где n – целое число.

Первое из данных соотношений представляет собой *условие баланса амплитуд*, а второе – *баланса фаз*.

Условие баланса фаз означает, что сигнал с выхода усилителя должен поступить на его вход в фазе с входным сигналом. Другими словами сдвиг фаз φ_k , создаваемых усилителем и сдвиг фаз φ_β , формируемый в цепи обратной связи был кратен 2π .

Условие баланса амплитуд означает, что усилитель должен полностью компенсировать потери сигнала, приобретаемые им в петле обратной связи.

При выполнении указанных условий самовозбуждения шумовое напряжение, которое всегда присутствует на входе усилителя, достаточно для перехода генератора в автоколебательный режим.

Ниже рассмотрим некоторые конкретные типы электронных и оптических автогенераторов.

4.2. Генераторы гармонических колебаний

Прежде всего рассмотрим генератор, в котором условие баланса фаз удовлетворяется за счет использования фазовращающей RC-цепочкой в цепи обратной связи усилительного

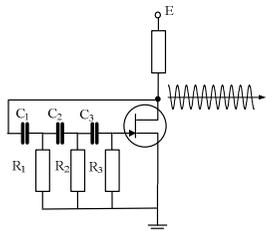


Рис.1. Гармонический RC-генератор

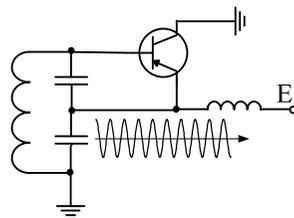


Рис.2. Гармонический LC-генератор

каскада (см. рис. 1 и 2). В полосе рабочих частот каскад усилителя вносит фазовый сдвиг равный π радиан. Дополнительный фазовый сдвиг равный тоже π радиан вносит и цепочка из трех RC-звеньев, так, чтобы удовлетворить условию баланса фаз $\varphi_k + \varphi_\beta = 2\pi n$. Каждое звено

RC-фильтра на рис.2, таким образом, сдвигает фазу на $\pi/3$ радиан. Важно отметить, что при этом указанные условия удовлетворяются только на одной частоте. Именно эта частота и будет формироваться генератором. Генераторы работающие на одной частоте называются *гармоническими генераторами*.

Упрощенная электрическая схема гармонического генератора другого типа представлена на рис.2. Здесь требуемый фазовый сдвиг в π радиан обеспечивается параллельным колебательным контуром, определяющим частоту автоколебаний.

4.3. Стабилизация частоты генераторов

Частота любого генератора в процессе его работы не остается постоянной, а меняется по случайному закону под действием различных причин, что приводит к нестабильности генерируемой частоты. Для оценки нестабильности вводят понятия абсолютной и относительной нестабильности частоты.

Абсолютная нестабильность есть разность Δf между текущим значением частоты f и средним номинальным значением частоты f_0 .

Относительная нестабильность есть отношение $\Delta f / f$. Из всех известных физических величин наименьшая величина $\Delta X / X$ — относительной погрешности любого измерения X — получена при измерении нестабильности частоты $\Delta f / f$. Нестабильность транзисторных генераторов $\Delta f / f = 10^{-2}$. У ламповых генераторов $\Delta f / f = 10^{-3}$, у кварцевых генераторов $\Delta f / f = 10^{-6}$. Иногда стабильностью частоты называют степень, с которой частота генератора остается постоянной за время его непрерывной работы. При определении величины стабильности (нестабильности) необходимо указывать временной интервал t , за который производилось измерение.

Нестабильность возникает вследствие изменения внешних условий генератора: температуры, механических воздействий, флуктуации напряжения питания, шума транзисторов. Все это приводит к фазовым сдвигам в отдельных элементах генератора, вызывает нарушение баланса фаз на частоте f_0 и изменяет частоту генерируемых колебаний.

В высокостабильных генераторах должны быть обеспечены следующие условия:

1. Высокостабильные индуктивность L и емкость C .
2. Незначительная мощность в нагрузке, чтобы активные элементы работали в линейном режиме.
3. Питание генератора должно осуществляться от высокостабильных источников питания.

4. Должно быть термостатированное для постоянства температуры среды, окружающей генератор.

При выполнении этих мер нестабильность частоты оказывается порядка 10^{-4} . Для дальнейшего повышения стабильности надо увеличивать добротность колебательного контура. Стабильность частоты f генератора в основном определяется добротностью Q его резонансного контура. Резонансная кривая имеет острую вершину и степень изменения фазы с частотой тем больше, чем больше Q . Оба эти фактора определяют стабильность частоты генератора. В связи с этим любое эквивалентное сопротивление, включенное параллельно с резонансным контуром, снижает его добротность. Генераторы на полевом транзисторе, имеющие резонансный контур в цепи затвора, обладают добротностью от 100 до 500. Для многих практических целей требуются более высокая добротность и стабильность частоты. Это достигается в кварцевых генераторах. Рассмотрим некоторые физические особенности кварца, которые объясняют его применение для стабильности частоты.

Определенные кристаллы, в частности кварц (SiO_2), обладают пьезоэлектрическими свойствами. Они выражаются в том, что когда кристалл подвергается механической деформации (сжатие, растяжение), то появляется электрический потенциал между его противоположными гранями, это прямой пьезоэффект. Справедливо и обратное явление. Если к кристаллу кварца приложить электрическое поле, то он испытывает деформацию. Это обратный пьезоэффект. В результате этих пьезоэлектрических свойств тонкая кварцевая пластинка, вырезанная соответствующим образом по отношению к кристаллографическим осям и снабженная проводящими электродами — так называемый кварцевый резонатор — будет механически вибрировать, если на его электроды подано переменное напряжение. Эти колебания, в свою очередь, создадут электрические колебания. Механические и электрические колебания будут максимальны на собственной резонансной частоте механических колебаний кристалла. Колебательное движение кварца описывается уравнением

$$m \frac{d^2 x}{dt^2} + h \frac{dx}{dt} + kx = F(t), \quad (1)$$

где m — масса кристалла, h — коэффициент механического трения, k — коэффициент упругости кристалла. Это уравнение аналогично уравнению для тока в последовательном контуре

$$L \frac{d^2 i}{dt^2} + R \frac{di}{dt} + \frac{i}{C} = E(t). \quad (2)$$

Сравнивая эти уравнения, видим, что масса кристалла аналогична индуктивности, механическое трение эквивалентно сопротивлению и коэффициент упругости соответствует емкости. В силу идентичности этих уравнений полезно ввести механическую добротность по аналогии с электрической

$$Q = \frac{\omega m}{h}. \quad (3)$$

Когда частота внешнего переменного напряжения приближается к собственной частоте кварцевого резонатора, то в результате обратного пьезоэффекта резко возрастает амплитуда механических колебаний. Это, в свою очередь, через прямой пьезоэффект приведет к росту тока, который достигает максимума при совпадении частот внешнего переменного напряжения с собственной частотой кварцевого резонатора. Следовательно, частотные характеристики кварцевого резонатора и последовательного LC -контура одинаковы и он может быть представлен последовательным эквивалентным контуром.

Внутренние потери кварца очень малы, вследствие чего Q достигает 100000. Более того, упругие параметры таковы, что резонансные частоты механических колебаний охватывают диапазон от 10 кГц до нескольких десятков МГц в зависимости от размеров и формы кристалла. Пьезоэлектрические свойства кварцевого резонатора обуславливают потенциал на его электродах, соответствующий механическим колебаниям. Независимо от пьезоэлектрических свойств кварцевый резонатор представляет собой конденсатор с емкостью C_0 . Поэтому надо добавить еще емкость между электродами с кристаллом кварца в качестве диэлектрика. Вследствие этого кварцевый резонатор эквивалентен последовательному контуру LC , шунтированному емкостью C_0 и его полная эквивалентная схема будет такая, как на рис. 8.15, а.

Примерные значения электрических параметров кварцевого резонатора с частотой 90 кГц такие: $L=137$ Гн, $C = 0,023$ пФ, $R = 15$ кОм и $C_0 = 3,5$ пФ. Символическое обозначение кварцевого резонатора дано на рис. 8.15,б.

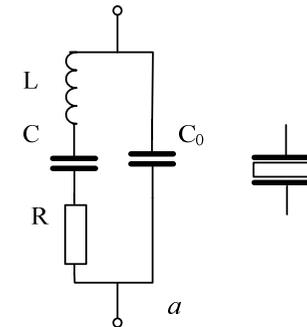


Рис. 1. Кварцевый резонатор:
а) эквивалентная схема; б) обозначение

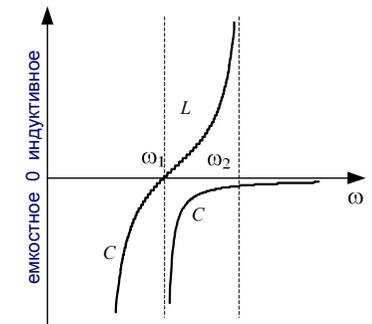


Рис. 2. Частотная характеристика сопротивления кварцевого резонатора

Частотная характеристика полной эквивалентной схемы имеет вид, показанный на рис. ????. Она имеет две резонансные частоты, соответствующие последовательному и параллельному

резонансам. Последовательный резонансный контур состоит из емкости и индуктивности, и его резонансная частота

$$\omega_1 = \frac{1}{\sqrt{LC}}. \quad (4)$$

Параллельный резонансный контур образован с учетом C_0 , и его резонансная частота

$$\omega_2 = \frac{1}{\sqrt{L \frac{CC_0}{C+C_0}}}. \quad (5)$$

Но так как $C_0 \gg C$, то приближенно $\omega_2 = 1/\sqrt{LC}$. Поэтому частоты последовательного и параллельного резонансов очень близки. Для любых частот кроме узкой полосы $\omega_2 - \omega_1$ электрическое сопротивление кварцевого резонатора имеет емкостной характер (в полосе $\omega_2 - \omega_1$ индуктивный). Это позволяет создавать схемы кварцевых генераторов.

4.4. Схемы кварцевых генераторов

Рассмотрим схему кварцевого генератора, где кварц включен между базой и коллектором (рис. 1). Это аналог емкостной трехточечной схемы генератора (рис.4.3.2), в которой роль резонансного контура играет кварцевый резонатор, Индуктивность L_{op} в цепи стока служит для подачи потенциала на сток, не допуская короткого замыкания

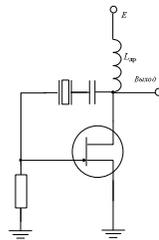


Рис. 1. Схема генератора полевом транзисторе с кварцевой стабилизацией

стока на землю по высокой частоте через источник питаия. Емкость C имеет малое сопротивление на частоте генерации и блокирует любой постоянный ток между стоком и затвором.

4.5. Релаксационные генераторы

В радиотехнике колебания, с которыми приходится иметь дело, делят на два типа: синусоидальные (гармонические) и релаксационные (разрывные). Эти колебания сильно различаются по форме. Осциллограмма релаксационных колебаний резко отличается от синусоиды и имеет характерную пилообразную форму, форму прямоугольника или другой более сложной

кривой. Электрические генераторы релаксационных колебаний характеризуются в общих чертах тем, что имеют накопитель энергии только одного типа, чаще всего это емкость. Процесс колебаний состоит из периодического притока энергии к накопителю и стока ее в какую-либо другую цепь. Обмен энергии осуществляется с помощью транзистора, играющего роль клапана, который закрыт, когда происходит накопление энергии и открыт, когда происходит передача энергии в другую систему. Транзистор работает, как своеобразный вентиль, переходя от закрытого состояния к открытому. Частота колебаний генератора определяется не только параметрами схем, но и в большой степени зависит от режима работы транзистора. Типичным примером автоколебательной системы релаксационных колебаний является мультивибратор. Колебания в мультивибраторе носят резко выраженный несинусоидальный разрывной характер и содержат большое количество гармоник. Этим объясняется название «мультивибратор», что означает генератор большого количества колебаний. Рассмотрим его работу (см рис.1).

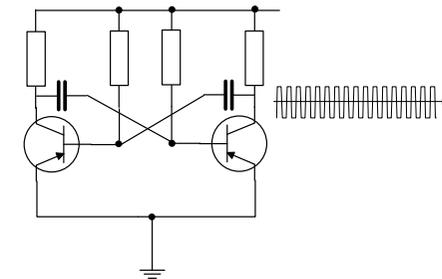


Рис.1. Схема мультивибратора

5. АКУСТО-ЭЛЕКТРИЧЕСКИЕ КОНВЕРТОРЫ СИГНАЛОВ

В этом разделе рассматриваются процессы распространения акустических волн в твердых телах и жидкостях. Нас интересует, какие типы волн могут существовать в различных материалах, как они отражаются и проходят через границы раздела двух сред, как эти волны возбуждаются и как их свойства связаны с плотностью и упругостью материалов.

Особый интерес представляют свойства пьезоэлектрических преобразователей, которые очень эффективны и поэтому широко используются в качестве ультразвуковых преобразователей электрической энергии в механическую.

5.1. Распространение акустических волн

Рассмотрим процессы распространения акустических волн в твердых телах и жидкостях. При этом нас будет интересовать типы волн, которые могут существовать в различных материалах, как они отражаются и проходят через границы раздела двух сред, как эти волны возбуждаются и как их свойства связаны с плотностью и упругостью материалов.

Определим вначале два основных типа волн. Первый — это *продольная волна*, в кото-

рой движение частиц среды происходит только в направлении распространения. Как показано на рис. 1.1.1, а, под действием сил среда расширяется или сжимается в направлении z - Второй тип — это *сдвиговая волна*, в которой движение частиц среды поперечно к направлению распространения (см. рис. 1,б). Сдвиговые волны связаны с кручением или изгибом

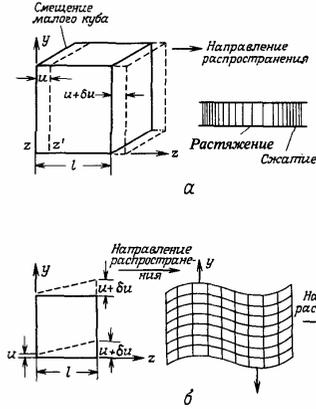


Рис. 1. а - распространение продольной волны; б - распространение поперечной (сдвиговой) волны.

Рис. 2. а — напряжение в продольном направлении в пластине длиной l ; б — напряжение в поперечном направлении.

материала (например, с кручением стержня). В случае сдвиговой моды не происходит изменений объема или плотности среды, как это видно из рис. 1,б. В общем случае акустическая волна в твердой среде может быть комбинацией сдвигового и продольного движения.

Выведем основное волновое уравнение для продольных акустических волн. Для этого введем сначала необходимые определения.

Напряжение. Сила, приложенная к твердому телу и отнесенная к единице площади, называется *напряжением*. В одномерном случае будем обозначать его символом T . Сила может растягивать или сжимать тело. Рассмотрим вначале слой материала бесконечно малой длины l , показанный на рис. 2 (а - случай продольного, б - сдвигового напряжения). Напряжение $E(z)$ — это сила, приходящаяся на единицу площади, действующая на частицы слева от плоскости z . Отметим, что продольное напряжение считается положительным, если внешнее напряжение, действующее на правую сторону слоя, имеет направление $+z$; внешнее напряжение, действующее на левую сторону слоя, равно $-T$ и имеет направление $-z$. Если считать напряжение положительным в направлении $+x$ или $+y$, то эти определения применимы также к сдвиговым напряжениям (см. рис. 2,б). Разность внешних напряжений, действующих на обе стороны слоя, равна $l dT/dz$. Таким образом, результирующая внешняя сила, приводящая в движение единичный объем среды относительно его центра масс, равна dT/dz .

Смещение и деформация. Пусть плоскость $z = \text{const}$ в материале под действием про-

дольного напряжения смещается в направлении оси z и занимает положение $z' = z + u$ (см. рис. 1, а). Параметр u называется *смещением* и в общем случае является функцией координаты z . В некоторой другой точке среды $z + \delta z$ смещение u изменяется и становится равным $u + \delta u$. Если u постоянно во всем объеме, то тело не деформируется и смещается как целое. Такие движения не представляют для нас интереса. Нас будет интересовать изменение смещений частиц с координатой z .

Пользуясь разложением в ряд Тейлора, можно показать, что в первом приближении изменение δu величины и на длине l дается выражением:

$$\delta u = \frac{\partial u}{\partial z} l = S l \tag{1}$$

где S — относительное удлинение материала

$$S = \frac{\partial u}{\partial z} \tag{2}$$

Параметр S называется *деформацией*.

Можно рассмотреть также одномерный случай сдвигового движения, при котором волна, распространяющаяся вдоль оси z , смещает частицы среды в направлении y . Как показано на рис. 1,б, частица с координатой $\mathbf{a}_z z + \mathbf{a}_y y$, где \mathbf{a}_z и \mathbf{a}_y — единичные векторы в направлениях y и z . По аналогии с рассмотрением, проведенным выше для продольных волн, можно определить сдвиговую деформацию как $S = \partial u / \partial z = \partial u / \partial z$. Единственное отличие в том, что смещение u происходит в направлении оси y . Рисунок показывает, что сдвиговое движение искажает форму выделенного в среде прямоугольника, но не изменяет его площади. Продольное же движение изменяет объем выделенного куба на $A \partial u$, где A — площадь в плоскости x, y . Поэтому относительное изменение объема равно $\partial u / l = S$.

Закон Гука и упругость. Закон Гука гласит, что в случае малых напряжений, приложенных к одномерной системе, напряжение пропорционально деформации:

$$T = c S, \tag{3}$$

где c — упругая постоянная материала. Величины T и c , вообще говоря, являются тензорными, но могут быть представлены одной компонентой для одномерной продольной или поперечной волны. Поскольку твердое тело легче изогнуть, чем растянуть, сдвиговая упругая постоянная обычно меньше, чем продольная.

Уравнение движения. Рассмотрим теперь уравнение движения точки в среде, в которой действует малое внешнее напряжение, изменяющееся во времени. Поскольку перемещающая сила на единицу площади равна $l \partial T / \partial z$, из второго закона Ньютона получаем:

$$\frac{\partial T}{\partial z} = \rho m \sigma U = \rho m \sigma$$

Здесь ν скорость частиц среды, σ плотность в стационарном состоянии.

Сохранение массы. Скорость частиц среды есть $v = u$, а поэтому изменение скорости на малой длине l в одномерном случае равно:

$$du = l \partial u / \partial z \quad (5)$$

В то же время из формулы (1.1.1) имеем

$$\partial v = \frac{\partial}{\partial t} \partial u = l \frac{\partial S}{\partial t} \quad (6)$$

Таким образом, из двух последних формул следует равенство

$$\frac{\partial S}{\partial t} = \frac{\partial v}{\partial z}$$

Это уравнение-одна из форм записи закона сохранения массы вещества в случае продольных волн. Аналогичное уравнение справедливо и для сдвиговых волн. Однако, поскольку при сдвиговом движении плотность среды не изменяется, уравнение (7) выполняется очевидным образом.

Напишем $\gamma_m = \gamma_{m0} + \gamma_{m1}$, где γ_{m0} — невозмущенная плотность, а γ_{m1} — ее возмущение. Тогда одномерное уравнение сохранения массы для продольных волн можно записать в виде

$$\frac{\partial}{\partial z} P_m v + \frac{\partial P_m}{\partial t} = 0 \quad (8)$$

Считая v и γ_{m1} малыми возмущениями и сохраняя только члены первого порядка малости, получаем:

$$\rho_{m0} \frac{\partial v}{\partial z} + \frac{\partial \rho_{m1}}{\partial t} = 0 \quad (9)$$

Таким образом,

$$\frac{\partial v}{\partial z} = - \frac{\partial}{\partial t} \left(\frac{\rho_{m1}}{\rho_{m0}} \right) = \frac{\partial S}{\partial t} \quad (10)$$

Волновое уравнение и определение постоянной распространения. Используя уравнения (3), (4), (7) и полагая $v = du/dt$, получаем волновое уравнение, описывающее распространение звуковой волны малой амплитуды:

$$\frac{\partial^2 T}{\partial z^2} = \rho_{m0} \frac{\partial^2 S}{\partial t^2} = \frac{\rho_{m0}}{c} \frac{\partial^2 T}{\partial t^2} \quad (12)$$

Решения этого уравнения для напряжений $T(z, t)$ имеют вид $F(t \pm z/V_a)$. Для волны с круговой частотой ω все полевые переменные изменяются как $\exp(i\omega t)$, а решения записываются в виде $\exp[j(\omega t \pm \beta az)]$. Здесь знак минус соответствует волне, бегущей в положительном направлении оси z , а знак плюс — обратной волне,

$$\beta a = \omega \sqrt{\frac{\rho_{m0}}{c}} = \frac{\omega}{V_a} \quad (12)$$

а $V_a = (c/\rho)^{1/2}$ — скорость распространения акустической волны. Параметр βa называется **постоянной распространения**.

Акустический импеданс. По аналогии с электромагнитной теорией и теорией линий передачи мы можем ввести для материала акустический импеданс Z , который точнее следует называть **удельным акустическим импедансом**. Напишем

$$z = - \frac{T}{v} \quad (32)$$

В случае плоской волны, распространяющейся в прямом направлении (величины, относящиеся к ней, будем помечать индексом F), определим **волновой импеданс** Z_0 :

$$Z_0 = - \frac{T_F}{v_F} = (p_{m0} c)^{1/2} = V_a \rho \quad (33)$$

Сдвиговые волны. Анализ распространения сдвиговых волн аналогичен случаю продольных волн. При распространении волны в z -направлении имеются компоненты напряжений как в z - так и в y -направлении. Например, изотропный материал может быть смещен в y -направлении, и тогда смещение u_y оказывается функцией координаты z . Сдвиговое движение частиц в материале эквивалентно вращению вокруг оси x . Сдвиговая деформация t определяется одномерным параметром $S = \partial u_y / \partial z$. Вывод уравнения движения для сдвиговых волн аналогичен описанному выше; единственное различие заключается в том, что в нем необходимо использовать соответствующий коэффициент упругости, связывающий сдвиговое напряжение и деформацию. Этр приводит к тому, что скорости звуковых волн и волновые импедансы оказываются меньшими, чем для продольных волн в тех же материалах.

Волны растяжения. Еще одним типом волн, важных для практики, которые можно описать представленной здесь простой одномерной теорией, являются **волны растяжения** в тонком стержне. Предполагается, что для этой моды, показанной на рис. 6, поперечные размеры стержня (в случае цилиндрического стержня — его диаметр) намного меньше длины рассматриваемой акустической волны. В этом случае единственной отличной от нуля компонентой напряжения является продольная компонента, параллельная направлению распространения волны в стержне. Внутри стержня не могут возникать напряжения, перпендикулярные его оси, так как нормальное напряжение должно обращаться в нуль на его поверхности. Поскольку стержень свободно увеличивается в поперечном сечении при сжатии в направлении оси z , с распространением волны вдоль стержня связано движение частиц как вдоль оси стержня, так и перпендикулярно ей. Однако компонента деформации $S = du_z / dz$, обусловленная движением в направлении оси z , непосредственно связана с напряжением в z -направлении законом Гука

(3), в котором необходимо использовать другой эффективный коэффициент упругости. Соответствующий коэффициент упругости называется *модулем Юнга*, который мы обозначим через E ; таким образом, можно написать $T = ES$. Модуль Юнга - это коэффициент упругости, для измерения которого необходимо длинную проволоку или стержень растягивать статической силой. В любом случае модуль Юнга меньше, чем продольный коэффициент упругости, поскольку для растяжения слоя материала, сжатие или расширение которого ограничено в поперечном направлении, с той же деформацией S необходима большая сила.

5.2. Возбуждение и прием акустических волн

Принцип действия преобразователей, служащих для возбуждения и приема акустических волн, основан на использовании ряда физических явлений - пьезоэффекта, магнитоэлектрических, электрострикции и др., связанных с возникновением заметных механических деформаций при воздействии на тело электрических, магнитных полей или различного рода излучений.

Способы возбуждения (приема), основанные на использовании - пьезоэлектрического эффекта. Наиболее эффективный и широко распространенный способ возбуждения (приема) объемных волн состоит в использовании пьезоэлектрических полуволновых слоев (пластин) с электродами, находящимися в акустическом контакте со звукопроводом. В основе этого метода лежит использование для излучения акустических волн обратного пьезоэффекта (деформации пластины под действием электрического поля, приложенного к электродам), а для приема - прямого пьезоэффекта (возникновение электрического заряда на обкладках - электродах деформированной пьезоэлектрической пластины).

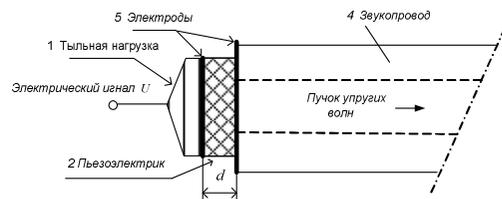


Рис. 1. Структура резонансного преобразователя для объемных акустических волн.

Устройство преобразователя схематически показано на рис. 1. Пьезоэлектрический слой 2 толщиной d с электродами 5 акустически связан со звукопроводом 4 и тыльной нагрузкой 1 через промежуточные слои. При приложении к электродам внешнего переменного электрического напряжения U в пьезоэлектрическом слое возникают колебания сдвига или сжатия - растяжения по толщине. В результате этого в звукопроводе возбуждаются соответственно сдвиговые (поперечные) или продольные объемные акустические волны. Эффективность вза-

имного преобразования электрической энергии в акустическую достигает максимума на частоте основного акустического резонанса $f_0 = V/2d$ (V - скорость звука в пьезоэлектрическом слое) и на нечетных высших гармониках $f_{2n+1} = (2n-1)f_0$, где $n=1,2,\dots$

Таблица 1.

мода	Материал	Скорость звука, м/с	Акустический импеданс, $10^6 \text{ Н} \cdot \text{с} / \text{м}^3$	КЭМС k
Продольная	LiNbO ₃	7340	34,5	0,49
	LiNbO ₃	7316	34,4	0,16
	Кварц	5747	15,2	0,092
	Керамика ЦТС	4560	34,2	0,51
Поперечная	LiNbO ₃	4560	21,4	0,62
	Кварц	3317	8,8	0,10
	Кварц	5105	13,5	0,037
	Кварц	3320	8,8	0,08
	Керамика ЦТС	2600	19,5	0,70

Пластины преобразователя изготавливаются из материалов с хорошими пьезоэлектрическими свойствами. Критерием при выборе материала служит коэффициент электромеханической связи (КЭМС) k , приведенный для различных пьезоэлектриков в приложении А. Окончательный выбор материала для преобразователя зависит от величины k и акустических параметров, а также от: стоимости, возможности получения заготовок необходимых размеров, простоты обработки и т. д. В настоящее время для изготовления преобразователей объемных волн наиболее часто используются кристаллы SiO₂, LiNbO₃, CdS, ZnO, пьезокерамики различных составов и другие материалы, основные данные для которых приведены в табл. 1.

В зависимости от электрофизических параметров используемых пьезоэлектриков, пьезопреобразователи можно условно подразделить на пьезодиэлектрические и пьезополупроводниковые.

Пьезодиэлектрические пластинчатые преобразователи могут быть изготовлены из материалов с большим k , например из пьезокерамики или LiNbO₃. При соблюдении необходимой кристаллографической ориентации пластин они имеют практически только одну моду колебаний. Обычно они применяются в двух модификациях. Для диапазона низких частот (менее 100 МГц) - в виде отдельных пластин или пластин, наклеенных на прочную основу - «пирамидку», одновременно служащую «тыльной» акустической нагрузкой пьезоэлектрического полуволнового слоя. В этом случае преобразователь может быть использован в качестве съемного элемента. Для высоких частот (более 100 МГц) вследствие больших потерь в промежуточном «склеивающем» слое и хрупкости тонких пластин применение съемных преобразователей за-

трудно. Поэтому заготовка преобразователя специальным образом приклеивается непосредственно к звукопроводу и после этого доводится до необходимой толщины.

Пьезополупроводниковые преобразователи изготавливаются известными методами полупроводниковой технологии и в зависимости от способа создания преобразующего слоя подразделяются на следующие основные типы а - преобразователи с диффузионным высокоомным слоем, полученным в приповерхностной области низкоомного пьезополупроводника диффузией компенсирующей примеси; б—пленочные, изготовленные нанесением хорошо ориентированных текстурированных пьезоэлектрических пленок на поверхность звукопровода; в — эпитаксиальные, полученные эпитаксиальным наращиванием тонких монокристаллических слоев пьезоэлектрика на низкоомные ориентирующие подложки.

Эпитаксиальные преобразователи, хотя и могут быть выращены только на ориентирующих подложках, вследствие монокристалличности структуры имеют ряд преимуществ по сравнению с текстурированными, так как они обладают большей величиной k , меньшим уровнем генерации колебаний неосновного типа и более высоким пробивным напряжением, позволяющим получать, большую интенсивность акустических волн.

Изготовление высокочастотных пьезополупроводниковых преобразователей значительно проще, чем пьезодиэлектрических. Они позволяют осуществить генерацию и прием объемных волн до частот порядка 70 ГГц независимо от пьезоэлектрических свойств звукопровода.

5.3. Другие способы возбуждения и приема звуковых волн

Существует множество механизмов и приборов, в которых механические системы и электрические цепи конструктивно объединены. Электрические машины, гидрофоны, измерители вибраций и другие приборы преобразуют механическую энергию в электрическую. В электромоторах, излучателях звука, различных электроизмерительных приборах происходит обратное преобразование - электрической энергии в механическую. В таких системах механические и электрические процессы взаимосвязаны, и поэтому системы называют *электромеханическими*.

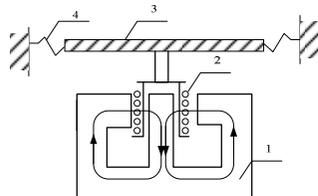


Рис. 1. Электродинамический преобразователь

Рассмотрим процесс электромеханического преобразования, происходящий в электродинамическом излучателе (приемнике), принцип действия которого основан на явлении взаимодействия тока с постоянным магнитным полем (рис. 1). В зазоре цилиндрического постоянного магнита 1 расположена катушка 2, состоящая из одного или нескольких рядов изолированного провода, намотанного на цилиндрический каркас, и жестко скрепленная с пластиной 3. Последняя способна совершать поршневые (как единое целое) колебания благодаря гибкому подвесу 4.

Когда преобразователь используют как излучатель акустических колебаний, по катушке пропускают переменный ток I и образуется магнитное поле. Переменное поле, созданное током, взаимодействует с постоянным полем, существующим в зазоре постоянного магнита, в результате чего на катушку действует переменная сила f , пропорциональная силе тока:

$$f = B_0 I l \quad (1)$$

где B_0 —магнитная индукция в зазоре; l — длина провода катушки.

При наличии механических колебаний электрическое сопротивление преобразователя складывается из сопротивления Z_0 при $V = 0$ и *динамического сопротивления* Z_M , обусловленного движением (динамикой) механической части. Сопротивление Z_M отражает взаимодействие электрической цепи и механической системы преобразователя, количественно выражает реакцию механических колебаний на электрические. Величина $K = B_0 l$ определяет количественную связь между динамическим и механическим сопротивлениями, равно как и между силой и током, электрическим напряжением и скоростью. Коэффициент K , называемый *коэффициентом электромеханического преобразования*, измеряют в ньютонах на ампер (Н/А) или в вольт-секундах на метр ($V \cdot c/m$). Коэффициент K показывает, какая сила действует на механическую систему преобразователя, когда по его электрической цепи протекает ток 1 А, или какое напряжение развивается на клеммах электрической цепи преобразователя при колебаниях его механической системы со скоростью 1 м/с.

5.4. Основные характеристики преобразователей акустических волн

Для описания свойств всех типов преобразователей акустических волн используются следующие основные характеристики.

Потери преобразования (ПП) —выраженное в децибелах отношение электрической мощности $P_p^э$, которую генератор может выделить в согласованную нагрузку, к мощности акустической волны $P_u^а$, излучаемой преобразователем, подключенным к этому генератору:

$$ПП = 10 \lg(P_p^э / P_u^а). \quad (1)$$

Величина $P_p^э$ - мощность генератора. При излучении преобразователем нескольких типов аку-

стических волн преобразование каждой из них описывается своей собственной величиной $ПП$.

Частотная зависимость потерь преобразования — зависимость величины $ПП$ от частоты в интересующем диапазоне частот. Определяется для полезно используемой акустической волны и других неосновных типов излучаемых акустических волн. Типичная форма частотных зависимостей величин $ПП$ основного и неосновных типов излучаемых акустических волн показана на рис. 3.

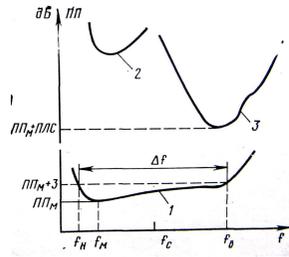


Рис. 3. Типичная форма частотных зависимостей потерь преобразования основного (1) и неосновных (2 и 3) типов излучаемых акустических волн

Минимальная величина потерь преобразования $ПП_M$

Частота минимума $ПП$, f_M

Ширина полосы пропускания Δf — область частот вблизи f_M , в пределах которой величина $ПП$ увеличивается по сравнению с $ПП_M$ не более чем на определенную величину (обычно на 3 дБ), $\Delta f = f_a - f_n$, где f_a, f_n — соответственно верхняя и нижняя граничные частоты полосы пропускания.

Относительная ширина полосы пропускания f_a / f_n — выраженное в процентах отношение ширины полосы пропускания Δf к средней частоте пропускания $f_c = 0.5(f_a - f_n)$.

Частотная зависимость входного полного электрического сопротивления преобразователя $Z_{вх}^3 = \text{Re}(Z_{вх}^3) + \text{Im}(Z_{вх}^3)$, где $\text{Re}(Z_{вх}^3)$ и $\text{Im}(Z_{вх}^3)$ — активная и реактивная составляющие полного электрического сопротивления.

5.5. Электрическое и акустическое согласование преобразователей

Рассмотрим излучающий и приемный преобразователи, включенные в радиотракт, с точки зрения оптимизации процесса передачи в них электрической мощности между источником сигнала и нагрузкой. При этом в соответствии с принципами теории электрических цепей излучающий преобразователь следует заменить его электрическим эквивалентом — пассивным двухполюсником с полным электрическим сопротивлением $Z_{вх}^3$, а приемный — активным

двухполюсником с полным электрическим сопротивлением $Z_{вых}^3$ и ЭДС ξ , определяемой амплитудой падающей акустической волны в звукопроводе. Очевидно, что в обоих случаях задача оптимизации передачи мощности сводится к анализу одной и той же эквивалентной электрической схемы, состоящей из последовательно соединенных с помощью отрезка длиной линии эквивалентного генератора и комплексного сопротивления нагрузки (рис. 1, а).

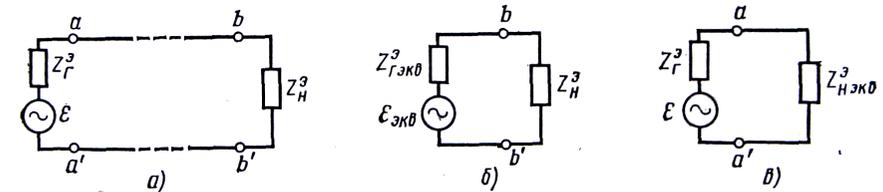


Рис. 1. К анализу процесса передачи мощности между эквивалентным питающим генератором и нагрузкой.

В рассматриваемом случае генератор обеспечит выделение максимальной мощности в нагрузке, если его эквивалентное внутреннее сопротивление, пересчитанное к точкам подключения нагрузки (b, b' рис 4. б), является величиной, комплексно-сопряженной полному сопротивлению нагрузки, т. е. $Z_{r экв}^3 = (Z_{н}^3)^*$. Условие выделения максимальной мощности в нагрузку можно сформулировать также в эквивалентной форме $Z_{н}^3 = (Z_{r экв}^3)^*$, пересчитав полное сопротивление нагрузки к точкам подключения к отрезку длиной линии генератора (a, a' , рис. 1, б).

Задача оптимизации процесса передачи мощности между источником сигнала и нагрузкой имеет самое непосредственное отношение к уменьшению потерь преобразования. Действительно, в любом излучающем преобразователе по крайней мере часть поглощенной внутри него мощности излучается в виде акустической волны. В связи с этим, передавая максимальную мощность в преобразователь, мы тем самым обеспечим излучение преобразователем максимума акустической мощности. Аналогично этому, в результате оптимизации передачи мощности от приемного преобразователя нагрузке, обеспечивается выделение в нагрузке максимальной электрической мощности и тем самым снижаются потери преобразования при приеме акустической волны.

Рассмотрим включение в согласованный радиотракт излучающего преобразователя «без потерь», единственным источником поглощения мощности в котором является полезное преобразование электрической энергии в акустическую и излучение последней в звукопровод. В согласованном радиотракте, в котором $Z_r^3 = \rho_0$, отраженная волна в отрезке длиной линии может появиться только при рассогласовании между характеристическим сопротивлением длиной линии ρ_0 и входным полным сопротивлением преобразователя $Z_{вх}^3$. Отношение

мощностей P_0^2 и P_{II}^2 , переносимых отраженной и падающей волнами в длину линии, характеризуется коэффициентом отражения

$$|r^3|^2 = P_0^2 / P_{II}^2 \quad (1)$$

Применяя закон сохранения энергии к преобразователю найдем, что поглощаемая в нем мощность $P_{\text{полл}}^2$ следующим образом связана с P_0^2 и P_{II}^2

$$P_{\text{полл}}^2 = P_{II}^2 - P_0^2 = P_P^2 - P_0^2 \quad (2)$$

Откуда вытекает, что

$$P_{\text{полл}}^2 / P_P^2 = 1 - |r^3|^2 \quad (3)$$

Поскольку в преобразователях с малыми потерями $P_{\text{полл}}^2 \approx P_{II}^2$, то в соответствии с соотношениями (3.1) и (4.36)

$$ППП = 10 \lg \frac{1}{1 - |r^3|^2} \quad (4)$$

Отсюда видно, что при $r^3=0$ излучающий преобразователь «без потерь» в случае полного электрического согласования с радиотрактом имеет нулевые потери преобразования.

Помимо электрического согласования, возможно также акустическое согласование преобразователей. Влияние акустического согласования, в принципе включающего не только выбор параметров промежуточного слоя, но также выбор акустических импедансов сред $Z=V\rho$, иллюстрируется рис. 2. Возможности акустического согласования можно существенно расширить, используя последовательность нескольких слоев из материалов с различными акустическими импедансами.

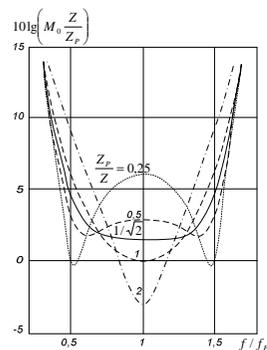


Рис. 2. Система преобразователь - звукопровод.

Если акустический импеданс звукопровода Z меньше, чем импеданс пьезопреобразователя Z_p , то преобразователь резонирует на частоте, соответствующей $\lambda_p/2$ (например, для $Z_p/Z = 2$). В противном случае резонанс имеет место на частотах, соответствующих $\lambda_p/4$ и $3\lambda_p/4$ ($Z_p/Z = 0,25$).

Ход этих зависимостей объясняется следующим образом. Если импеданс звукопровода мал по сравнению с импедансом пьезоэлектрика, то поверхности преобразователя можно считать практически свободными, и преобразователь резонирует на половине длины волны (например, для $Z_p/Z = 2$). Поэтому при $f = f_p$ сопротивление излучения имеет минимум, который соответствует максимуму мощности возбуждаемых упругих волн. Если же, наоборот, Z велико по сравнению с Z_p (например, $Z_p/Z = 0,5$ или $0,25$), преобразователь, у которого одна поверхность свободна, а другая оказывается достаточно закрепленной, резонирует на частотах, соответствующих $\lambda_p/4$ и $3\lambda_p/4$. Сопротивление излучения оказывается минимальным вблизи значений $f = 0,5 f_p$ и $f = 1,5 f_p$. Значение $Z_p/Z = 1/\sqrt{2}$ является критическим. При меньших значениях отношения Z_p/Z кривые на рис. 2 имеют уже два минимума. Вблизи критического значения Z_p/Z форма частотной характеристики преобразователя приближается к прямоугольной с достаточно широкой плоской частью по обе стороны от резонансной частоты f_p .

6. ЭЛЕМЕНТЫ ОПТОЭЛЕКТРОНИКИ И ИНФРАКРАСНОЙ ТЕХНИКИ

Под оптоэлектроникой понимается направление электроники, охватывающее вопросы использования оптических и электрических методов обработки, хранения и передачи информации. оптоэлектроника возникла как этап развития радиоэлектроники и вычислительной техники, тенденцией которых является непрерывное усложнение систем при возрастании их информационных и технико-экономических показателей (увеличение надёжности, быстродействия, уменьшение размеров и веса). Идея использования света для обработки и передачи информации уже давно реализована: большая группа фотоприёмников (фотоэлементов, фотоэлектронных умножителей, фоторезисторов, фотодиодов, фототранзисторов и пр.) служит для преобразования световых сигналов в электрические. Существуют также и преобразователи последовательности электрических сигналов в видимое изображение. Вся же обработка информации в электрических трактах радиоэлектронных устройств осуществлялась вакуумными и полупроводниковыми приборами.

Оптоэлектроника отличается от вакуумной и полупроводниковой электроники наличием в цепи сигнала оптического звена или оптической (фотонной) связи. Достоинства оптоэлектроники определяются в первую очередь преимуществами оптической связи по сравнению с электрической, а также теми возможностями, которые открываются в результате использования разнообразных физических явлений, обусловленных взаимодействием световых полей с твёрдым телом.

Из-за электрической нейтральности фотонов в оптическом канале связи не возбужда-

ются электрические и магнитные поля, сопутствующие протеканию электрического тока. Иными словами, фотоны не создают перекрестных помех в линиях связи и обеспечивают полную электрическую развязку между передатчиком и приёмником, что принципиально недостижимо в цепях с электрической связью. Передача информации с помощью светового луча не сопровождается накоплением и рассеиванием электромагнитной энергии в линии. Отсюда - отсутствие существенного запаздывания сигнала в канале связи, высокое быстродействие и минимальный уровень искажения передаваемой информации, переносимой сигналом.

Высокая частота оптических колебаний (10^{14} - 10^{15} гц.) обуславливает большой объём передаваемой информации и быстродействие. Соответствующая оптической частоте малая длина волны (до $\lambda \sim 1^{-10}$ мкм.) открывает пути для микроминиатюризации передающих и приёмных оптоэлектронных устройств, а также линии связи. Минимальные поперечные размеры светового луча - порядка длины волны λ . Информационная ёмкость такого канала вследствие его большой широкополосности чрезвычайно высока.

Идеи оптоэлектроники возникли ещё в 1955, но известные в то время средства для взаимного преобразования электрических и оптических сигналов и для осуществления оптической связи не обеспечивали необходимых эффективности, быстродействия, мощности светового потока, возможности микроминиатюризации. оптоэлектроника начала интенсивно развиваться лишь с 1963-65, после того как появились лазеры, полупроводниковые светоизлучающие диоды и волоконная оптика.

6.1. Источники оптического излучения для систем передачи

В качестве источника оптического излучения в оптоэлектронике используются приборы, преобразующий электрическую энергию возбуждения в энергию оптического излучения заданного спектрального состава и пространственного распределения.

В технике связи в настоящее время наибольшее распространение получили следующие типы оптоэлектронных излучателей:

-светоизлучающие полупроводниковые диоды (СИД);

-инжекционные полупроводниковые лазерные диоды (ППЛД);

Оба типа излучателей является основными приборами в передающей связной аппаратуре.

Параметрами, которыми характеризуются указанные оптоэлектронные источники излучения, являются:

1. Эффективность преобразования энергии возбуждения в энергию излучения.
2. Ширина спектральной линии излучения.
3. Направленность излучения и связанная с ней концентрация мощности излучения на единице площади, т.е. интенсивности излучения I :

$$I = \frac{cn}{4\pi} E^2 \quad [\text{Вт/см}^2]$$

где n - показатель преломления, c - скорость света, E - напряженность светового поля.

4. Быстродействие источника излучения при модуляции.
6. Степень когерентности излучения.
7. Габаритные размеры источника.
8. Технологичность изготовления и стоимость.
9. Срок службы (обычно $\sim 10^5$ часов)
10. Устойчивость к механическим, тепловым, радиационным перегрузкам.

6.2. Светоизлучающие диоды.

Светодиод (СИД) представляет собой полупроводниковый прибор с р - n переходом, протекание электрического тока через который вызывает интенсивное спонтанное излучение.

6.2.1. Конструкции светодиодов для оптической связи

В технике оптической связи наибольшее применение получили две конструкции СИД: поверхностный (рис. 1) и торцевой (рис. 2)

В поверхностном СИД волоконный световод присоединяется к поверхности излучения через специальную выемку в полупроводниковой подложке. Такой способ стыковки СИД и стекловолокна обусловлен необходимостью ввода максимальной мощности спонтанного излучения в световод

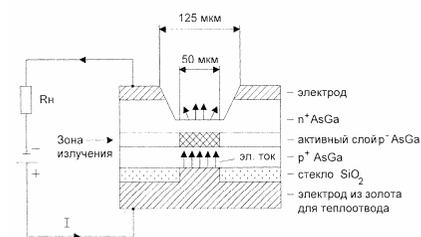


Рис. 1. Конструкция светодиода поверхностного типа (диод Баррасса).

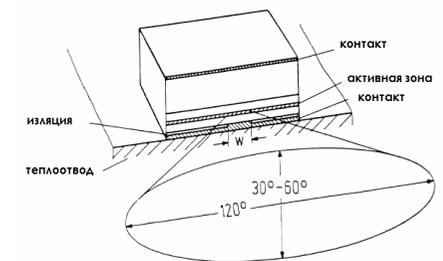


Рис. 2. Конструкция торцевого СИД

В конструкции торцевого светодиода предусмотрен вывод оптической мощности излучения через один из торцов. При этом другой торец выполнен в виде зеркала, которое отражает фотоны в активный слой. В приборе применяются дополнительные слои полупроводникового материала GaAlAs, который отличается от активного слоя показателем преломления и

шириной запрещенной зоны. Это создает в активном слое оптический волновод, способствующий концентрации фотонов и усилению бегущей волны в инверсной насыщенной зарядом среде. Светоизлучающий торец СИД согласуется с волоконным световодом и линзовой системой.

6.2.2. Принцип действия светодиода

Работа светодиода основана на случайной рекомбинационной люминесценции избыточных носителей заряда, инжектируемых в активную область светодиода. В результате инжекции неосновных носителей заряда и дрейфа основных в активном слое происходит накопление и рекомбинация этих зарядов с выделением квантов энергии $h\nu$, которые примерно соответствуют ширине E_g запрещенной зоны активного слоя:

$$E_{\phi} \approx E_g = h\nu \quad (1)$$

При этом фотоны, случайно образовавшиеся, могут двигаться в любом случайном направлении, отражаться от границ различных слоев полупроводников, поглощаться кристаллами и излучаться с поверхности или из торца СИД (рис. 2.1). Величина излучаемой мощности СИД примерно линейно зависит от величины тока инжекции. Благодаря некоторым вышеперечисленным особенностям конструкции торцевого СИД в нем может происходить образование небольшого числа стимулированных фотонов. Это способствует увеличению общей мощности излучаемой энергии и ее концентрации в пространстве. По этой причине торцевые СИД называются слабокогерентными источниками света - *суперлюминесцентными* диодами (СЛД).

6.2.3. Основные характеристики светодиодов

а) Ватт-амперная характеристика светодиодов показывает зависимость излучаемой мощности от тока, протекающего через прибор (рис. 2.3).

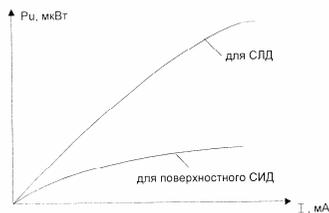


Рис. 1. Ватт-амперная характеристика СИД

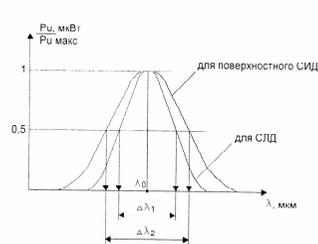


Рис. 2. Спектральная характеристика СИД

Характеристика имеет линейный и нелинейные участки. Нелинейность обусловлена предельными возможностями СИД по спонтанной рекомбинации электронов и дырок и их ог-

раниченным числом, зависящим от насыщенности примесными компонентами и общего объема активного слоя. Ватт-амперная характеристика зависит от температуры кристалла. С ее повышением мощность излучения может значительно снижаться.

б) Спектральная характеристика светодиодов показывает зависимость излучаемой мощности от длины волны излучения (рис. 2.4). По спектральной характеристике можно определить ширину спектра излучения на уровне половинной от максимальной мощности излучения. Ширина спектра СЛД $\Delta\lambda_1$, (около $10 \div 30$ нм), для поверхностного СИД $\Delta\lambda_2$ (около $30 \div 60$ нм).

Относительно узкий спектр излучения СЛД объясняется волноводным эффектом и некоторой согласованностью (когерентностью) излучательных рекомбинаций. При этом характер излучения остается спонтанным и ширина спектра определяется разбросом энергетических состояний рекомбинирующих электронов и дырок.

в) Диаграмма направленности излучения светодиода показывает распределение энергии излучения в пространстве (рисунок 2.5)

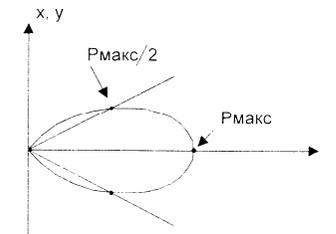


Рис. 3. Угловая расходимость излучения СИД.

Угловая расходимость излучения $\Delta\phi$ оценивается на уровне уменьшения мощности в пространстве в два раза ($P_{\max}/2$, что отмечено на рисунке точками на пересечении лучей и кривых распределения мощности. Для поверхностного СИД величина $\Delta\phi$ и может составлять $110^\circ \dots 180^\circ$. Для СЛД величина $\Delta\phi$ составляет примерно: $60^\circ \div 30^\circ$.

г) Внешняя квантовая эффективность светодиода показывает долю выводимой мощности излучения от полученной в результате спонтанной рекомбинации

$$\eta_{\text{внеш}} = \frac{P_{\text{излуч}}}{P_{\text{рекомб}}} \quad (1)$$

Обычно эта доля не превышает 2-10 %, что обусловлено большими потерями из-за рассеяния мощности внутри прибора и отражением фотонов на границе "полупроводник - воздух" и "полупроводник - световод" из-за различных показателей преломления полупроводника ($n = 3,5$) и среды ($n = 1,5$).

6.3. Лазеры для систем связи.

Лазер (Laser - Light Amplification by Stimulated Emission of Radiation) или *оптические квантовые генераторы* – это современные когерентные источники излучения, обладающие целым рядом уникальных свойств. Создание лазеров явилось одним из самых замечательных достижений физики второй половины XX века, которое привело к революционным изменениям во многих областях науки и техники. К настоящему времени создано большое количество лазеров с различными характеристиками – газовых, твердотельных, полупроводниковых, излучающих свет в различных оптических диапазонах. Лазеры могут работать в импульсном и непрерывном режимах. Мощность излучения лазеров может изменяться в пределах от долей милливатта до 10^{12} – 10^{13} Вт (в импульсном режиме). Лазеры находят широкое применение в военной технике, в технологии обработки материалов, в медицине, в оптических системах навигации, связи и локации, в прецизионных интерференционных экспериментах, в химии, просто в быту и т. д. Хотя первый лазер был построен сравнительно недавно (1960 г.), современную жизнь уже невозможно представить без лазеров.

Одним из важнейших свойств лазерного излучения является чрезвычайно высокая степень его монохроматичности, недостижимая в излучении нелазерных источников. Это и все другие уникальные свойства лазерного излучения возникают в результате согласованного, кооперативного испускания световых квантов многими атомами рабочего вещества.

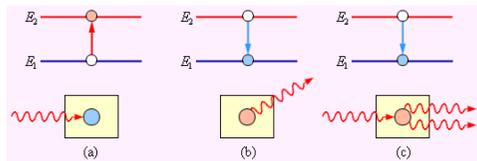


Рис. 1. Условное изображение процессов (a) поглощения, (b) спонтанного испускания и (c) индуцированного испускания кванта.

Чтобы понять принцип работы лазера, нужно более внимательно изучить процессы поглощения и излучения атомами квантов света. Атом может находиться в различных энергетических состояниях с энергиями E_1 , E_2 и т. д. В теории Бора эти состояния называются стабильными. На самом деле стабильным состоянием, в котором атом может находиться бесконечно долго в отсутствие внешних возмущений, является только состояние с наименьшей энергией. Это состояние называют основным. Все другие состояния нестабильны. Возбужденный атом может пребывать в этих состояниях лишь очень короткое время, порядка 10^{-8} с, после этого он само-

произвольно переходит в одно из низших состояний, испуская квант света, частоту которого можно определить из второго постулата Бора. Излучение, испускаемое при самопроизвольном переходе атома из одного состояния в другое, называют *спонтанным*. На некоторых энергетических уровнях атом может пребывать значительно большее время, порядка 10^{-3} с. Такие уровни называются *метастабильными*.

Переход атома в более высокое энергетическое состояние может происходить при резонансном поглощении фотона, энергия которого равна разности энергий атома в конечном и начальном состояниях.

Переходы между энергетическими уровнями атома не обязательно связаны с поглощением или испусканием фотонов. Атом может приобрести или отдать часть своей энергии и перейти в другое квантовое состояние в результате взаимодействия с другими атомами или столкновений с электронами. Такие переходы называются *безизлучательными*.

В 1916 году А. Эйнштейн предсказал, что переход электрона в атоме с верхнего энергетического уровня на нижний может происходить под влиянием внешнего электромагнитного поля, частота которого равна собственной частоте перехода. Возникающее при этом излучение называют *вынужденным* или *индуцированным*. Вынужденное излучение обладает удивительным свойством. Оно резко отличается от спонтанного излучения. В результате взаимодействия возбужденного атома с фотоном атом испускает еще один фотон той же самой частоты, распространяющийся в том же направлении. На языке волновой теории это означает, что атом излучает электромагнитную волну, у которой частота, фаза, поляризация и направление распространения точно такие же, как и у первоначальной волны. В результате вынужденного испускания фотонов амплитуда волны, распространяющейся в среде, возрастает. С точки зрения квантовой теории, в результате взаимодействия возбужденного атома с фотоном, частота которого равна частоте перехода, появляются два совершенно одинаковых фотона-близнеца.

Именно индуцированное излучение является физической основой работы лазеров. На рис.1 схематически представлены возможные механизмы переходов между двумя энергетическими состояниями атома с поглощением или испусканием кванта.

Рассмотрим слой прозрачного вещества, атомы которого могут находиться в состояниях с энергиями E_1 и $E_2 > E_1$. Пусть в этом слое распространяется излучение резонансной частоты перехода $\nu = \Delta E / h$. Согласно распределению Больцмана, при термодинамическом равновесии большее количество атомов вещества будет находиться в нижнем энергетическом состоянии. Некоторая часть атомов будет находиться и в верхнем энергетическом состоянии, получая необходимую энергию при столкновениях с другими атомами. Обозначим населенности нижнего и верхнего уровней соответственно через n_1 и $n_2 < n_1$. При распространении резонансного излучения в такой среде будут происходить все три процесса, изображенные на рис.2 Эйнштейн показал, что процесс (a) поглощения фотона невозбужденным атомом и процесс (c)

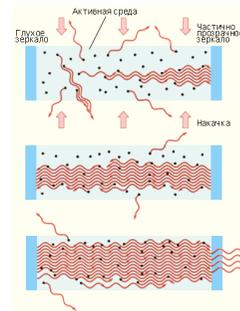


Рис.2. Развитие лавинообразного процесса генерации в лазере.

индуцированного испускания кванта возбужденным атомом имеют одинаковые вероятности. Так как $n_2 < n_1$ поглощение фотонов будет происходить чаще, чем индуцированное испускание. В результате прошедшее через слой вещества излучение будет ослабляться. Излучение, возникающее в результате спонтанных переходов, некогерентно и распространяется во всевозможных направлениях и не дает вклада в проходящую волну.

Чтобы проходящая через слой вещества волна усиливалась, нужно искусственно создать условия, при которых $n_2 > n_1$, т. е. создать *инверсную населенность уровней*. Такая среда является термодинамически неравновесной. Идея использования неравновесных сред для получения оптического усиления впервые была высказана В. А. Фабрикантом в 1940 году. В 1954 году русские физики Н. Г. Басов и А. М. Прохоров и независимо от них американский ученый Ч. Таунс использовали явление индуцированного испускания для создания микроволнового генератора радиоволн с длиной волны $\lambda = 1,27$ см. За разработку нового принципа усиления и генерации радиоволн в 1964 году все трое были удостоены Нобелевской премии.

Среда, в которой создана инверсная населенность уровней, называется *активной*. Она может служить резонансным усилителем светового сигнала. Для того, чтобы возникла генерация света, необходимо использовать *обратную связь*. Для этого активную среду нужно расположить между двумя высококачественными зеркалами, отражающими свет строго назад, чтобы он многократно прошел через активную среду, вызывая лавинообразный процесс индуцированной эмиссии когерентных фотонов. При этом в среде должна поддерживаться инверсная населенность уровней. Этот процесс в лазерной физике принято называть *накачкой*.

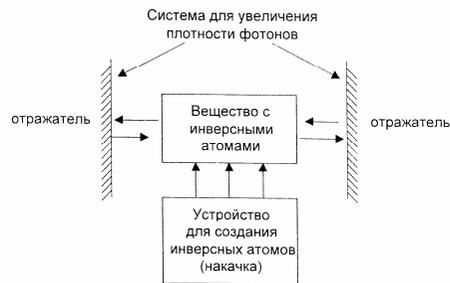


Рис. 3. Обобщенная конструкция лазера

Начало лавинообразному процессу в такой системе при определенных условиях может положить случайный спонтанный акт, при котором возникает излучение, направленное вдоль оси системы. Через некоторое время в такой системе возникает стационарный режим генерации. Это и есть лазер. Лазерное излучение выводится наружу через одно (или оба) из зеркал, обладающее частичной прозрачностью. На рис. 2. схематически представлено развитие лавинообразного процесса в лазере.

Некоторые конструкции лазерных диодов представлены на рис.4

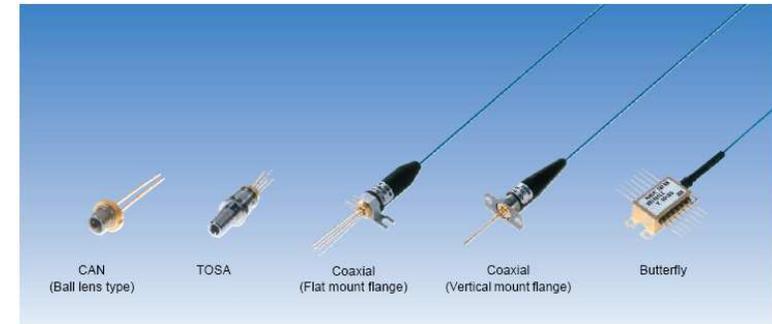


Рис.4. Конструкции лазерных диодов для систем оптической связи

6.4. Фотоприёмники оптических систем связи

6.4.1 Характеристики фотоприемников

Фотоприемники (фотодетекторы) предназначены для преобразования оптического сигнала в электрический. Так же они осуществляют детектирование оптического сигнала, т. е. его демодуляцию, и по этому носят название *фотодетекторы*. Фотоприемник как оптический детектор должен обладать:

- а) большим откликом на входное воздействие оптического сигнала;
- б) низким уровнем собственных шумов;
- в) широкой полосой пропускания, согласованной со спектром входного сигнала;
- г) линейностью характеристик.

Эти требования противоречивы и их невозможно в максимальной степени реализовать в одном приборе, поэтому невозможно создать один универсальный фотоприемник.

Фотоприемник является первым и основным элементом системы демодуляции и обработки оптического сигнала. Как и в радиоэлектронике, системы *детектирования* в оптической электронике разделяют на две группы: а) непосредственного детектирования и б) детектирования с преобразованием. Последняя аналогична гетеродинному приему. Она применима для детектирования лишь когерентного светового потока, который предварительно смешивается с когерентным излучением опорного сигнала, генерируемого лазером и выполняющим функцию местного гетеродина. Оптическое детектирование с преобразованием, как правило, осуществляется в длинноволновом оптическом диапазоне с $\lambda > 10$ мкм. Наиболее часто применяют схемы прямого детектирования.

Для определения технических возможностей конкретного фотоприемника наиболее важными являются следующие характеристики:

1. **Спектральная (монохроматическая) чувствительность** S_λ — мера реакции фотоприемника на оптическое излучение с длиной волны λ :

$$S_\lambda = \frac{dU}{d\Phi_\lambda}, [B / Bm]$$

$$S_\lambda = \frac{dI}{d\Phi_\lambda}, [A / Bm]$$

Для тепловых приемников S_λ не зависит от длины волны, а для фотонных приемников существует максимальная (пороговая) длина волны λ_m , выше которой энергии фотона $h\omega = hc/\lambda$ недостаточно для возникновения фотоэффекта.

2. **Интегральная чувствительность** S — мера реакции фотоприемника на световой поток Φ заданного спектрального состава

$$S = \frac{dU}{d\Phi}, [B / Bm]$$

$$S = \frac{dI}{d\Phi}, [A / Bm]$$

3. **Минимально различимый сигнал** Φ_{\min} — та величина светового потока, измеряемая в [Вт], которая на выходе фотоприемника создает сигнал, равный шуму.

4. **Эквивалентная мощность шума** NEP^* — та величина светового потока, которая на выходе фотоприемника в единичной полосе частот вызывает сигнал, равный шуму,

$$NEP^* = \frac{\Phi}{\Delta f} \cdot \frac{U_{ш}}{U_\phi}, [Bm / Гц^{1/2}]$$

7. **Инерционность** — способность фотоприемника без искажения регистрировать быстрые изменения интенсивности светового потока. Характеризуется граничной частотой λ_{\max} при которой чувствительность фотоприемника падает в заданное число раз (обычно в 2 или e раз), или постоянной времени (для линейных процессов).

Для детекторов, отклик которых имеет экспоненциальные законы нарастания и спада с постоянной времени τ , зависимость чувствительности S или S_λ от частоты модуляции света f имеет вид

$$S(f) = \frac{S(0)}{\sqrt{1 + 4\pi^2 f^2 \tau^2}}$$

Принцип действия **тепловых фотоприемников** основан на регистрации изменения свойств материала при изменении его температуры вследствие поглощения оптического излучения. Инерционность тепловых приемников велика (более 10 мс), а чувствительность сравни-

тельно низка $D^* = 10^8 \dots 10^{10} [см \cdot Гц^{1/2} / Bm]$. Поэтому в системах передачи информации они не используются. Тепловые приемники применяются там, где необходимо обеспечить постоянство спектральной чувствительности, а также в далекой ИК-области спектра.

Фотонные приемники эффективно работают в той области спектра, где энергия фотона существенно превышает kT . Принцип действия основан на использовании внешнего или внутреннего фотоэффектов. Они обладают малой инерционностью, большой чувствительностью и высокой обнаружительной способностью. В случае, когда тепловая энергия сравнима или превышает энергию фотона $h\omega$, тепловое возбуждение действует активнее оптического и эффективность фотонного приемника резко падает. Поэтому фотонные приемники, предназначенные для работы в области $\lambda > 3$ мкм, как правило, требуют охлаждения тем более глубокого, чем больше рабочая длина волны.

Световые сигналы в оптической электронике, можно разделить на два вида: 1) оптические сигналы, развивающиеся только во времени; 2) многомерные оптические сигналы, представляющие световые картины. В соответствии с этим фотоприемники можно разбить на две группы.

1. **Дискретные**, как правило, одноэлементные фотоприемники с малой рабочей площадью, предназначенные для приема коротких оптических импульсов, обладающие высокой спектральной чувствительностью S_λ в заданной области спектра, большой детектирующей способностью D^* и малой инерционностью τ . Лучшим сочетанием параметров в этой группе приборов обладают фотодиоды, особенно **p-i-n** и **лавинные фотодиоды**.

2. Фотоприемники, предназначенные для регистрации двухмерных световых сигналов. Как такие приборы представляют собой **многоэлементные фотоприемники** с самосканированием и высокой пространственной разрешающей способностью, обладающие хорошей чувствительностью в сравнительно широком спектральном интервале. Лучшими характеристиками из этой группы приемников обладают фоточувствительные приборы с зарядовой связью.

6.4.2. Принцип работы PIN-фотодиода.

Структура и принцип действия pin-фотодиода пояснены на рисунке 2. Он состоит из низкоомной p^+ -подложки, слабо легированного (собственного) i -слоя и тонкого низкоомного p^+ -слоя толщиной до 0,3 мкм, через которые производится освещение. Низкоомные p^+ и p^+ -области выполняют роль контактов. Наличие центрального высокоомного i -слоя приводит к увеличению ширины слоя объемного заряда по сравнению с обычным p-n-переходом. Его толщина d_i , подбирается так, чтобы поглощение света происходило в этом i -слое, совпадающем со слоем объемного заряда. При приложении обратного смещения U обедненный слой распространяется на всю i -область. Это приводит к уменьшению емкости перехода, повыше-

нию чувствительности и быстродействия.

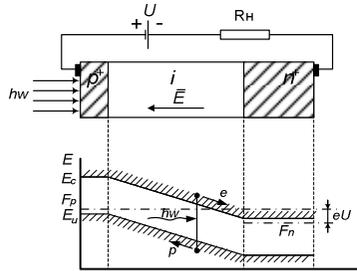


Рис.1 – структура и зонная диаграмма p-i-n-фотодиода

Падающий свет, затухая по экспоненциальному закону с постоянной, определяемой показателем поглощения κ_{ω} для данной длины волны, вызывает генерацию носителей заряда преимущественно в i-слое. Фотогенерированные носители ускоряются электрическим полем до скорости насыщения дрейфа ($\sim 10^5$ м/с), поскольку напряженность электрического поля в обедненном слое обычно превышает 1 кВ/см. Эта скорость дрейфа примерно на три порядка превышает скорость диффузии. Поэтому p-i-n фотодиод конструктивно выполняется так, чтобы максимально уменьшить долю поглощенного света вне i-слоя. С этой целью переход формируется у самой поверхности кристалла.

Фотодиоды рассматриваемого типа могут работать в двух режимах: в вентильном (без приложения внешнего напряжения к p-n переходу – область 1 на рис.1) и в фотодиодном режиме (при приложении напряжения в запирающем направлении – область 2 на рис.1). Фотодиодный режим используется в p-i-n-фотодиодах.

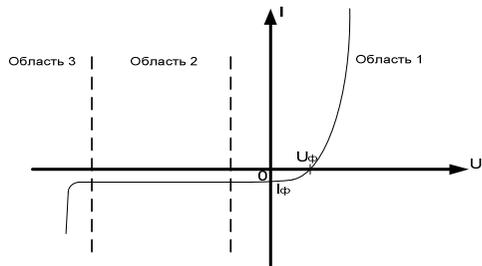


Рис.2 Вольт-амперная характеристика p-n перехода.

Из рис. 2 видно, что фотодиод может по-разному использоваться для детектирования оптического излучения. В простейшем случае (область 1) диод непосредственно подключается ко входу усилителя напряжения с высоким входным сопротивлением, который измеряет изменение $U\phi$ (см. рис. 2). В другом случае ток диода усиливается усилителем тока, имеющим низ-

кое входное сопротивление, т. е. напряжение на диоде поддерживается вблизи нуля. При этом оказываются весьма малыми шумы диодного тока. На практике, однако, фотодиоды в системах оптической связи почти всегда работают в фотодиодном режиме. При этом в p-n переходе существует зона, в которой нет свободных носителей заряда (обедненная зона) и в которой все определяется электрическим полем. Каждая возникающая в этой зоне пара электрон – дырка разделяется и дает вклад в ток, когда электрон и дырка выходят за границы обедненной области.

Уменьшив степень легирования слоя n-типа, можно увеличить ширину слоя объемного заряда при том же рабочем напряжении. В пределе мы получаем беспримесный материал с собственной проводимостью (обозначаемый буквой «i»), к которому добавляется слой материала n-типа с малым удельным сопротивлением для обеспечения омического контакта. Такая структура p-i-n, обеспечивающая квантовый выход, близкий к единице, и очень высокую чувствительность.

6.4.3. Лавинные фотодиоды

Лавинные фотодиоды (avalanche photodiodes APDs) внутренне увеличивает первичный фототок сигнала прежде, чем он входит во входную схему следующего усилителя. Это увеличивает чувствительность приемника, так как фототок усилен перед столкновением с тепловым шумом, связанным с схемой приемника. Для увеличения носителей, фотопроизведенные носители должны пересечь область, где присутствует очень высокое электрическое поле. В этой

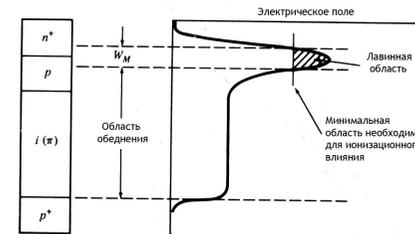


Рис. 1. Устройство и структура поля в области умножения лавинного фотодиода.



Рис. 2. Лавинный фотодиод

высоко-полевой области фотопроизведенный электрон или дырка могут получить достаточно энергии, чтобы это ионизировало связанные электроны в зоне валентности после столкновения с ними. Этот механизм увеличения носителей известен как *ионизация воздействия*. Недавно созданные носители также ускорены высоким электрическим полем, таким образом получая достаточно энергии причинить далее ионизацию воздействия. Это явление - *эффект лавины*. Ниже напряжения пробоя диода общее количество порожденных носителей ко-

нечно, тогда как выше пробоя их число может быть бесконечно. Обычно эта структура используется для того, чтобы достигать увеличения носителей с очень небольшим дополнительным шумом – *пробой базы* - через устройство показано на рис. 1. Пробой базы - через лавинный фотодиод (RAPD) составлена из материала р-типа высокого удельного сопротивления, нанесённого как эпитаксиальный слой на $p +$ (тяжелые примеси р-типа) подложку. Диффузия р-типа или инжекция ионов производят в материале высокого удельного сопротивления, сопровождаемом наращиванием $n +$ (тяжелых примесей n-типа) слой. Для кремния обычно формируются легирующие примеси, эти слои - обычно бор и фосфор. Эта конфигурация упоминается как $p+\pi rp +$ - через устройство. π слой это основной материал, который имеет p примесь из-за несовершенной очистки.

Термин "пробой базы" является результатом действия фотодиода. Когда приложено обратное напряжение смещения, большая его часть падение напряжения приложенного параллельно pn переходу.

Слой истощения расширяется с увеличивающимся смещения, пока не будет достигнуто некоторое значение напряжения, в некоторой пиковой электрической области в $pn +$ переходе будет приблизительно на 5 - 10 процентов ниже этого, должен произойти лавинный пробой. В этой точке слой истощения только "пробой базы" к почти свойственной области π .

При нормальном использовании RAPD полностью работает в режиме истощения. Свет попадает в устройство через $p +$ область и поглощается в π , который действует как область концентрации для фотопроизведенных носителей. Будучи поглощенным фотон отдаёт свою энергию, таким образом создавая электронно-дырочные пары, которые отделены электрическим полем в области π . Фотопроизведенные электроны дрейфуют через область π к $pn +$ переходу, где существует высокое электрическое поле. Именно в этой высоко-полевой области происходит умножение носителей.

Среднее число электронно-дырочных пар, созданных носителем на расстояние единицы длины хода, называется *коэффициентом ионизации*. Большинство материалов показывает различные *коэффициенты ионизации электрона α* и *коэффициенты ионизации дырки β* . Экспериментально полученные величины α и β для пяти различных полупроводниковых материалов, которые показаны на рис. 6-6. Отношение $k = \beta/\alpha$ двух коэффициентов ионизации - мера работы фотоприёмника. Поскольку мы будем рассматривать в главе. 6.4, лавинные фотодиоды состоящие из материалов, в которых тип носителей доминирует в значительной степени, при этом ударная ионизация обеспечивает малый шум и большую эффективность усиления. Как показано на рис. 6-6, всех исследованных материалов пока только кремний имеет существенное различие между электроном и коэффициентом ионизации дырки. Увеличение M для всех носителей, произведенных в фотодиоде определяется как:

$$M = \frac{I_m}{I_p} \quad (1)$$

где I_m - средняя величина общего количества усиленного выходного тока, и I_p - первичный немноженный фототок, определенный по формуле. (1). Практически, механизм лавины - статистический процесс, не каждая пара носителей, произведенная в диоде испытывает то же самое умножение. Таким образом, взвешенная величина M . выражена как среднее количество.

Аналогично пин-фотодиоду, характеристики APD охарактеризованы его чувствительностью R_{APD} , которым дают

$$R_{APD} = \frac{\eta q}{h\nu} M = R_0 M \quad (2)$$

где R_0 – это единица чувствительности одного APD (усилителя).

6.5. Оптические волокна

6.5.1. Конструкция оптических волокон

Наиболее широко распространенные оптические волноводы изготавливаются либо в виде оптического волокна из кварцевого стекла, либо путем создания в кристалле волноводного слоя или канала (рис. 1), в котором показатель преломления n_1 выше, чем в окружающем этот канал материале n_2 . Последний вариант мы не будем рассматривать, так как кристаллические волноводы имеют относительно высокие потери и из них нельзя изготовить длинных волноводов. Обычно длина таких волноводов не превышает нескольких сантиметров, и они используются в устройствах интегральной оптики в качестве модуляторов, переключателей света и т.п.

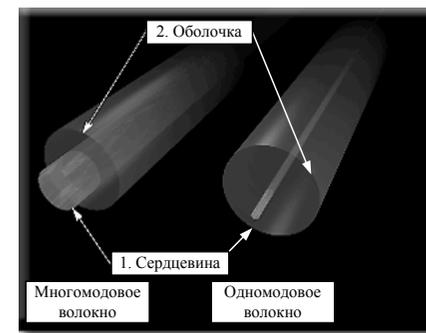


Рис. 1. Структура оптических многомодовых и одномодовых волокон

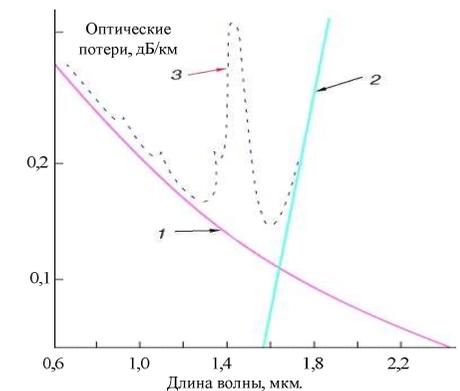


Рис. 2. Качественная схема зависимости оптических потерь от длины волны света.

Вернемся теперь к стекловолкну. Типичный стекловолноконный волновод состоит из

трех компонентов (слоев). Первый слой — это сердцевина, то есть центральная световедущая жила из кварцевого стекла. Характерный радиус сердцевины 2 – 5 мкм в так называемых одномодовых волокнах и 25 – 30 мкм в многомодовых волокнах. Затем идет оболочка. Оболочка также является стеклянной, характерное значение ее наружного радиуса 50 – 70 мкм. Затем идет защитная оболочка из полимерного материала, которая может быть многослойной. Наружный диаметр волокна с защитным покрытием составляет 0,5 – 1 мм.

Принципиально важным является то, что показатель преломления сердцевины должен быть несколько выше, чем у оболочки. Разница $n_1 - n_2$ обычно порядка одного процента. Такие волокна называют *слабонаправляющими*. Превышение n_1 над n_2 необходимо для того, чтобы обеспечить полное внутреннее отражение света на границе сердцевина—оболочка.

6.5.2. Изготовление оптических волокон

Существуют разные методы изготовления стекловолокна. Одной из наиболее эффективных и распространенных является технология изготовления волокна из кварцевого стекла по методу химического осаждения из газовой фазы. При изготовлении необходимо решить по крайней мере две центральные задачи. Во-первых, обеспечить чрезвычайно высокую химическую чистоту материала, составляющего сердцевину, и, во-вторых, обеспечить высокую однородность вытягиваемого волокна. Делается это следующим образом. Сначала изготавливают заготовку — стеклянный стержень диаметром 5 – 15 мм, имеющий распределение показателя преломления в поперечном сечении такое же, как у будущего световода. Для получения заготовки берут вначале кварцевую трубку из высококачественного стекла. Через трубку продувают смесь газообразного тетрахлорида кремния SiCl_4 и кислорода, а трубку прогревают газовой горелкой до высоких температур (вплоть до 1500°C). Тогда в результате химической реакции на внутренней поверхности трубки осаждается чистый кварц SiO_2 . Высокая степень химической чистоты этого кварца обеспечивается высокой чистотой газовых компонент SiCl_4 и O_2 . Полученный тонкий слой будет служить составной частью в будущей оболочке. Затем процесс нанесения слоев на внутреннюю поверхность повторяется, но в состав газовой смеси добавляется еще тетрахлорид германия GeCl_4 . Тогда в осажденном слое кварца содержится некоторое количество германия, что обеспечивает более высокий показатель преломления в этом слое, который служит основой будущей сердцевины.

Затем полую стеклянную трубку с нанесенными на внутреннюю поверхность слоями кварца прогревают до размягчения, и она за счет сил поверхностного натяжения сжимается в стержень. Заготовка готова. Для того чтобы из нее получить волокно, заготовку помещают в устройство для вытяжки, где она устанавливается вертикально, и нижний конец нагревают до плавления. За этот конец производится вытяжка. Для обеспечения нужной толщины произво-

дится регулировка скорости вытяжки и температуры нагревания. Правильный подбор и поддержание нужной скорости вытяжки важны для получения волокна высокой однородности. Вытягиваемое волокно покрывают защитной оболочкой и наматывают на барабан. Если в процессе изготовления заготовки производится нагрев до более высоких температур и выбирается более высокое давление газовой фазы, так что химическая реакция происходит не только на внутренней поверхности трубки, но и в газовом потоке, то указанный метод называют модифицированным способом осаждения из газовой фазы. Если слои с примесью германия наносятся так, что показатель преломления n_1 в сердцевине постоянен по сечению сердцевины, то тогда распределение показателя преломления в волокне называется ступенчатым. Если концентрация германия изменяется так, что показатель преломления сердцевины плавно уменьшается вдоль радиуса, проведенного от центральной оси к оболочке, то такое волокно называется градиентным

Из отдельных волокон могут быть изготовлены кабели различного профиля, содержащие до нескольких десятков отдельных волокон.

6.5.3. Оптические потери в световолокне

Имеется несколько источников затухания света при его распространении по волокну. Некоторые из них устранимы, а некоторые нет. Волокна низкого качества содержат примеси, которые поглощают свет, а также дефекты (мелкие пузырьки, трещины, неоднородности толщины и т.п.), которые вызывают рассеяние света. Технологии изготовления волокна, существовавшие до конца шестидесятых годов, не обеспечивали изготовление волокна высокого качества и потери составляли порядка 100 – 1000 дБ/км. В таких волокнах свет затухает в два раза на расстоянии порядка десятка метров. В шестидесятые годы ученые предсказывали, что для применения световолокон для целей связи необходимо, чтобы затухание не превосходило 20 дБ/км. В 1970 году компания “Корнинг гласс компани” сумела изготовить волокно с такими потерями, после чего интерес к оптическим волокнам резко возрос. Чтобы устранить поглощение, обусловленное примесями, необходимо иметь концентрацию примесей ионов Zd -группы (железо, хром, марганец, кобальт и др.) в количестве менее 1 атома на 10^9 атомов основного материала. Технология осаждения из газовой фазы тем и хороша, что газовая фаза может быть очищена от примесей в нужной степени. Следующая примесь, которая играет исключительно важную роль, – это пары воды, точнее, гидроксильная группа OH^- . Колебания атомов O-H вызывают сильное поглощение не только на основной частоте, соответствующей длине волны 2,73 мкм, но и на комбинационных частотах вблизи 1,2 – 1,4 мкм.

Необходимо обеспечивать очень хорошее обезвоживание материала при изготовлении волокна (концентрация OH^- менее 10^{-8} [2]), чтобы обеспечить потери менее 1 дБ на длине волны 1,5 мкм. Более того, даже готовое волокно требует дополнительной защиты от проникно-

вения молекул воды, так как за длительный период эксплуатации пары воды могут за счет диффузии проникать в оболочку и сердцевину. Несмотря на столь жесткие требования, успехи в технологии изготовления волокон таковы, что в лучших образцах потери составляют около 0,2 дБ/км

Но можно ли создать кварцевое волокно с еще меньшими потерями? Практически нельзя, так как указанная величина потерь близка к теоретическому пределу, обусловленному уже фундаментальными, а не технологическими причинами. Наиболее важными фундаментальными причинами затухания света являются рэлеевское рассеяние и поглощение света атомами, из которых состоит волокно. Рэлеевское рассеяние представляет собой рассеяние света на микронеоднородностях стекла, размер которых меньше длины волны света. Микронеоднородности возникают из-за того, что при температурах изготовления волокна (выше точки плавления) неизбежно существуют неоднородности плотности и химического состава вещества за счет хаотического движения молекул при высокой температуре. И эти флуктуации тем больше, чем выше температура. При охлаждении ниже точки плавления (для кварца это 1940°C) эти микронеоднородности как бы замораживаются [3]. Наличие неоднородностей плотности вещества и его состава приводит к неоднородностям показателя преломления, на которых и происходит рассеяние света. Потери за счет рэлеевского рассеяния обратно пропорциональны длине волны в четвертой степени. Поэтому они наиболее существенны для коротких длин волн (на рис. 2 это длины волн короче 1,5 мкм). Для больших длин волн более существенную роль играют потери за счет поглощения света при колебаниях атомов, из которых состоит кварц. Хотя фундаментальные частоты колебаний пар атомов Si–O, Ge–O лежат в области 10 мкм, тем не менее края полос поглощения простираются до единиц микрон. В результате комбинации рэлеевского рассеяния и инфракрасного поглощения имеет место абсолютный минимум потерь вблизи 1,5 мкм.

Заметим, что этот минимум характерен для кварцевых волокон. Для волокон другого состава зависимость потерь от длины волны света будет иной.

6.5.4. Волноводные моды в оптическом волокне

Если пользоваться представлениями и терминами геометрической оптики, то в основе распространения света по волокну лежит эффект полного внутреннего отражения (рис. 3). Этот эффект наблюдается тогда, когда свет падает из более плотной оптической среды (то есть среды, обладающей большим показателем преломления) на оптически менее плотную среду ($n_1 > n_2$). Тогда при углах θ , больших некоторого критического угла θ_c , который определяется из условия

$$\sin \theta_c = n_2 / n_1, \quad (1)$$

свет не проникает во вторую среду, а полностью отражается от границы раздела. Заметим, что

более детальное рассмотрение показывает, что в некоторой степени свет во вторую среду проникает, но так, что его амплитуда очень быстро (по экспоненциальному закону) падает от границы раздела. К этому вопросу мы еще вернемся в дальнейшем.

Если слой с показателем преломления n_1 находится между двумя слоями с n_2 , то свет будет распространяться только внутри центрального слоя с n_1 . Это и есть пример оптического волновода. Такой волновод обладает очень интересным свойством. Оказывается, что распространяющимися в нем будут только такие лучи, которые падают под определенными (дискретными) значениями угла θ . Связано это вот с чем. Вспомним, что луч света - это световая волна. Если в волновод, изображенный на рис. 3б, вводится свет в виде плоской волны, то в результате отражений от верхней и нижней границы свет в слое l можно считать суперпозицией (наложением) двух волн: одна распространяется снизу вверх, а другая сверху вниз. Между этими волнами возникает интерференция, то есть в некоторых точках колебания электрического поля складываются, а в некоторых вычитаются. Может случиться так, что одна волна полностью погасит другую. Тогда свет вообще не будет распространяться вдоль волновода. Однако при специальных углах θ накладывающиеся волны интерferируют конструктивно, не гасят друг друга, и поэтому свет беспрепятственно распространяется по слою l . В строгой теории при решении уравнений Максвелла для диэлектрического волновода каждому такому углу соответствует собственное решение. Это решение называется собственной модой (или просто модой) данного волновода. Каждой моде соответствует свое распределение поля по сечению волновода. Иногда оно имеет простой вид, например, изменяется как функция $\sin K_m x$ или $\cos K_m x$, где K_m - соответствующий волновой вектор для моды с "номером" m , иногда весьма сложный, если в волноводе может распространяться много мод и сам волновод не плоский. Каждая мода распространяется со своей скоростью.

Если волновод идеальный, то моды не взаимодействуют между собой, то есть свет, сосредоточенный в одной моде, не переходит в другую. Но если создать какие-либо дефекты или неоднородности (например, изгибы, скрутки волокна), то моды будут взаимодействовать между собой.

Число распространяющихся мод N при данной длине волны ограничено и определяется размерами волновода (радиусом сердцевины R_c), длиной волны l и разностью $n_1 - n_2$.

Для круглого волокна со ступенчатым распределением

$$N = \frac{2\pi^2 R_c^2}{\lambda^2} (n_1^2 - n_2^2). \quad (2)$$

Полезно упомянуть также о таком параметре волновода, как числовая апертура NA , который указывает, какой максимальный угол ввода света из воздуха в волновод возможен для возбуждения распространяющихся мод (рис. 3):

$$NA = \sin \alpha = \sqrt{n_1^2 - n_2^2} \quad (3)$$

Представление поля в волноводе в виде взаимодействующих мод дает очень удобную возможность расчета поля на выходе волокна. Если есть распределение поля на входе, то его можно представить как суперпозицию мод волновода, затем учесть, какие разности фаз возникают на выходе за счет разной скорости распространения, и затем снова просуммировать распределение полей разных мод на выходе. Однако для точных расчетов удобно пользоваться понятием мод только для волноводов с относительно небольшим числом мод. В случае большого числа мод анализ методами геометрической оптики (решение уравнения Эйконала) более удобен.

Какими свойствами обладают моды в световолокне? Рассмотрим сначала одномодовое волокно. Стеклоподобный, как и любой другой диэлектрический волновод, обладает замечательным свойством. Формально в нем всегда, то есть при любой длине волны, может распространяться хотя бы одна мода. Эту моду называют фундаментальной. Условием, при котором в волокне распространяется только фундаментальная мода, является требование, чтобы так называемая нормализованная частота V была меньше 2,405.

Нормализованная частота

$$V = \frac{2\pi R_c}{\lambda} \sqrt{n_1^2 - n_2^2} \quad (4)$$

Профиль поля для фундаментальной моды по сечению волокна имеет максимум в центре, а затем поле экспоненциально спадает по радиусу, проникая и в оболочку.

Если обеспечить такие условия, что V становится равной или несколько большей 2,405 (например, уменьшая длину волны света, вводимого в волокно), то появляются условия для возбуждения второй моды. При дальнейшем уменьшении длины волны будут возбуждаться моды более высокого порядка.

В многомодовом волокне могут существовать моды, глубоко проникающие в оболочку, а также собственные моды оболочки. Эти моды, хотя и являются разновидностью распространяющихся мод, достаточно сильно затухают. Их называют модами утечки. При наличии макронеоднородностей в волокне (изгибы, скрутки) слабозатухающие распространяющиеся моды могут взаимодействовать с модами утечки, отдавая им часть энергии, что повышает дополнительные потери для распространяющихся мод.

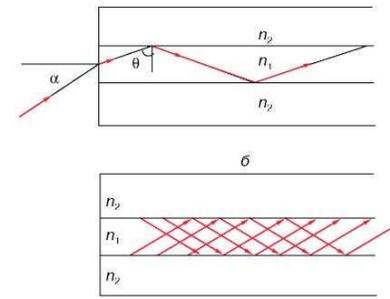


Рис. 3. Распространение света в плоском волноводе.

a – Траектория луча света при полном внутреннем отражении.
б – суперпозиция плоских волн в волноводе

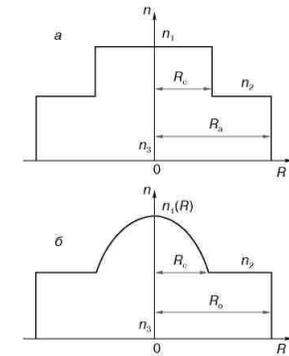


Рис. 4. Распределение показателя преломления по сечению волокна.

a – Ступенчатое волокно
 R_c – радиус сердцевины R_0 – радиус оболочки;
б – градиентное волокно.

Как уже упоминалось, каждая мода в волноводе имеет свою скорость распространения и соответственно свой набег, то есть приращение фазы в зависимости от длины пути распространения. Это приводит, в частности, к тому, что картина распределения поля, которая была на входе многомодового волокна, не сохраняется. Например, если на входе свет был сфокусирован в точку в центре сердцевины, то на выходе достаточно длинного многомодового волокна мы увидим сложную пятнистую картину. Однако имеется один замечательный тип волокна, где можно обеспечить равенство (с некоторой ограниченной точностью, конечно) скоростей распространения мод. Это так называемые градиентные волокна (рис. 4). В градиентном волокне показатель преломления меняется не ступенькой при переходе от сердцевины к оболочке, а плавно уменьшается от центра сердцевины к границе раздела. Оказывается, что можно подобрать профиль изменения показателя преломления так, что удастся выровнять скорости мод распространения в волокне. Качественно этот эффект можно пояснить следующим образом. Если луч света падает под большим углом к оси волновода, то он отражается от слоя, который находится ближе к оболочке. Луч, падающий под меньшим углом, отражается от слоя, расположенного ближе к оси. Тогда первый луч имеет большую геометрическую длину пути распространения, но зато проводит больше времени в области, где показатель преломления меньше. В итоге этот луч имеет такую же оптическую длину распространения, как и второй луч, у которого геометрический путь короче, но зато он распространяется в области с большим показателем преломления. Хотя время распространения от входа до выхода для таких лучей не идеально точно совпадает, тем не менее это свойство выравнивания оптических путей или времени распространения в определенном виде градиентных волокнах чрезвычайно по-

лезно для оптических систем связи.

6.5.5. Дисперсия групповых скоростей

Информация по волоконным линиям связи передается в основном в виде коротких световых импульсов (импульсно-кодовая модуляция). Очевидно, что чем короче импульс, тем больше импульсов в единицу времени можно передать, то есть тем выше будет пропускная способность линии. Но чем короче импульс, тем шире его спектр. Ширина спектра импульса $\Delta f \approx 1/t_u$, где t_u — длительность импульса

Потенциально световой луч может иметь ширину спектра, сравнимую с его несущей частотой, то есть t_u могут быть порядка $10^{-13} - 10^{-14}$ секунд. Вопрос в том, можно ли такие короткие импульсы использовать на практике для передачи информации.

Трудности здесь возникают разного характера. Во-первых, как создать такие короткие импульсы и обеспечить их кодовую модуляцию; во-вторых, как зарегистрировать эти импульсы на выходе и, в-третьих, может ли волокно пропускать такие импульсы без искажений, то есть без потери информации? Остановимся только на последнем вопросе.

Световой импульс распространяется со скоростью, которая называется групповой скоростью V_g . Напомним, что в соответствии с теорией Фурье любой, в том числе и световой, импульс может быть представлен в виде совокупности (суперпозиции) большого числа простых волн с различными частотами. Общий спектр этих частот приблизительно Δf . В зависимости от свойств среды, в которой распространяется импульс, скорости распространения этих элементарных волн могут быть равны между собой (например, в вакууме), а могут отличаться в зависимости от частоты каждой волны. Скорость распространения импульса, то есть групповая скорость, может не совпадать со скоростью распространения элементарных волн. Более того, если скорости элементарных волн не равны между собой, то это приводит к искажению формы импульса и его уширению. Круг явлений, связанных с изменением формы импульса вследствие зависимости скорости распространения элементарных компонент от частоты, называют дисперсией групповых скоростей (ДГС). Количественно ДГС определяется производной $dV_g/d\omega$, где ω — частота света. Экспериментально ДГС можно характеризовать параметром

$$D = dt_g/d\lambda$$

Здесь $t_g = 1/V_g$ — это так называемое групповое время, то есть время, за которое импульс проходит единицу длины, например один километр. Величина D тогда характеризует, как изменится t_g , если изменить длину волны на $d\lambda$.

Поскольку размерность $[D]$ есть пикосекунда/(нанометр \times километр), то для наглядности этот параметр можно интерпретировать как величину уширения импульса на D пикосекунд, если импульс проходит длину пути в один километр, а спектр импульса в единицах длин волн равен одному нанометру (этому спектру соответствует длительность импульса порядка единиц пикосекунд, то есть 10^{-12} с, при длине волны в 1 мкм).

Эффект ДГС играет важную и крайне отрицательную роль в системах связи (хотя в других случаях эффект может быть полезен). Рассмотрим основные источники ДГС. Первый, наиболее важный источник — это межмодовая дисперсия. Как уже упоминалось, в многомодовом волокне со ступенчатым показателем преломления скорости распространения мод различны. Отсюда следует, что если на входе импульс света возбudit много мод, то сигнал разобьется на такое же число импульсов, сколько возбуждено мод, и каждый импульс будет двигаться со своей групповой скоростью, присущей данной моде. Естественно, что после прохождения некоторого расстояния одни импульсы уйдут вперед, другие отстанут, и в целом произойдет уширение исходного импульса. Межмодовая дисперсия настолько велика, что если взять ступенчатое волокно с разумным $\Delta n = n_1 - n_2 \sim 0,01$, то разброс групповых скоростей или времен для разных мод будет также порядка одного процента. Это гигантская величина, из которой следует, что импульс с длительностью в 10^{-9} с (1 наносекунда) уширится вдвое уже на расстояниях порядка 30 м. Отсюда следует, что многомодовые волокна со ступенчатым показателем преломления могут использоваться только на коротких расстояниях и для передачи достаточно длинных (по оптическим стандартам) импульсов порядка микро- или наносекунд.

Значительно лучше обстоит дело у градиентных волокон, поскольку в них скорости распространения разных мод приблизительно равны. Межмодовая дисперсия в градиентных волокнах может быть в десять – сто раз меньше. Соответственно, длины линий связи могут достигать километров при длительностях импульса порядка наносекунд.

Радикальным решением проблемы межмодовой дисперсии является переход к одномодовому волокну. Поскольку здесь отсутствует межмодовая дисперсия, то наиболее существенными оказываются два других типа дисперсии: материальная и внутримодовая. Материальная дисперсия определяется зависимостью показателя преломления от частоты света и прямо связана с химическим составом материала (в данном случае кварцевого стекла). Величина $|D| = 0 - 15$ пс/(км \times нм) в интервале длин волн между 1,2 и 1,5 мкм.

Внутримодовая дисперсия возникает из-за того, что одна и та же мода, но возбужденная на разных частотах света, будет иметь несколько разную скорость распространения. Обычно материальная дисперсия больше внутримодовой, но в кварцевом волокне вблизи длины волны в 1,3 мкм материальная дисперсия также мала и даже меняет знак, проходя через нулевое значение. Вблизи этой длины волны возможна взаимная компенсация обоих видов дисперсии. Наличие области длин волн, где дисперсия близка к нулю, делает очень привлекательной дли-

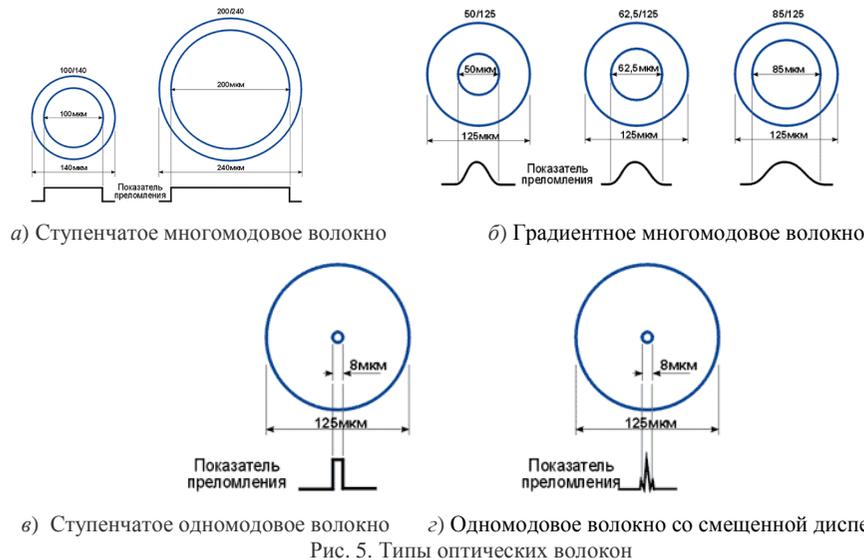
ну волны в 1,3 мкм для длинных линий связи, где расстояние между ретрансляторами составляет десятки, а в будущем, возможно, и более сотни километров.

6.5.6. Типы оптических волокон

Все оптические волокна делятся на две основные группы: многомодовые MMF (multi mode fiber) и одномодовые SMF (single mode fiber).

Многомодовые волокна подразделяются на ступенчатые (step index multi mode fiber) и градиентные (graded index multi mode fiber).

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 fiber)



а) Ступенчатое многомодовое волокно б) Градиентное многомодовое волокно
в) Ступенчатое одномодовое волокно г) Одномодовое волокно со смещенной дисперсией
Рис. 5. Типы оптических волокон

Типы и размеры волокон приведены на рис. 5. 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

Ступенчатое одномодовое волокно. В ступенчатом одномодовом волокне (SF) диаметр светонесущей жилы составляет 8-10 мкм и сравним с длиной световой волны. В таком волокне при достаточно большой длине волны света $\lambda > \lambda_{CF}$ (λ_{CF} - длина волны отсечки) 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.

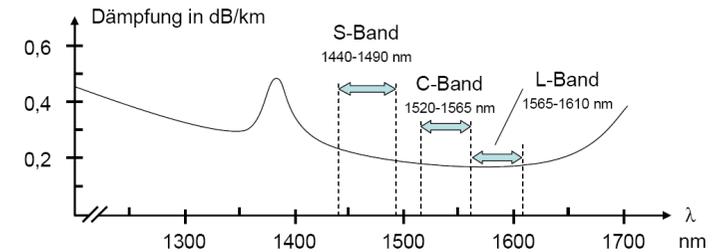


Рис. 6. Дисперсия коэффициента затухания света в кварцевых волокнах

Одномодовое волокно со смещенной дисперсией. В одномодовом волокне со смещенной дисперсией (DSF) длина волны, на которой результирующая дисперсия обращается в ноль, - длина волны нулевой дисперсии λ_0 - смещена в окно 1550 нм. Такое смещение достигается благодаря специальному профилю показателя преломления волокна, рис. 5 г. Таким образом, в волокне со смещенной дисперсией реализуются наилучшие характеристики как по минимуму дисперсии, так и по минимуму потерь. Поэтому такое волокно лучше подходит для строительства протяженных сегментов с расстоянием между ретрансляторами до 100 и более км. Разумеется, единственная рабочая длина волны берется близкой к 1550 нм. **Одномодовое волокно с ненулевой смещенной дисперсией.** Одномодовое волокно с ненулевой смещенной дисперсией NZDSF в отличие от DSF оптимизировано для передачи не одной длины волны, а сразу нескольких длин волн (мультиплексного волнового сигнала) и наиболее эффективно может использоваться при построении магистралей "полностью оптических сетей" - сетей,

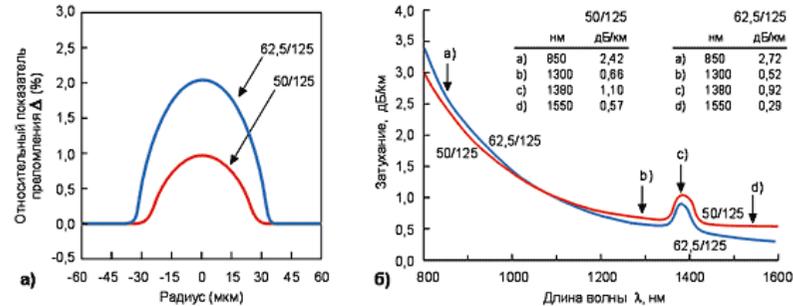
на узлах которых не происходит оптоэлектронного преобразования при распространении оптического сигнала.

- В одномодовых оптических волокнах (ОВ):

$$\Delta t = D \cdot L \cdot \Delta\lambda, \tag{1}$$

где Δt величина дисперсионное уширение импульса, $D \left[\frac{нс}{км \cdot нм} \right]$ - дисперсионный коэффициент ОВ; L – длина линии связи;

где $\Delta\lambda$ – ширина спектра оптического сигнала.



а) профили показателей преломления волокон 50/125 и 62,5/125;

б) характерные кривые спектральных потерь мощности

Рис.7. Многомодовые градиентные волокна

- Для ОВ со ступенчатым профилем:

$$\Delta t = \frac{L \cdot \xi \cdot \Delta n}{c}, \tag{2}$$

где Δn - разность показателей преломления сердцевины n_0 и оболочки n_1 ОВ:

$$\Delta n = n_0 - n_1 = \frac{NA^2}{2n_0}, \tag{3}$$

NA – числовая апертура ОВ;

ξ - параметр связи мод (если связи нет, то $\xi = 1$; при полной связи $\xi = \frac{1}{2}$).

- Для градиентного ОВ:

$$\Delta t = \frac{L}{2c} \cdot \frac{\Delta n^2}{n_1}. \tag{4}$$

- При эффективности ввода излучения в коннекторе многомодового ОВ $\eta < 1$ в системе возникает модовый шум (МШ), характеризуемый относительной флуктуацией мощности сигнала P_c :

$$\Delta P/P_c \approx (1-\eta)/\eta M,$$

где M – число мод в ОВ. В волокне со ступенчатым профилем число M определяется значениями показателей преломления сердцевины и оболочки ОВ n_0, n_1 , а также радиусом ОВ:

$$M \approx V^2 / 2,$$

где параметр V , называемый нормированной частотой ОВ, определяется как [7]:

$$V = \frac{2\pi}{\lambda} r \sqrt{n_0^2 - n_1^2}$$

7. СИСТЕМЫ ПРИЕМА И ПЕРЕДАЧИ ИНФОРМАЦИИ

7.1. Обобщенная схема аппаратуры передачи цифровой информации

Обобщенная структурная схема системы передачи цифровых данных (СПД) одностороннего действия (рис. 1.1) состоит из оконечного оборудования, выполняющего функции отправителя (ООД—ОС) и получателя (ООД—ПС) сообщений, устройства защиты от ошибок (УЗО), устройства преобразования сигналов (УПС) и канала связи.

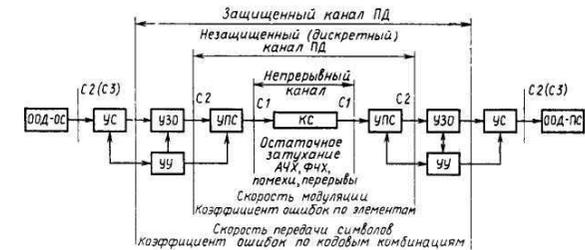


Рис. 1.1. Структурная схема системы аппаратуры передачи данных (СПД)

Рис.1. Структурная схема СПД

В УЗО происходит кодирование — декодирование данных. УПС осуществляет преобразование сигналов данных в форму, удобную для передачи по каналу связи, а устройства сопряжения (УС) обеспечивают обмен информационными и управляющими сигналами между АПД и ООД. Координация взаимодействия составных частей СПД обеспечивается специальными импульсами, вырабатываемыми устройством управления (УУ). Совокупность непрерывного канала с включенными на его входе и выходе УПС называют незащищенным (ПК) или дискретным каналом передачи данных (ПД), а объединение НК с УЗО — защищенным от ошибок (ЗК) каналом ПД. Для обмена информацией между абонентами используются две односторонние СПД, которые конструктивно могут быть выполнены в виде одной дуплексной СПД.

Основными характеристиками, определяющими качество и эффективность передачи данных, являются надежность, скорость и верность передачи, а также удельная скорость. Не-

маловажным показателем является сложность аппаратурной реализации, определяющая ее стоимость. При выборе вариантов проектируемой аппаратуры необходимо остановиться на том, который при равенстве одного или нескольких показателей обеспечивает более высокие остальные. Однако в зависимости от назначения технических средств часто приходится ухудшать одни показатели (например, стоимость или скорость) с целью обеспечения более высоких других показателей.

Надежность АПД определяет ее способность работать без отказов при сохранении заданных параметров аппаратуры. Важнейшими показателями надежности являются среднее время наработки на один отказ T_H и коэффициент готовности K_r . Нарботка на отказ T_H представляет собой среднее значение длительности непрерывной работы аппаратуры между двумя отказами. Коэффициент готовности K_r характеризует вероятность того, что АПД будет находиться в исправном состоянии в произвольно выбранный момент времени

$$T_H = \left(\sum_{i=1}^n T_i \right) \quad (1)$$

где T_i — время безотказной (исправной) работы АПД между i -м и $(i+1)$ -м отказами; n — общее число отказов за время измерений; T_0 — время восстановления аппаратуры, определяемое как среднее время простоя, вызванное отысканием и устранением данного отказа.

На практике для АПД массового применения необходимо, чтобы время наработки на отказ составляло не менее 5000 ч, а коэффициент готовности каналов передачи данных без резервирования находился в пределах 0,9—0,98. Для АПД, применяемой в АСУ, K_r должен быть не менее 0,99.

Скорость передачи информации V равна количеству информации, передаваемой по каналу связи за единицу времени [бит/с]:

$$V = (\log_2 m_c) / \tau_0 \quad (2)$$

Где m_c — количество позиций сигнала, τ_0 — длительность единичного элемента сигнала.

Для двухпозиционных сигналов

$$V_2 = 1 / \tau_0 \quad (3)$$

где $(\tau_0)^{-1}$ — определяет количество элементов, передаваемых по каналу связи в секунду, и носит название **скорости модуляции** B (Бод). Таким образом, для двоичных систем скорости передачи информации и модуляции совпадают. Применение многопозиционных сигналов позволяет при одной и той же скорости модуляции повысить, по сравнению с двухпозиционными системами, скорость передачи.

При разработке АПД следует иметь в виду, что соответствующим ГОСТ 17422—72 установлен ряд скоростей передачи: для телеграфных каналов 50, 75, 100 и 200 бит/с; для кана-

лов тональной частоты 300(200), 600, 1200, 2400, 3600, 4800, 7200, 9600 бит/с; для широкополосных 6, 12, 24, 48, 72 и 96 кбит/с. При разработке аппаратуры передачи дискретной информации по физическим линиям необходимую скорость передачи также следует выбирать из приведенного ряда скоростей.

Верность передачи данных количественно оценивается вероятностями ошибочного приема единичных элементов P_0 и кодовой комбинации P_{kk} , которые определяются следующим образом:

$$P_0 = \lim(n_{ош} / n) \quad (4)$$

где $n_{ош}$, $N_{ош}$ — количество ошибочно принятых единичных элементов и кодовых комбинаций соответственно; n , N_n — количество переданных единичных элементов и кодовых комбинаций соответственно.

В связи с ограниченным числом n и N_n на практике вместо вероятностей P_0 и P_{kk} используют **коэффициенты ошибок по элементам** K_0 и по **кодовым комбинациям** K_{kk} :

$$K_0 = n_{ош} / n \quad (5)$$

Для телефонных каналов коэффициент K_0 нормируется рекомендацией МККТТ V53. Его величина зависит от типа канала, скорости модуляции. Значения K_0 приведены в табл. 1.

Коэффициент ошибки по кодовым комбинациям независимо от типа канала и скорости передачи должен быть не более $1 \cdot 10^{-6}$

Удельная скорость передачи γ_s характеризует эффективность использования канала связи и численно равна количеству передаваемых бит на 1 Гц полосы:

$$\gamma_s = V / \Delta F_k \quad (6)$$

где ΔF_k — эффективная полоса пропускания канала.

Таблица 1

Тип канала связи	Скорость модуляции, Бод	Вероятность (коэффициент) ошибки
Коммутируемый канал (КК)	300	$1 \cdot 10^{-4}$
	600	$1 \cdot 10^{-3}$
	1200	$1 \cdot 10^{-3}$
Некоммутируемый (выделенный) канал (НКК)	300; 600	$5 \cdot 10^{-3}$
	1200	$5 \cdot 10^{-3}$

Без применения амплитудно-фазовых корректоров γ_s , как правило, не превышает 0,39 бит/Гц, а с применением автоматических корректоров — 3,2 бит/Гц и выше. Повышение удельной скорости связано с увеличением сложности аппаратурной реализации и соответственно ее стоимости, поэтому при проектировании АПД для реальных условий необходимо учитывать объем передаваемых сообщений, стоимость аренды каналов и т. д. Задачей учебно-

го проектирования является разработка АПД, удовлетворяющей заданным требованиям при минимальной стоимости аппаратурной реализации.

7.2. Параметры и характеристики каналов связи

Для организации обмена и передачи данных в автоматизированных информационных системах наиболее широко в настоящее время используются стандартные телефонные каналы тональной частоты (ТЧ), а также кабельные линии городской телефонной сети (ГТС). Каналы тональной частоты предназначались для организации телефонной связи между абонентами одного города или различных населенных пунктов и не учитывали особенностей передачи дискретных сигналов. Параметры каналов, которые практически не оказывали влияния на качество передачи телефонных сообщений, заметно ухудшают качество передачи дискретной информации, особенно при передаче на скоростях следующими формулами соответственно:

$$P_H = 201 \lg(U_x / U_{исх}) \quad (1)$$

где P_x , U_x и I_x — мощность, напряжение и ток в измеряемой точке x ; $P_{исх}$, $U_{исх}$, $I_{исх}$ — мощность, напряжение и ток в точке канала, принятой за исходную.

Если в качестве исходных приняты $P_{исх}=1$ мВт; $U_{исх} = 0,775$ В; $I_{исх}=1,29$ мА, то такие уровни называют *абсолютными*. Напряжение и ток в этом случае определены для сопротивления $R_H = 600$ Ом. Если сопротивление нагрузки канала не равно 600 Ом, то исходные значения напряжения тока определяют по формулам:

$$U_{исх} = \sqrt{10^{-3} R_H} \quad (2)$$

$$I_{исх} = \sqrt{10^{-3} / R_H} \quad (3)$$

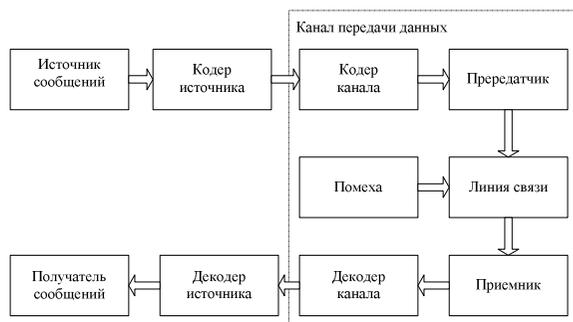


Рис. 1. Обобщенная структурная схема системы передачи дискретных сообщений

Измерительный уровень- абсолютный уровень в данной точке, когда уровень на входе канала равен 0 дБ. От относительных уровней можно легко перейти к абсолютным по формулам полученных из (1.7)

$$P_x = U_{исх} 10^{0,1P_M} \quad (4)$$

$$U_x = U_{исх} 10^{0,05P_H} \quad (5)$$

Величина $\alpha = 101 \lg(P_{вх}/P_x) = 201 \lg(U_{вх}/U_x)$ называется *затуханием* участка линии связи. Соотношение между уровнями на входе $p_{вх}$ и выходе p_v канала определяет его *остаточное затухание*:

$$\alpha_{ост} = p_{вх} - p_{вых} = 101 \lg(P_{вх} - P_{вых}) \quad (6)$$

Общее затухание кабельной линии связи длиной l_K равно $\alpha_l = \alpha_K l_K$, где α_l [дБ/км] - затухание единицы длины линии. Если канал связи состоит из i участков, вносящих затухание α_i и j участков с усилением S_j , то остаточное затухание всего канала вычисляется по формуле

$$a_{ост} = \sum_i a_i - \sum_j S_j \quad (7)$$

где $S_j = 101 \lg(P_i/P_{вх})$ — усиление i -го участка.

Обобщенная модель цифровой системы передачи (ЦСП) информации включает три фундаментальных процесса: кодирование-декодирование источника, кодирование-декодирование канала, модуляция-демодуляция при передаче по каналу (рис. 1). На передающей стороне все виды обработки информационных сообщений служат цели преобразования их в сигналы, наиболее подходящие для передачи по каналу конкретного типа. На приемной стороне производятся обратные операции, направленные на восстановление в исходном виде с минимально возможными искажениями. При этом искажения обусловлены либо неидеальностью процессов прямого - обратного преобразования, либо неидеальностью характеристик тракта (канала связи), включая воздействие помех.

Процесс кодирования источника имеет своей главной целью сокращение объема передаваемой информации, т.е. снижение требований к таким ресурсам системы, как время передачи, полоса пропускания, объем памяти при обработке или при хранении информации. Кодирование канала используется для исправления ошибок, возникающих при приеме цифрового сигнала из-за действия различных помех и искажений. В трактах вещания информации программных служб применяется только прямое исправление ошибок, а в обратных каналах интерактивных систем, особенно в телефонных каналах, возможно также использование перезапросов. В любом случае кодирование канала приводит к увеличению объема передаваемых данных, т.к. алгоритмы обнаружения и исправления ошибок требуют добавления специальных служебных символов, а повторы перезапрошенных блоков непосредственно увеличивают время передачи.

Модуляция используется для преобразования сигналов, представленных в основной (исходной) полосе частот, в радиосигналы заданной полосы частот, что обеспечивает возможность их передачи по конкретному физическому каналу. Дополнительным свойством сложных

видов модуляции является более плотная упаковка данных в частотной области, когда на единицу полосы приходится больше передаваемой информации.

В цифровых системах передачи процесс модуляции-демодуляции можно рассматривать как способ преобразования кода в сигнал и обратно. Конкретный метод модуляции выбирается, исходя из особенностей построения системы, требуемой скорости передачи по предоставленному каналу, заданной вероятности приема (включая возможности системы защиты от ошибок) и пр. Таким образом, постановка проблемы совместной оптимизации модема и кодека направлена на решение одной важной задачи – наилучшего согласования сигнала с характеристиками канала. При поиске оптимального варианта согласования чаще всего останавливаются на выборе одного из двух критериев:

- высокой *спектральной эффективности*, т.е. передачи с высокой скоростью в узкой полосе;
- высокой *энергетической эффективности*, т.е. передачи с низким отношением несущая/шум и с максимальным занятием всей доступной полосы.

В первом случае применяют плотные созвездия сигналов (например, модуляция 64 QAM или 16 QAM) совместно с мало избыточными кодами, исправляющими ошибки. Во втором случае используются разреженные созвездия (QPSK) совместно с высоко избыточными корректирующими кодами. С учетом реальных ограничений на допустимую полосу канала и достижимое отношение несущая/шум выбирается необходимый компромисс между спектральной и энергетической эффективностями.

7.3. Классификация систем радиосвязи по частотному диапазону.

Как известно электроны при движении образуют электромагнитные колебания. Это явление Максвелл предсказал в 1865, а Генрих Герц экспериментально обнаружил в 1887.

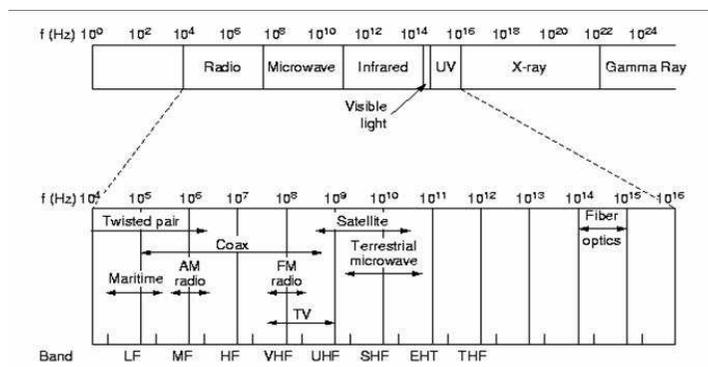


Рис. 1 Использование электромагнитного спектра для передачи данных

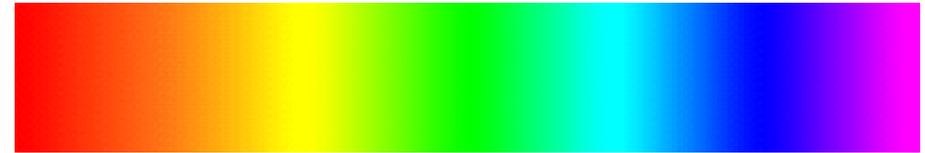


Рис. 2. Электромагнитный спектр оптического диапазона

Напомним, что число колебаний электромагнитной волны в секунду называется частотой волны (f). Расстояние между ее гребнями - длиной волны (λ). Если к источнику электромагнитных волн подключить антенну соответствующего размера, то волны будут распространяться и регистрироваться приемниками. При определенных условиях, о которых мы будем разговаривать ниже, волны будут распространяться в строго определенном направлении. В этом случае антенна приемника должна быть должным образом ориентирована в пространстве по отношению к антенне передатчика, чтобы принимать сигналы. При других условиях антенна передатчика распространяет электромагнитные волны во всех направлениях.

В вакууме электромагнитная волна распространяется со скоростью света $c = 3 \times 10^8$ м/сек. В медном проводнике эта скорость составляет $2/3$ от скорости в вакууме. Фундаментальное соотношение соединяющее f , c и λ , таково:

$$f \lambda = c \quad (1)$$

Поскольку c - константа, то зная λ , знаем f , и наоборот. Например, волны с частотой в 1 МГц, согласно этому соотношению имеют длину волны 300 метров, а волны длиной в 1 см имеют частоту 30 ГГц. На рис.1,2 представлен электромагнитный спектр. Для передачи информации из всего этого спектра используется только диапазоны: радио, микроволновый, инфракрасный, видимый и, частично, ультрафиолетовый. Диапазоны рентгеновского излучения, гамма излучения и большая часть ультрафиолета, хотя и имеют большие частоты, а потому и более предпочтительны для передачи, однако требуют сложной аппаратуры для генерации и модуляции, плохо преодолевают препятствия и, что самое главное, опасны для живой материи.

Таблица 1 Классификация диапазона частот систем связи

№ диапазона	Диапазон частот	Наименование частот	Диапазон волн	Наименование волн
4	3÷30 кГц	ОНЧ Очень низкие	100÷10 км.	Мириаметровые
5	30÷300 кГц	НЧ Низкие	10÷1 км.	Километровые
6	300÷3000 кГц	СЧ Средние	10000÷100 м.	Гектометровые
7	3000÷30 МГц	ВЧ Высокие	100÷10 м.	Декаметровые
8	30÷300 МГц	ОВЧ Очень высокие	10÷1 м.	Метровые
9	300÷3000 МГц	УВЧ Ультравысокие	100÷10 см.	Дециметровые
10	3000÷30 ГГц	СВЧ Сверхвысокие	10÷1 см.	Сантиметровые
11	30÷300 ГГц	КВЧ Крайне высокие	10÷1 мм.	Миллиметровые

12	300÷3000	ГГц	ГВЧ Гипервысокие	1÷0,1 мм.	Децимиллиметровые
----	----------	-----	------------------	-----------	-------------------

Таблица 2 Классификация диапазона частот систем связи

Обозначение диапазона	Частоты
L-band	1,4 - 1,7 ГГц
S-band	1,9 - 2,7 ГГц
C-band low	3,4 - 5,25 ГГц
C-band high	5,725 - 7,075 ГГц
X-band	7,25 - 8,4 ГГц
Ku-band	10,7 - 14,8 ГГц
K-band	15,4 - 27,5 ГГц
Ka-band	27 - 50 ГГц
W-band	65 - 110 ГГц

В нижней части рисунка диапазоны, используемые для передачи перечислены в соответствии с их официальными названиями МСС (ITU): так LF диапазон, то есть длинные волны, соответствует волнам длиной от 1 до 10 км (примерно от 30 кГц до 300 кГц). Аббревиатуры LF, MF, HF, что соответствует отечественным ДВ, СВ, КВ (т.е. длинные волны, средние и короткие) появились тогда, когда никто и не думал о частотах больше 10 МГц. Позднее появились VHF, UHF и т.д.

Системы радиосвязи можно классифицировать по различным признакам; по принадлежности служб; по рабочему диапазону частот; по физическим процессам в среде распространения; по пропускной способности и передаваемым сигналам и т. д.

Классификация по используемым диапазонам частот проводится в соответствии с Регламентом радиосвязи, который является основным международным документом, определяющим условия работы различных радиосредств. Согласно этому документу радиоспектр от 3 кГц до 3000 ГГц разбивается на девять диапазонов ($N = 9$), каждый из которых занимает полосу от $0,3 \cdot 10^N$ до $3 \cdot 10^N$ Гц. В таблицах 1,2 приведены диапазоны частот и их наименования.

В настоящее время во всем мире работает огромное количество радиосредств, число которых с каждым годом увеличивается. Это привело к тому, что в различных диапазонах многие службы радиосвязи в настоящее время вынуждены работать в одной и той же полосе частот.

В связи с этим возникла серьезная и сложная проблема электромагнитной совместимости, решать которую необходимо усилием всех стран мира.

7.4. Иерархический принцип построения цифровых телекоммуникационных систем

Цифровые системы передачи, используемые на сетях связи, соответствуют определен-

ной иерархической структуре. Иерархия ЦСП должна удовлетворять следующим требованиям:

- системы передачи должны обеспечивать передачу всех видов аналоговых и дискретных сигналов;
- скорость передачи информации должна быть выбрана с учетом использования оборудования, как аналого-цифрового преобразования, так и временного группообразования сигналов;
- должно обеспечиваться простое объединение, разделение и транзит передаваемых сигналов;
- при передаче типовых сигналов необходимо, чтобы пропускная способность ЦСП использовалась наилучшим образом;
- должна быть предусмотрена возможность взаимодействия цифровых систем передачи с аналоговыми системами;
- параметры ЦСП должны выбираться с учетом характеристик существующих и перспективных линий связи.

В настоящее время наиболее широкое распространение получили две технологии временного мультиплексирования ВОСП: плезихронная цифровая иерархия PDH и синхронная цифровая иерархия SDH.

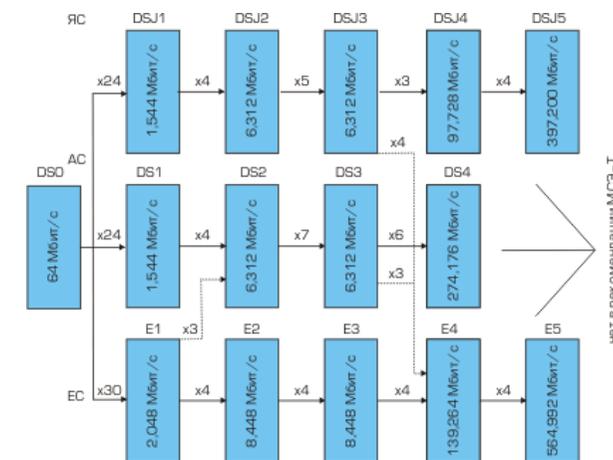


Рис.3. Плезихронная цифровая иерархия PDH

В принятой в нашей стране европейской системе PDH мультиплексор на первом уровне использует в качестве входных сигналов 30 каналов ОЦК (каналов нулевого уровня) и формирует первичный цифровой информационный поток со скоростью передачи 1920 кбит/с (30 кана-

лов по 64 кбит/с). Для осуществления синхронизации, сигнализации и контроля ошибок к нему добавляют еще два канала ОЦК, в результате чего первичный цифровой канал (ПЦК) имеет скорость 2048 кбит/с = 2,048 Мбит/с (или приблизительно 2 Мбит/с). Далее в схеме последовательного каскадного мультиплексирования, использующей мультиплексоры типа 4:1, можно сформировать цифровые каналы 2-го, 3-го, 4-го и 5-го уровней иерархии со скоростями передачи 8,448; 34,368; 139,264; 564,992 Мбит/с (или приблизительно 8; 34; 140; 565 Мбит/с). Указанная иерархия позволяет передавать соответственно 30, 120, 480, 1920 и 7680 каналов ОЦК. В странах Европы и Латинской Америки она получила обозначение E1 (ПЦК), E2, E3, E4 и E5.

Таблица 1. Стандарты цифровых систем передачи плезихронной иерархии

Уровень иерархии	Скорость передачи В (кбит/с)		
	Американская система	Японская система	Европейская система
Основной цифровой канал (ОЦК)	64	64	64
Первичный цифровой канал (ПЦК)	DS1(T1) 1544	DS1 1544	E1 2048
Вторичный цифровой канал (ВЦК)	DS2(T2) 6312	DS2 6312	E2 8448
Третичный цифровой канал (ТЦК)	DS3(T3) 44736	DSJ3 32064	E3 34368
Четвертичный цифровой канал (ЧЦК)	-	DSJ4 97728	E4 139264
Уровни сопряжения с ЕС	DS3→E4 45→140	DSJ3→E4 32→140	

Первичные ЦСП используются на местных сетях связи. В этих системах осуществляется восьмиразрядное кодирование телефонных сигналов. Оконечное оборудование таких систем в цифровой сети связи может использоваться не только для каналообразования, но и для организаций временного коммутационного поля в коммутационном оборудовании. Существуют модификации первичной системы, например: системы с дельта-модуляцией, осуществляющие передачу 40—60 телефонных сигналов; системы, осуществляющие цифровую передачу 12-канального группового сигнала с частотным разделением каналов (исходный спектр 60—108 кГц) и нескольких (до 10) телефонных сигналов, передаваемых на основе временного разделения; системы, в которых передаются как телефонные сигналы, так и сигналы вещания.

Первичные системы обычно предназначаются для уплотнения многопарных низкочастотных кабелей городских телефонных сетей и одночетверочных кабелей сельских сетей. Длина участка регенерации при этом составляет 1,5—3 км.

Цифровые системы передачи высших порядков, принятые в СССР,—вторичные, третичные, четверичные и пятеричные — используют метод двустороннего согласования скоростей, имеющий ряд существенных преимуществ по сравнению с методом одностороннего согласования скоростей, применяемым в ряде стран.

Вторичные ЦСП, скорость группового потока в которых составляет 8448 кбит/с, используются на местных и зональных сетях связи. По принципу построения различают вторичные системы следующих видов: с объединением цифровых сигналов четырех первичных систем, что обеспечивает организацию 120 каналов ТЧ; с непосредственным кодированием 120 телефонных сигналов; с кодированием группового 60-канального сигнала с частотным разделением (исходный спектр 312—552 кГц) и совместной передачей этого кодированного сигнала с цифровым потоком первичной системы.

Вторичные ЦСП предназначаются для работы по симметричным междугородным и коаксиальным кабелям 0,7/3,0 и 1,2/4,4 мм, а также волоконно-оптическим, радиорелейным и спутниковым линиям связи. Кроме того, сигналы вторичной ЦСП могут передаваться по трактам систем передачи с ЧРК при использовании многоуровневых кодов. При работе по кабельным линиям длина участка регенерации составляет 3—6 км.

Третичные ЦСП, скорость группового потока которых 84 368 кбит/с, рассчитаны на объединение сигналов четырех вторичных систем и используются на местных, зональных и магистральных сетях связи. Третичная ЦСП обеспечивает организацию до 480 каналов ТЧ. По этой системе могут передаваться телевизионные вещательные сигналы, преобразованные в цифровую форму с использованием методов устранения избыточности. Разновидностью третичной ЦСП является система, осуществляющая кодирование и передачу сигналов 300-канальной группы с частотным разделением (исходный спектр 812—2048 кГц). С помощью третичных систем уплотняются коаксиальные кабели с парами 1,2/4,4 мм, симметричные междугородные кабели, а также волоконно-оптические и радиорелейные линии. Длина участка регенерации при работе по коаксиальному кабелю составляет 2,3—3,2 км.

Четверичные ЦСП, скорость группового потока которых 139 264 кбит/с, осуществляют объединение сигналов четырех третичных систем и предназначены для работы на местных, зональных и магистральных сетях связи. Четверичная система обеспечивает организацию 1920 телефонных каналов и может использоваться для передачи сигналов телевизионного вещания. Система предназначена для работы по коаксиальным кабелям с парами 2,6/9,5 и 1,2/4,4 мм, а также по волоконно-оптическим линиям связи. Длина регенерационного участка при работе по кабелю с парами 2,6/9,5 мм составляет 3—3,5 км, а при работе по кабелю 1,2/4,4 мм -1,5-2,0 км. На городских сетях четверичная ЦСП позволяет создать многоканальные соединительные линии между АТС большой емкости, а также обеспечить высококачественную передачу телевизионных сигналов от вынесенных передающих станций к телецентру и передатчику.

Объединение сигналов нескольких четверичных систем позволит получить более мощную пятеричную ЦСП на 7680 каналов ТЧ. Известны предложения по созданию сверхширокополосных систем с тактовыми частотами до 1,6 ГГц.

7.5. Системы плезиохронных цифровых иерархий

7.5.1. Общие особенности плезиохронных цифровых иерархий

Наличие стандартных скоростей передачи и фиксированных коэффициентов мультиплексирования дало возможность на практике не только по отдельности использовать три схемы мультиплексирования АС, ЯС и ЕС (или СЕРТ), но и обеспечить их взаимодействие.

При формировании первичного уровня иерархии использовалась схема мультиплексирования с байт-интерливингом (чередованием байтов). В результате этого сформированный фрейм должен был бы иметь длину, кратную байту. Однако добавление синхронизирующих и сигнализирующих бит нарушает эту кратность не только на уровне фрейма, но и на уровне мультифрейма (например, в системе D2). Кроме того, при использовании каскадного мультиплексирования даже при достаточно жесткой, но локальной синхронизации входных последовательностей, подаваемых на мультиплексор от разных абонентов/пользователей при приеме/передаче, приходится (с целью общей синхронизации цифровых потоков) выравнивать цифровые последовательности перед мультиплексированием, добавляя выравнивающие биты к входным цифровым последовательностям (т.е. осуществляя так называемую процедуру стаффинга). Это приводит к тому, что начиная с формирования вторичных цифровых каналов приходится использовать схему мультиплексирования с бит-интерливингом (чередованием битов), что делает невозможным идентификацию байтов каждого канала в общем потоке. Итак, на верхних уровнях иерархии используется внутренняя побитовая синхронизация, при которой мультиплексор сам выравнивает скорости входных потоков путем добавления нужного числа выравнивающих бит в каналы с относительно меньшими скоростями передачи. Этот наиболее простой вариант называют положительным выравниванием, хотя могут использоваться и другие варианты. На каждом последующем уровне мультиплексирования эта схема повторяется, добавляя новые выравнивающие биты. Эти биты затем удаляются при демultipлексировании на приемной стороне для восстановления исходной цифровой последовательности. Такой процесс передачи получил название плезиохронного (т.е. почти синхронного), а цифровые иерархии АС, ЯС и ЕС получили соответственно название плезиохронных цифровых иерархий ПЦИ. Кроме синхронизации на уровне мультиплексора второго порядка также происходит формирование фреймов и мультифреймов, которые позволяют структурировать последовательность в целом, что позволяет получать информацию о сигнализации, аварийных ситуациях и кодовых комбинациях избыточных кодов CRC, используемых для обнаружения ошибок в принятых фреймах.

7.5.2. Недостатки плезиохронной цифровой иерархии

Общая схема канала передачи с использованием технологии PDH даже для простой топологии сети “точка – точка”, но при скорости 140 Мбит/с должна включать три уровня муль-

типлексирования на передающей стороне (для ЕС, например, 2—>8, 8—>34 и 34—>140) и три уровня демultipлексирования на приемной стороне, что приводит к достаточно сложной аппаратурной реализации таких систем. Однако существенное удешевление цифровой аппаратуры за последние двадцать лет и использование ВОК в качестве среды передачи привели к тому, что PDH получила значительное распространение уже в восьмидесятые годы. Эти системы позволяли осуществить передачу большого количества каналов высококачественной цифровой телефонной связи.

Однако передача данных на скорости 64 кбит/с на основе протокола пакетной коммутации X.25 выявила ряда недостатков PDH технологии. Их суть в том, что добавление выравнивающих бит делает невозможным идентификацию и вывод (или ввод), например, потока 64 кбит/с, или даже 2 Мбит/с, “зашитого” в поток 140 Мбит/с, без полного демultipлексирования этого потока и удаления выравнивающих бит. Осуществляя, и достаточно часто, такой ввод/вывод, приходится проводить относительно сложную операцию трехуровневого демultipлексирования (“расшивания”) PDH сигнала с удалением/добавлением выравнивающих (на всех трех уровнях) бит и его последующего трехуровневого мультиплексирования (“сшивания”) с добавлением новых выравнивающих бит.

Другое узкое место технологии PDH – слабые возможности в организации служебных каналов для целей контроля и управления потоком в сети и практически полное отсутствие средств маршрутизации мультиплексированных потоков нижних уровней, что крайне важно для использования систем PDH в сетях передачи данных.

7.6. Синхронная цифровая иерархия – SDH

Системы PDH стали применяться не только для организации телефонных каналов, но и для передачи данных. Однако наличие в PDH потоках выравнивающих битов, делает невозможным прямое извлечение из потока, составляющих его компонентов. Так, чтобы извлечь из потока E4 поток E1 необходимо демultipлексировать E4 на четыре E3, затем один из E3 на четыре E2, и только после этого можно вывести требуемый E1. А для организации ввода/вывода требуется трехуровневое демultipлексирование, а затем трехуровневое мультиплексирование. Понятно, что использование систем PDH в сетях передачи данных, требует большого количества мультиплексоров, что значительно удорожает сеть и усложняет ее эксплуатацию.

Этот недостаток можно устранить, если объединяемые потоки будут синхронными, для чего оборудование этих сетей должно использовать синхронизацию от одного высокоточного источника.

В 80-х годах были разработаны стандарты (ITU-T G.707, G.708, G.709) на иерархию

синхронной оптической сети SONET в США и в Европе на синхронную цифровую иерархию SDH для использования на волоконно-оптических линиях связи, а также стандарты на соответствующие интерфейсы, необходимые для стыковки оборудования.

В сетях SDH, как и в SONET, используются синхронные схемы передачи с байт-интерливингом при мультиплексировании. В качестве формата основного сигнала первого уровня в иерархии SDH был принят синхронный транспортный модуль STM-1, скорость передачи которого 155.52 Мбит/сек. Мультиплексирование с коэффициентом кратности 4 дает следующий ряд скоростей SDH иерархии: STM-4, STM-16, STM-64 или соответственно 622.08, 2488.32, 9953.28 Мбит/сек. Ряд скоростей SONET начинается с сигнала OC-1, имеющего скорость 51.84 Мбит/сек, а далее сигналы OC-3, OC-12, OC-48 и т.д. совпадают по скорости с STM-1, STM-4, STM-64 и т.д. Отметим, что выбор скорости 155.52 Мбит/сек для STM-1 (или 51.84 для OC-1) был неслучаен. Это было сделано при разработке стандартов SONET/SDH для "примирения" американской и европейской систем PDH и позволило мультиплексировать потоки, как одной, так и другой иерархий в STM-1.

Мультиплексирование STM-1 в STM-N может осуществляться как каскадно: $4 \times \text{STM-1} = \text{STM-4}$, $4 \times \text{STM-4} = \text{STM-16}$, $4 \times \text{STM-16} = \text{STM-64}$, так и непосредственно: $4 \times \text{STM-1} = \text{STM-4}$, $16 \times \text{STM-1} = \text{STM-16}$, $64 \times \text{STM-1} = \text{STM-64}$.



Рис.3 Структура фрейма STM-N

7.6.1. Схема мультиплексирования PDH сигналов в технологии SDH.

Кадр STM-1 имеет 2430 байт и представляется в виде матрицы размером в 9 строк и 270 столбцов (Рис. 3). Первые 9 столбцов составляют байты секционного заголовка (Section Overhead – SOH), которые содержат байты служебных каналов, индикации ошибок передачи, управления синхронизацией, идентификации наличия STM-1 в кадре STM-N, а также указатель на начало административного блока в поле полезной нагрузки, в котором размещаются виртуальные контейнеры четвертого или третьего уровней (речь о них пойдет ниже). Поле полезной нагрузки состоит из оставшихся 261 столбца. Учитывая, что частота следования кадров

STM-1 равна 8 КГц, получаем скорость полезной нагрузки $9 \times 261 \times 8 \times 8 = 150336$ Кбит/сек.

Синхронные сети не смогли бы получить широкого распространения, если бы они не обеспечивали поддержки скоростей PDH. Поэтому стандартами SDH определена поддержка в сетях доступа сигналов PDH (исключение для европейской иерархии составил сигнал E2). Для передачи в общем случае не синхронных PDH потоков в поле полезной нагрузки синхронного транспортного модуля, была разработана специальная схема мультиплексирования. По этой схеме кадры потоков PDH отображаются в контейнеры фиксированного размера C-n (C-11 для T1, C-12 для E1, C-31 для E3, C-32 для T3 и C-4 для E4). Контейнеры C-n снабжаются маршрутными заголовками (Path Overhead – POH), содержащими управляющую информацию и статистику о прохождении контейнера. Такие контейнеры (снабженные POH) получили название виртуальных контейнеров (VC) и, в зависимости от уровня соответствующего им PDH сигнала, различают виртуальные контейнеры первого, третьего и четвертого уровней. В виртуальные контейнеры третьего и четвертого порядка вместо C-n могут входить группы компонентных блоков (TUG). Для устранения несинхронности входных потоков, каждому виртуальному контейнеру присписывается указатель, содержащий информацию о положении виртуального контейнера в поле полезной нагрузки, т.е виртуальный контейнер не привязан жестко к определенному месту в поле полезной нагрузки, а может как бы «плавать» в нем. На Рис. 4. приведена европейская схема мультиплексирования потоков T1, E1, E3, T3 и E4 в STM-N, которая является частью общей схемы мультиплексирования PDH сигналов в технологии SDH.

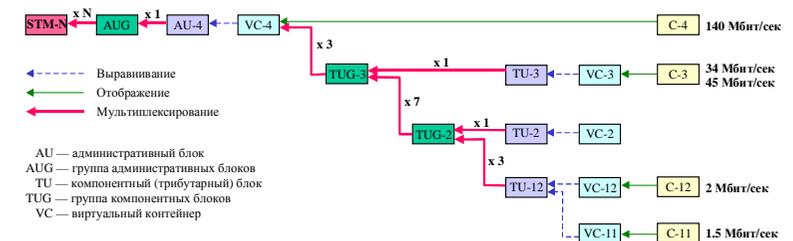


Рис. 4. Схема мультиплексирования PDH сигналов в технологии SDH

Из этой схемы легко увидеть, что в STM-1 могут быть мультиплексировано до 63 потоков E1, до трех E3 и один E4.

7.6.2. Структура цикла SDH

Сигнал SDH-1 транспортируется, как синхронная структура, которая включает набор байтов (по 8 бит), организованный в виде двухмерного массива

Цикл SDH состоит из 2-х частей:

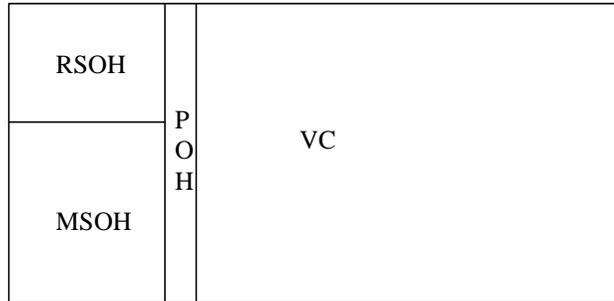


Рис. 1. Цикл SDH

1. Секционный заголовок (SOH=RSOH+MSOH) – область сигнала, которая обеспечивается в каждом цикле SDH для выполнения функций, поддерживающих и обслуживающих транспортировку «виртуальных контейнеров» между смежными узлами сети
2. Виртуальный контейнер(VC+POH) – включает “контейнерную” область, которая несет
3. трафик клиента – полезную нагрузку, и трактовый заголовок POH

Байты в цикле передаются слева-> направо, сверху ->вниз, т.е цикл передается как последовательность 9 строк.

В соответствии с данными рис.4, цикл SDH можно представить как двухмерный массив из N-строк и M-столбцов ячеек, каждая из которых – отдельный байт синхронного сигнала. Идентичность каждого байта известна, и сохраняется относительно байтов цикловой синхронизации, известных как A1 и A2, расположенных в самом начале массива и обеспечивающих точку отсчета, от которой определяются все остальные байты.

Для сигнала STM-1: N=9 M=270.

Расчет базовой скорости SDH производится следующим образом:

$$V=N(\text{строк}) \cdot M(\text{столбцов}) \cdot 8 \text{ бит (размер ячейки)} \cdot 8000 \text{ циклов/с} = 155,52 \text{ Мбит/с}$$

*-согласно теории Найквиста (удвоенная самая высокая частота канала ТЧ 4кГц)

7.6.3. Сети SDH

SDH (SONET- североамериканский аналог)– это стандарт для высокоскоростных оптических сетей связи; более известный, как синхронная цифровая иерархия (**S**ynchronous **D**igital **H**ierarchy), предназначен для обеспечения простой, экономичной и гибкой инфраструктуры сети связи.

Цифровые сети PDH были асинхронными системами, так как не использовали внешнюю синхронизацию от центрального опорного источника. В них потеря бит приводила не только к потере информации, но и к нарушению синхронизации.

В сетях SDH средняя частота всех местных таймеров или одинакова (синхронна) или близка к синхронной (плезизохронна) благодаря использованию центрального таймера (источника) с точностью не хуже 10^{-9} . В этой ситуации необходимость выравнивания фреймов или мультiframeов стоит не так остро, а диапазон выравнивания значительно уже.

Кроме того, упрощается процедура выделения определенного фрагмента потока, если ввести указатели начала этого фрагмента в структуре инкапсулирующего его фрейма. Использование указателей позволяет гибко компоновать внутреннюю структуру контейнера-носителя. Сохранение указателей в буфере и их дополнительная защита кодами с коррекцией ошибок позволяют получить исключительно надежную систему локализации внутренней структуры передаваемой по сети полезной нагрузки. Сегодня сети SDH используются во всех традиционных областях.

Также как и PDH сети SDH построены по иерархическому принципу. Самый низкий уровень SDH - иерархии назван «Синхронным Транспортным Модулем» первого уровня или STM-1, имеющий скорость 155 Мб/с. Сигналы более высокого уровня получают мультиплексированием с «чередованием байтов» сигналов низшего уровня. Линейная скорость более высокого уровня STM-N сигнала равна произведению N на 155.52 Мбит/с, т.е. линейную скорость сигнала самого низкого уровня.

Таблица. 1

Модуль STM (SDH)	Битовая скорость, Мбит/с
STM-1	155,52
STM-4	622,08
STM-16	2488,32
STM-64	9953,28
STM-256	39813,12

Основные типы оборудования, применяемые в сетях SDH, условно можно разделить на терминальные мультиплексоры, мультиплексоры ввода/вывода и кросс-коннекторы или коммутаторы. Это деление условно, так как выпускаемые производителями SDH оборудования устройства, как правило, могут сочетать в себе несколько функций.

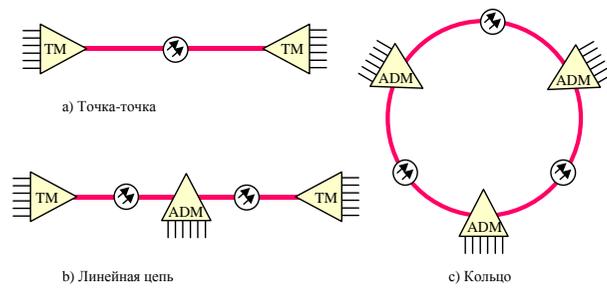


Рис. 1. Базовые топологии SDN сетей

Терминальный мультиплексор (TM) предназначен для мультиплексирования и демultipлексирования сигналов PDH и STM (в терминологии SDN их называют трибутарными или компонентными интерфейсами) в агрегатный поток STM-N. Он также может осуществлять локальную коммутацию с одного трибутарного интерфейса на другой. Мультиплексоры ввода/вывода (ADM) имеют на входе те же наборы интерфейсов, что и TM, и, как правило, два агрегатных потока STM-N (условно называемых "восточный" и "западный"). В этих мультиплексорах часть данных из агрегатных потоков выводится на трибутарный интерфейс, другая часть вводится (добавляется) с трибутарного интерфейса для последующей передачи, остальной трафик проходит через устройство без какой-либо дополнительной обработки. Кросс-коннектор (DSX) — распределительный узел сети, имеющий несколько, как правило, однотипных интерфейсов STM-N, осуществляющий между ними коммутацию виртуальных контейнеров различного уровня.

Основные конфигурации сетей SDN это точка-точка, линейная цепь и кольцо, хотя применяются и другие топологии, например ячеистая (Рис. 5).

Защитные свойства сетей SDN достигаются за счет средств контроля, переносимых в служебных байтах заголовков (SOH, POH). Один из способов — резервирование участков сети по схеме 1+1, используя разнесенные трассы для прокладки оптического кабеля. В этом случае при повреждении кабеля на одном участке сети, происходит автоматическое переключение на резервный участок. Однако наиболее популярная топология — самовосстанавливающееся кольцо. При разрыве в кольце служебные байты пострадавших линий сигнализируют о невозможности прохождения трафика. В ответ на это ближайшие узлы перенаправляют потоки с поврежденного участка в обход по кольцу. Этот процесс занимает доли секунды.

Надежность сетей SDN, их способность к самовосстановлению, возможность поддержки потоков в широком диапазоне скоростей, развитые средства конфигурирования, мониторинга и администрирования, привели к тому, что эти сети широко используются в качестве транспортных сетей поддерживающих передачу различных типов трафика (голос, данные, видео и т.д.) от широкого круга технологий (Frame Relay, ISDN, ATM и др.) Тенденция послед-

них лет — вытеснение существующих систем PDH сетями SDN, а также использование этой технологии не только операторами связи, но и для построения магистралей корпоративных информационных систем.

Однако по мере распространения SDN технологий, при объединении сетей различных операторов связи, все более остро встает проблема глобальной синхронизации узлов, и этот момент нельзя недооценивать.

7.7. Преобразование сигналов в ЦСП

7.7.1. Принципы цифрового преобразования аналоговых сигналов

Как следует из обобщенной структурной схемы любой СП первыми и последними операциями совершаемыми в этих системах передачи являются аналого-цифровые и цифро-аналоговые преобразования сигналов. Эти преобразования являются важнейшими не только для ЦСП, но и в решении задач современной измерительной техники, обработки сигналов и вообще для любых радиоэлектронных систем, которые в результате своей работы должны формировать качественные данные.

Импульсно-кодовая модуляция (ИКМ) (pulse code modulation - PCM) является наиболее распространенным методом цифрового преобразования аналоговых сигналов. При ИКМ, как и при других видах аналого-цифрового преобразования, происходит дискретизация во времени передаваемого сигнала. Сформированные при дискретизации отсчеты преобразуются затем в группы кодовых символов. Если каждый символ, входящий в состав кодовой группы, может принимать любое из $г$ значений (0, 1, 2, ..., $г-1$), а кодовая группа содержит n символов, то возможно формирование $г^n$ различных кодовых групп, где $г$ - основание кода, n - число разрядов.

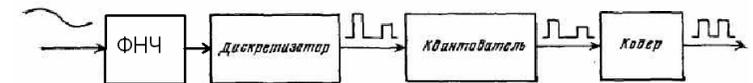


Рис. 1. Структурная схема импульсно-кодового модулятора.

Аналоговые сигналы на входе цифровой системы передачи принимают любые значения в пределах заданного амплитудного диапазона. Используя n -разрядные кодовые группы, можно передать информацию не более чем о $г^n$ различных значениях сигнала. Поэтому при цифровой передаче происходит амплитудное квантование передаваемого сигнала. Таким образом, при импульсно-кодовой модуляции осуществляются три вида преобразований:

- дискретизация во времени исходного сигнала,

- квантование амплитуд дискретных отсчетов сигнала

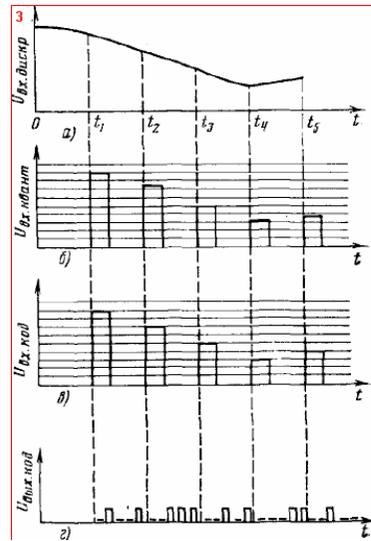


Рис. 2. Временные диаграммы работы импульсно-кодového модулятора:
 а) входной аналоговый сигнал; б) последовательность отсчетов;
 в) последовательность квантованных отсчетов; г) последовательность кодовых групп

кодирование, т. е. формирование кодовых групп, соответствующих квантованным значениям отсчетов. На рис. 1 Структурная схема импульсно-кодového модулятора и на рис. 1.2 - временные диаграммы, иллюстрирующие отдельные виды преобразований. В реальных аналого-цифровых преобразователях квантование и кодирование, как правило, осуществляются одновременно. Возможно также совмещение операций дискретизации и квантования. Известны аналого-цифровые преобразователи, в которых временная дискретизация производится после квантования и кодирования. Временная дискретизация сигналов. Представляет собой амплитудно-импульсную модуляцию аналогового сигнала.

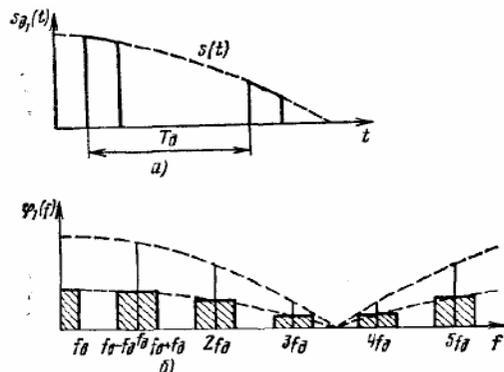


Рис. 3. Амплитудно-импульсная модуляция первого рода:
 а) отсчеты аналогового сигнала; б) спектр последовательности отсчетов

На рис. 3 и 4 показаны временные диаграммы и спектры, соответствующие различным способам формирования отсчетов. Рисунок 1.3 иллюстрирует амплитудно-импульсную модуляцию первого рода (АИМ-1), при которой напряжение за время отсчета изменяется в соответствии с изменениям входного сигнала. При амплитудно-импульсной модуляции второго рода (АИМ-2) напряжение отсчета пропорционально значению сигнала в моменты дискретизации (рис.1.4); изменения входного сигнала в течение отсчета не влияют на модулированный сигнал. Спектр модулированного сигнала в общем случае содержит гармоники частоты дискретизации, каждой из которых соответствуют верхняя и нижняя боковые полосы (если сигнал $s(t)$ и последовательность модулируемых импульсов не содержат постоянной составляющей, то в спектре модулированного сигнала отсутствуют частота дискретизации f_d и гармоники этой частоты).

Амплитуды гармоник частоты дискретизации f_d изменяются пропорционально спектральной плотности модулируемого импульса.

Спектральные плотности боковых полос, соответствующих определенной гармонике частоты дискретизации, при АИМ-1 пропорциональны спектральной плотности модулируемого импульса на частоте данной гармоники. При АИМ-2 спектральные плотности боковых полос пропорциональны значениям спектральной плотности модулируемого импульса на частотах этих полос и распределение спектральной плотности в пределах боковых полос отличается от подобного распределения в исходном сигнале. Значение отсчета должно быть постоянным в процессе кодирования. В противном случае могут возникать ошибки при формировании кодовой группы. Поэтому временная дискретизация сигнала в системах с ИКМ соответствует АИМ-2.

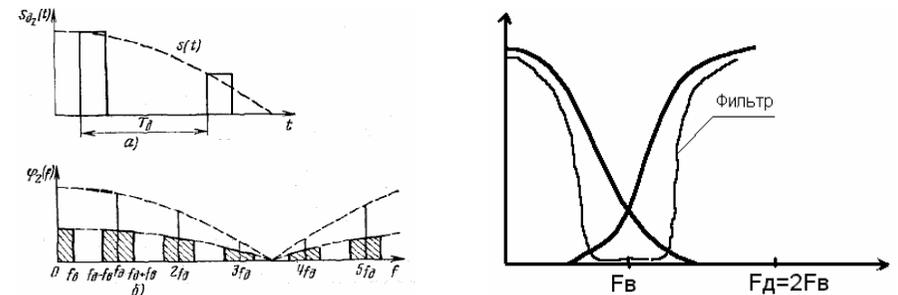


Рис. 4. Амплитудно-импульсная модуляция второго рода:
 а) отсчеты аналогового сигнала;
 б) спектр последовательности отсчетов
 в) АЧХ фильтра

Как видно из рис. (3,б) и рис (4,б), исходный сигнал может быть выделен из последовательности $s_n(t)$, если боковые полосы не накладываются друг на друга. Для этого необходимо выполнение условия

$$f_o \geq 2 \cdot f_e \quad (1)$$

Условие (1) представляет собой теорему Котельникова, согласно которой непрерывный сигнал $s(t)$ может быть восстановлен без искажений из последовательности дискретных отсчетов $s_n(t)$ этого сигнала, если частота дискретизации f_d по крайней мере в 2 раза выше наибольшей частоты, содержащейся в спектре исходного сигнала. Поскольку все реально существующие непрерывные сигналы связи представляют собой случайные процессы с бесконечно широким спектром, причем основная энергия сосредоточена в относительно узкой полосе частот, перед дискретизацией необходимо с помощью фильтра нижних частот ограничить спектр сигнала некоторой частотой F_b . рис (1.4,в). К фильтру очень высоки требования, его фронт должен быть как можно положе. Основную информацию несёт первая гармоника и если частоты f_e и $f_o - f_e$ очень близки, то фильтра не справится, поэтому приходится делать зазор между этими частотами больше.

Увеличение частоты дискретизации позволяет упростить фильтры, ограничивающие спектр аналогового сигнала в передающем, оборудовании, а также фильтры, выделяющие спектр исходного сигнала при демодуляции в приемном оборудовании.

В процессе квантования по уровню значение каждого АИМ - отсчета заменяется ближайшим разрешенным значением. характеристиками квантующего устройства являются следующие:

- число уровней квантования $N_{кв}$;
- шаг квантования ΔC разность между двумя соседними разрешенными уровнями;
- напряжение ограничения $U_{огр}$ максимальное значение амплитуды отсчета, подвергаемого квантованию.

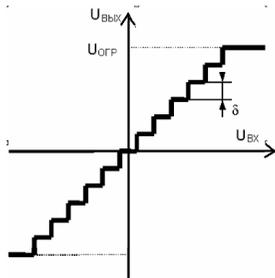


Рис. 5. Амплитудная характеристика равномерного квантователя.

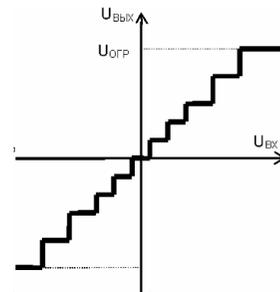


Рис. 6. Амплитудная характеристика неравномерного квантователя.

Если $\Delta C = \text{const}$, то квантование называют равномерным. Амплитудная характеристика равномерного квантователя показана на рис. 5.

Если $\Delta C = \text{var}$, то квантование называют неравномерным. Амплитудная характеристика неравномерного квантователя показана на Рис. 6.

Ошибка квантования - разность между истинным значением отсчета и его квантованным значением. При равномерном квантовании величина ошибки квантования не превышает половины шага квантования.

На рис. 7 показаны амплитудное квантование аналогового сигнала и вид функции, определяющей ошибку квантования.

Ошибка квантования представляет собой функцию с большим числом резких скачков, частота следования которых существенно выше частоты исходного сигнала. Поэтому при амплитудном квантовании расширяется спектр сигнала. При квантовании сигнала, прошедшего временную дискретизацию, соседние боковые полосы вследствие расширения спектра будут накладываться друг на друга. В результате в полосу приемного фильтра, выделяющего спектр исходного сигнала (при ширине спектра $0.5f_d$) будут попадать составляющие всего расширенного спектра квантованного аналогового сигнала. Следовательно, при оценке искажений исходного сигнала следует определять полную мощность искажений квантования. При квантовании гармонического сигнала энергетический спектр искажений квантования является дискретным. При квантовании реальных сигналов, занимающих определенную полосу частот, энергетический спектр искажений квантования в полосе сигнала принимается равномерным.

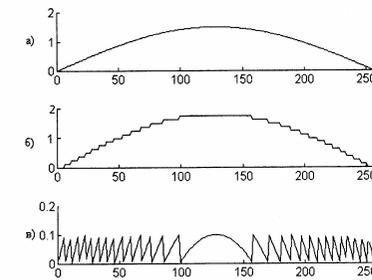


Рис. 1. Сигнал до квантования (а), сигнал после квантования (б), ошибка квантования (в)

Рис. 7 а) Амплитудное квантование сигнала. б) Ошибка квантования.

Мощность искажений квантования определяется выражением:

$$P_{ш.к} = \frac{\Delta C^2}{12} \quad (2)$$

При равномерном квантовании, когда все шаги квантования равны по величине,

$$P_{кв} = \frac{\delta^2}{12}. \quad (3)$$

При равномерном квантовании ошибка квантования ограничена пределами $-\frac{\delta}{2} \leq \Delta_{кв} \leq \frac{\delta}{2}$

и в этих пределах плотность распределения значений ошибки постоянна. Искажения квантования действуют одновременно с передачей сигнала.

Поэтому влияние этих искажений удобно оценивать по отношению сигнал-шум квантования. Определим максимальную величину отношения сигнал-шум, достижимую при равномерном квантовании и использовании n-разрядного двоичного кода. При передаче двуполярных сигналов с симметричным распределением мгновенных значений шаг квантования

$$\delta = \frac{U_{огр}}{2^{n-1}}, \quad (4)$$

а при передаче сигналов одной полярности

$$\delta = \frac{U_{огр}}{2^n} \quad (5)$$

где 2^n – число уровней квантования

7.7.2. Оценка отношения уровня сигнала к шуму квантования

Численной характеристикой влияния шумов квантования на точность кодирования аналогового сигнала может служить отношение мощностей сигнала и шума квантования. Для двуполярного сигнала это отношение равно:

$$\frac{P_c}{P_{ш.к.}} = 10.8 + 20 \lg\left(\frac{U}{U_{огр}}\right) + 6n - 6, \quad (\text{дБ}). \quad (1)$$

Для однополярного сигнала:

$$\frac{P_c}{P_{ш.к.}} = 10.8 + 20 \lg\left(\frac{U}{U_{огр}}\right) + 6n, \quad (\text{дБ}).$$

Изменение мощности передаваемого сигнала приводит к снижению величины сигнал-шум. На рис. 1.8 приведена зависимость отношения сигнал-шум (с учетом искажений, как квантования, так и ограничения)

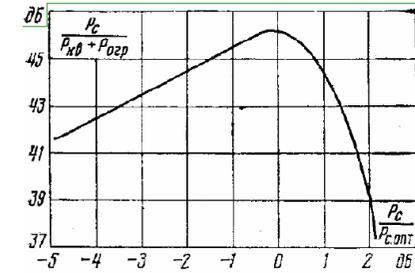


Рис. 1. Зависимость отношения сигнал-шум от уровня сигнала

от уровня сигнала с нормальным законом распределения мгновенных значений. Резкое падение величины отношения сигнал-шум при превышении сигналом оптимального уровня вызвано ростом мощности искажений ограничения.

Уменьшение сигнала вызывает пропорциональное снижение отношения сигнал-шум, т. е. передача слабых сигналов сопровождается большими искажениями.

Если изменять шаг квантования пропорционально значению квантуемого сигнала, то отношение сигнал-шум при изменении сигнала будет сохраняться постоянным. Переменную величину шага можно получить, например, с помощью устройства с нелинейной амплитудной характеристикой и равномерного квантователя (рис. 1.9) Исходный сигнал передается через устройство, коэффициент передачи которого обратно пропорционален величине сигнала, квантуется и проходит через схему, нелинейность которой обратная нелинейности входного устройства. Таким образом, перед квантованием осуществляется компрессия (сжатие) динамического диапазона входного сигнала, а после квантования - его расширение, что обеспечивает общую линейность системы передачи.



Рис.2. Структурная схема устройства неравномерного квантования с компрандированием аналогового сигнала

Совокупность операций сжатия диапазона, осуществляемого компрессором, и расширения, производимого экспандером, называется **компрандированием** сигнала.

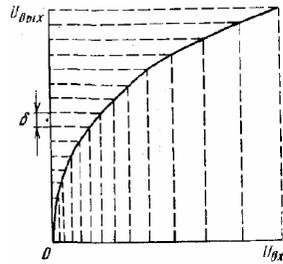


Рис.3. Характеристика компрессии

Амплитудная характеристика входного устройства $y(U_{вх})$ (рис. 1.10) называется характеристикой компрессии [Характеристика компрессии связывает шкалу равномерного квантования (ось ординат) со шкалой неравномерного квантования и характеризует любую систему неравномерного квантования даже при отсутствии компрессора как отдельного узла]. Шаг неравномерного квантования определяется из выражения.

$$\delta(U_{вх}) = \frac{\delta_p}{y'(U_{вх})},$$

где $\delta_p = \frac{U_{ошп}}{2^{n-1}}$ величина шага при равномерном квантовании. Постоянство отношения сигнал-шум квантования соответствует условию

$$U_{вх} y'(U_{вх}) = U_{вх} \frac{dy}{dU_{вх}} = const$$

После разделения переменных и интегрирования получим

$$y \approx c_1 \ln U_{вх} + c_2 = c_1 \ln \mu U_{вх} \tag{2}$$

где c_1 и c_2 - постоянные интегрирования; $\mu = e^{\frac{c_1}{c_2}}$ Функция $y(U_{вх})$ должна удовлетворять условиям:

$$y(0) = 0; \quad y(1) = 1, \tag{3}$$

условия (1.3) не выполняются при конечных значениях c_1 и c_2 Если видоизменить выражение (1.2) и принять

$$y = c \cdot \ln(\mu U_{вх} + b) \tag{4}$$

то условия (1.3) выполняются .при

$$b = 1 \text{ и } c = \frac{1}{\ln(1 + \mu)}.$$

Подставляя b и c в выражение (1.4), получаем искомую зависимость

$$y = \frac{\ln(1 + \mu|U_{вх}|)}{\ln(1 + \mu)},$$

которая широко используется в АЦП ИКМ.

Зависимости величины $\frac{U_{вх}}{\delta(U_{вх})}$ от входного сигнала для различных μ приведены на рис.

4. При увеличении μ расширяется область значений входного сигнала, в которой отношение сигнал-шум квантования приблизительно постоянно при $(\mu \cdot U_{вх}) \gg 1$.

В современных ЦСП находят применение две логарифмические характеристики компрессии (типов А и μ), которые удобно изображать и описывать в нормированном виде $y=f(x)$, где $y = U_{вх}/U_{огр}$, $x = U_{огр}/U_{вх}$:

$$y = \begin{cases} \frac{A|x|}{1 + \ln A}; & 0 \leq |x| \leq \frac{1}{A} \\ \frac{1 + \ln(A|x|)}{1 + \ln A}; & \frac{1}{A} \leq |x| \leq 1 \end{cases},$$

$$y = \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}; \quad 0 \leq |x| \leq 1,$$

где $A=87,6$ и $\mu=255$ - параметры компрессии.

Характеристика компрессии типа А используется в ЦСП, соответствующих европейской ПЦИ, а типа μ - в ЦСП, соответствующих североамериканской ПЦИ.

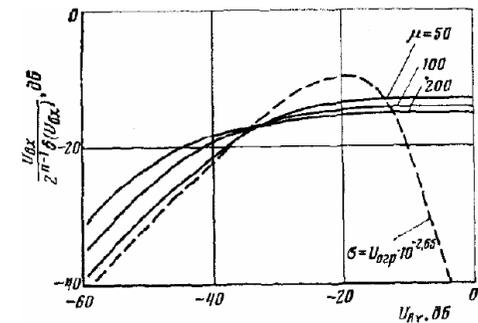


Рис. 4. Зависимость $\frac{U_{вх}}{\delta(U_{вх})}$ от величины входного сигнала.

Оценим выигрыш от компрессии с μ и А законами

Если δ - выигрыш,

$$\delta = \frac{\frac{\Delta U_c}{\Delta U_{ax}}}{\frac{U_c}{\Delta U}} \Big|_{U_c \rightarrow 0} = y'(0)$$

$$y'(0) = \frac{1}{\ln(1 + \mu)} \cdot \frac{1}{1 + \mu \cdot |U|} \Big|_{U=0} = \frac{\mu}{\ln(1 + \mu)}$$

μ- закон:

При μ=250

$$\delta = 20 \cdot \lg\left(\frac{100}{\ln(101)}\right) = 26.7(\text{дБ})$$

A- закон:

При A=87.6

$$\delta(A) = y'(0) = 20 \lg\left(\frac{A}{1 + \ln A}\right) = 24.08(\text{дБ}).$$

7.8. Волоконно- оптические системы передачи (ВОСП)

7.8.1. История развития ВОСП

Сейчас уже общеизвестно насколько важно для решения телекоммуникационных задач внедрение в практику волоконно-оптических систем передачи информации. Действительно, требования к каналам связи по их информационной емкости, качеству связи и ее надежности, помехозащищенности, функциональным возможностям, эксплуатационным и стоимостным показателям столь высоки, что удовлетворить им, без использования новых информационных технологий на основе волоконно-оптических каналов связи, невозможно. Уже сейчас скорости передачи информации 2,5 Гбит/с и 10 Гбит/с являются стандартными для нашедших самое широкое применение цифровых систем связи синхронной цифровой иерархии (СЦИ или SDH систем). Отметим, что эти системы почти полностью ориентированы на использование ВОЛС. Учитывая, что на стандартный телефонный канал требуется 64 кбит/с, легко подсчитать, что число телефонных каналов передаваемых в такой системе по одному волокну имеет

порядок величины 10^5 . Огромные возможности появляются при использовании вместе с временным еще и спектрального уплотнения каналов. В этом случае пропускная способность может быть в принципе повышена до 1,5 Тбит/с. Число телефонных каналов, передаваемых в волокне при этом увеличивается еще на два порядка. Такие параметры недостижимы для других типов систем передачи информации. Помимо высокой информационной емкости ВОСП обладают целым рядом других уникальных качеств. Среди них - очень малое затухание в линии связи и как следствие большие длины регенерационных участков, радикальный уход от проблем электромагнитной совместимости очень остро стоящих в системах связи других типов. Волоконная линия не требует дорогостоящей и дефицитной меди, но использует в основе своей один из наиболее распространенных в природе элементов - кремний. Конечно ВОСП дорогие системы, но в пересчете на стоимость передачи одного бита информации - это самые дешевые и экономически наиболее выгодные системы. Именно поэтому ВОСП являются наиболее интенсивно внедряющимися в практику системами. В настоящее время около 70% всей информации передается с помощью ВОСП.

Назовем основные физические причины бурного развития ВОСП в технике связи. Таких причин несколько.

Во-первых, это очень высокая частота света, выполняющего роль несущей информационного сигнала. Действительно, длина волны оптических волн на четыре - пять порядков меньше длины волны излучения СВЧ диапазона и соответственно частоты оптических волн $f_{\text{опт}} = (10^4 - 10^5) f_{\text{СВЧ}}$. Высокая несущая частота $f_{\text{опт}} = 10^{14} - 10^{15}$ Гц дает возможность использовать очень короткие импульсы излучения вплоть до импульсов пикосекундного и даже фемтосекундного диапазона, что является необходимым условием получения высокого быстродействия цифровых ВОСП.

Во-вторых, важнейшим свойством лазеров источников светового излучения в ВОСП является когерентность их излучения. В результате энергия лазерного излучения не размазывается по многим частотам, а сосредоточена в узкой полосе частот. Благодаря этому свойству, резко могут быть снижены потери на ввод излучения и потери на распространение. Например, длина регенерационного участка линии связи при использовании лазерного источника излучения достигает 100 км, тогда как для некогерентных источников она имеет порядок величины 1 км. Когерентность света оценивается шириной его спектра $\Delta\lambda$. Для лазерного диода величина $\Delta\lambda \cong 0.1$ нм, в то время как для светодиода она по крайней мере на два порядка большая.

В-третьих, ограничительную роль таких важных физических процессов как дисперсия волн в волоконном световоде и затухание волн в световодах современная технология позволяет свести к минимуму. Так, например, на длине волны 1.55 мкм затухание в промышленном кабеле ОКЛ-01-0.3 составляет всего 0.3 дБ/км. Дисперсия волн в световоде приводит к расплыванию световых импульсов и следовательно ограничивает быстродействие ВОСП. Вели-

чина распывания определяется дисперсионным коэффициентом β определяющим уширение импульса излучения с шириной спектра 1 нм на расстоянии 1 км. Величина β зависит от длины волны, материала и конструкции волокна. Пример типичного значения $\beta=3$ пс/ нм •км. Такие значения затухания и дисперсии позволяют реализовать ВОСП с очень высокими энергетическими показателями.

Рассматривая историю развития оптической связи следует прежде всего отметить, что световое излучение для создания сигналов на большие расстояния человек использовал уже очень давно. Однако современная история оптической связи фактически началась с изобретения в 1960 году первого лазера причем первые работы по лазерной связи связаны с использованием лазерного луча непосредственно в атмосфере. Были разработаны основные элементы таких систем связи, в частности, модуляторы лазерного излучения, приемно-передающая аппаратура, проведены исследования систем в различных диапазонах волн видимого и инфракрасного излучения. Перспективность оптических систем связи стала очевидной и основные трудности подобных систем были связаны с сильной зависимостью их работы от состояния атмосферы. Туман, дождь и другие атмосферные осадки приводили фактически к срыву связи, что естественно было недопустимо. Поэтому реальное применение такие оптические системы нашли только на очень коротких трассах, в космосе, закрытых производственных помещениях и какого-либо широкого применения не нашли.

Первые предложения применения оптического волокна для реализации линий связи были сделаны в 1966 году Као и Хокманом в США и одновременно Вертсом. Тогда это предложение носило чисто научный характер и казалось не имело какого-либо серьезного практического значения. Волноводы, изготовленные из имеющегося тогда стекла, обладали потерями превышающими 1000 дБ/км и это обстоятельство казалось непреодолимым препятствием. Затухание в 20 дБ/км было мечтой авторов идеи. Тем не менее практическая значимость таких систем была столь притягательной, что она сразу инициировала поток исследований в этой области и уже в 1970 г. американская фирма Corning Glass Works, являвшаяся известным производителем оптического стекла, нашла путь удаления примесей из материала волокна и поставила на рынок волокна с затуханием не превышающем 20 дБ/км. К 1976 г. величина потерь было уменьшено до 2 дБ/км. Особенно большие успехи в создании оптического волокна достигла Япония- 1979 г. было достигнуто рекордно малое затухание - всего 0.2 дБ/км. Однако и это не предел и работы по дальнейшему уменьшению потерь в волоконных линиях связи продолжают.

Интегральным параметром, определяющим уровень развития системы связи, является произведение длины линии связи l на её быстродействие V . Именно этим параметром $V \cdot l$ характеризуют обычно уровень достижений в развитии ВОСП. График этого параметра приведен на рис.1. Как видно из рис. 1, развитие ВОСП идет не только за счёт совершенствования

волоконных кабелей. Это важнейший, особенно на раннем этапе развития ВОСП, но только один из аспектов развития. В качестве других аспектов необходимо отметить следующие:

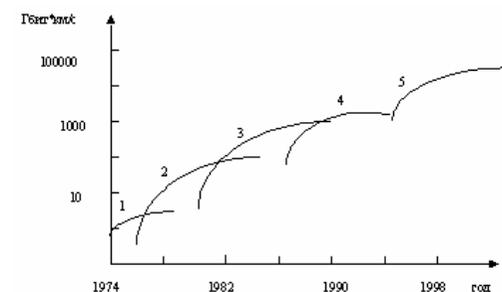


Рис. 1. Эволюция параметра $V \cdot l$ ВОСП.

Цифры у кривых характеризуют ВОСП и означают:

- 1 — $\lambda=0.8$ мкм, многомодовый кабель
- 2 — $\lambda=1.3$ мкм, одномодовый кабель
- 3 — $\lambda=1.55$ мкм, одномодовый кабель
- 4 — $\lambda=1.55$ мкм, одномодовый кабель с оптическими усилителями
- 5 — $\lambda=1.55$ мкм, одномодовый кабель с оптическими усилителями и спектральным уплотнением.

1. Быстрое развитие элементной базы электронных систем ВОСП. В частности, широкое внедрение арсенид-галлиевой технологии позволило создать транзисторы и микросхемы, работающие в СВЧ диапазоне обеспечивающие быстродействие (фронты импульсов) в десятки и даже единицы пс. Это очень важное достижение, без которого впечатляющих успехов ВОСП просто не было бы. В самом деле, уже в аппаратуре с быстродействием 2.5 Гбит/с необходимы фронты переключения импульсов не превышающие 100пс. В нашей стране серийно выпускаются цифровые микросхемы 6500 серии, имеющие время переключения не превышающее 200пс. Такие микросхемы вполне пригодны для использования в связной аппаратуре со скоростью передачи 622 Мбит/с. В этой связи отметим, что выпускаемая в России аппаратура в настоящее время имеет быстродействие не превышающее 622 Мбит/с. И хотя лабораторные образцы элементов, часто находятся на уровне самых высоких мировых достижений, технологические трудности не позволяют наладить их серийное производство.

2. Существенное значение имеет развитие лазерной техники. Созданы надёжные со сроком службы 10^5 и более часов полупроводниковые лазеры с выходной мощностью до 10 Вт и более и быстродействием в сотни пикосекунд. В последние годы всё более широкое внедрение получают оптические усилители, на основе эрбиевого волокна, позволившие качественно улучшить параметры системы связи. Разработаны удовлетворяющие требованиям практики приемники лазерного излучения. Эти достижения получены на основе результатов широкого

круга исследований в области квантовой электроники и лазерной техники во всех промышленно развитых странах мира.

3. Третий аспект развития ВОСП является системным. Он связан как с быстрым развитием цифровых систем передачи информации, так и с развитием аналоговых систем связи. Когда было впервые предложено использовать волокно для передачи информации, цифровых систем связи ещё не существовало. В России, например, цифровые методы передачи информации начали внедряться в практику, начиная только с середины семидесятых годов, тогда как за рубежом они появились на десять лет раньше. Между тем, именно цифровые методы наиболее подходят для применения в ВОСП. Поэтому именно развитие сетевых цифровых технологий, в значительной мере стимулировало развитие волоконно-оптических каналов связи в части повышения скорости передачи в них до 100 Мбит/с и выше. Такие скоростные технологии, как например Fast Ethernet, FDDI и др., полностью ориентированы на применение волоконной связи. Созданы глобальные сетевые технологии, такие, как отмеченные ранее сети SDH, новая технология асинхронной передачи ATM, которая принципиально может поддерживать скорость передачи от 1.5Мбит/с до 40Гбит/с. Одной из задач нашего курса является изучение существа организации этих новых перспективных технологий, в значительной степени определяющих уровень развития ВОСП. Отметим, что аналоговые методы передачи данных также могут использовать волокно. Такие системы находят применение прежде всего для передачи телевизионных сигналов, хотя цифровое телевидение на ВОСП имеет возможно ещё большие перспективы.

4. Нельзя не привести параллель в развитии волоконно-оптической связи и компьютерной техники. Эта параллель не случайна. Как персональные компьютеры, так и ВОСП используют одни и те же достижения микроэлектроники и ведут обработку и преобразование близких по формату цифровых данных. Поэтому все достижения компьютерной техники находят немедленное применение в ВОСП и обуславливают быстрое развитие цифровых систем связи. Функционирование современных ВОСП невозможно без самого широкого применения компьютеров и их узлов для решения разнообразных задач обеспечения связи.

7.8.2. Структура цифровых волоконно-оптических линейных трактов [2]

Современные волоконно-оптические системы передачи (ВОСП) обычно являются цифровыми. Для построения цифровых ВОСП используется аппаратура плейохронных, синхронных и асинхронных систем передачи. Кроме того, широкое внедрение оборудования WDM и DWDM способствовало распространению спектрального мультиплексирования в оптических линиях связи. Wavelength Division Multiplexing и Dense WDM, волновое мультиплексирование и плотное волновое мультиплексирование, сочетающееся с оборудованием оптического усиления. На рисунке 2.1 представлена структурная схема цифровой ВОСП, в которой

укрупнено отражены следующие возможные виды оборудования систем передачи.



Рис. 1. Структурная схема волоконно-оптической системы передачи

Мультиплексор, MUX (Multiplexer) – устройство обеспечивающее объединение / разделение нескольких низкоскоростных каналов передачи данных в один высокоскоростной канал на одном конце и восстановление низкоскоростных каналов на другом. Аппаратная реализация мультиплексора может быть электрической и оптической (WDM).

В электрической реализации нашли широкое распространение мультиплексоры технологий:

- PDH, Plesiochronous Digital Hierarchy – плейохронной цифровой иерархии;
- SDH, Synchronous Digital Hierarchy – синхронной цифровой иерархии;
- ATM, Asynchronous Transfer Mode – асинхронного режима передачи.

В оптической реализации применяются мультиплексоры оптических несущих для отдельных окон прозрачности оптического волокна (например: 1528.77 нм; 1529.55 нм; ...; 1560.61 нм – всего 41 частота). Частотный интервал между несущими определяет вид мультиплексирования: WDM и DWDM. Плотное мультиплексирование предлагает расстояние между соседними длинами волн не более 2 нм.

Оптический конвертор, OC (Optical Converter) – устройство, состоящие из оптического передатчика с преобразователями кода передачи и оптического приемника с преобразователями кода приема. В качестве оптического передатчика чаще всего применяются полупроводниковые лазеры (одномодовые и многомодовые), однако возможно применение различных светодиодов (поверхностных и торцевых). В качестве оптического приемника применяются фотодиоды типа р-і-п и лавинные (ЛФД). Преобразование линейного кода необходимо для помехоустойчивой передачи цифровых сигналов по оптической линии.

Оптический усилитель, OA (Optical Amplifier) – устройство, обеспечивающее увеличение мощности оптического излучения. В составе аппаратуры могут применяться полупроводнико-

вые и волоконные усилители. Оптические усилители могут использоваться в качестве усилителей мощности передачи, предусилителей на приеме и промежуточных усилителей вместо электронных регенераторов. Оптические усилители, как правило обеспечивают увеличение мощности оптических сигналов в широком спектре (в окне прозрачности) и могут использоваться совместно с оборудованием WDM.

Регенератор, Reg (Regeneration) – устройство, восстанавливающее длительность и амплитуду электрических импульсов. Регенератор используется в приемной части мультиплексора, иногда он является составной частью оптического конвертора, а также применяется на промежуточных станциях в виде отдельного устройства и совместно с мультиплексорами WDM.

Мультиплексоры ввода / вывода: оптический и электрический – **ADM, OADM** (Add / Drop Multiplexer, Optical ADM) – используются в качестве промежуточных станций для создания доступа к отдельным каналам передачи (электрическим и оптическим). Мультиплексоры ввода / вывода могут использоваться совместно с аппаратурой оперативного переключения (АОП) – кроссовыми коммутаторами.

Оптический кроссовый коммутатор, ОХС (Optical Cross - Connect) – устройство, обеспечивающее выделение, ввод, переключение оптических несущих в узловых станциях. ОХС служит для установления постоянных и полупостоянных оптических соединений между отдельными оконечными станциями или группами оконечных и промежуточных станций при создании оптических сетей.

Точки **А, В, С, D...I**, обозначенные на рисунке 1, называются точками сопряжения.

Если параметры этих точек определены стандартами международной организации по стандартизации, то их называют интерфейсами.

Использование различных видов оборудования в линиях передачи позволяет создавать различные каналы, тракты и секции.

Каналы передачи (transmission circuit) – комплекс технических средств и среды распространения, обеспечивающий передачу сигнала электросвязи в определенной полосе частот или с определенной скоростью передачи между сетевыми станциями. Каналы передачи подразделяют на аналоговые, цифровые и смешанные. Цифровые каналы в зависимости от скорости передачи, называют основным (64 кбит/с), первичным (2,048 Мбит/с), вторичным (8,448 Мбит/с), третичным (34,368 Мбит/с), четвертичным (139,264 Мбит/с). Каналы подразделяют на коммутируемые, арендуемые и виртуальные.

Тракт передачи (transmission path) – комплекс технических средств цифровой системы передачи, предназначенный для передачи стандартизованных блоков (циклов) данных и сигналов для их обслуживания.

Волоконно-оптический тракт – комплекс технических средств волоконно-оптической системы передачи, предназначенный для передачи оптических сигналов и сигналов обслужива-

ния тракта. В некоторых реализациях аппаратуры нет достаточно определенного выделения оптического тракта.

Секция (Section) – комплекс технических средств, предназначенный для передачи стандартизованных циклов (кадров, модулей) данных с соответствующими скоростями.

Секции подразделяют на:

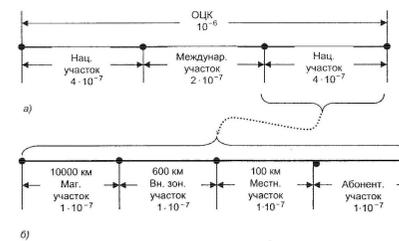
- секции мультиплексирования;
- секции регенерации;
- секции оптического мультиплексирования;
- секции оптической регенерации.

Любая из секций представляет собой объект технического обслуживания и снабжается сигналами контроля и управления.

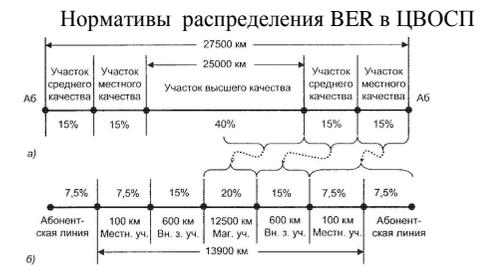
7.8.3. Помехоустойчивость цифровых ВОСП.

Помехоустойчивость цифровых систем передачи определяется способностью системы противостоять вредному действию помех. Воздействие помехи проявляется в том, что принятый цифровой сигнал может существенно отличаться от переданного. Поэтому помехоустойчивость характеризует степень соответствия принятого и переданного сообщений при заданном уровне помех. В бинарных системах в качестве меры помехоустойчивости используется вероятность ошибочного приема сигнала $P_{ош}$. В этом случае влияние помехи сводится к тому, что вместо фактически переданного двоичного символа, может быть принят альтернативный двоичный символ. На практике существует несколько способов измерения степени искажений цифрового потока данных. Наиболее распространенный подход состоит в статистической оценке числа ошибок. В этом случае используется битовый коэффициент ошибок (**bit error ratio - BER**), равный отношению числа ошибок N_e , происходящих в течении временного интервала t к общему числу символов N_t (единиц или нулей), переданных в течении указанного интервала. Обозначив битовую скорость передачи символом B , получим:

$$BER = \frac{N_e}{N_t} = \frac{N_e}{Bt},$$



Распределение коэффициента ошибок в ОЦК:



Распределение параметров качества сигнала в каналах:

а) международный участок
б) национальный участок

а) международный участок
б) национальный участок

7.8.4. Оценка коэффициента битовых ошибок

Ошибки возникают вследствие случайных флуктуаций сигнала. Виды ошибок:

- 1) пропуск сигнала ПС, при передаче «1» фиксируем «0»;
- 2) ложный сигнал ЛС.

$$P_{ПС} = P(0,1) = P(0) \cdot P(0/1)$$

$$P_{ЛС} = P(1,0) = P(1) \cdot P(1/0)$$

где $P(0), P(1)$ - априорные вероятности появления нуля или единицы на текущем ТИ; $P(0/1), P(1/0)$ - условные вероятности появления 0, при передаче 1 и 1 при передаче 0 соответственно. Так, что:

$$BER = P_{ПС} + P_{ЛС} = P(0) \cdot P(0/1) + P(1) \cdot P(1/0) \quad (1)$$

В общем случае вероятность $P(\alpha/\beta)$ определяются статистикой Пуассона:

$$P(n) = \frac{\langle n \rangle^n \cdot e^{-\langle n \rangle}}{n!}, \quad (2)$$

где $\langle n \rangle$ - среднее число сигнальных фотоэлектронов, зарегистрированных в течении времени T одного ТИ,

$$\langle n \rangle = \frac{P_0}{h\nu} \cdot \eta T \quad (3)$$

Известно, что в случае распределения Пуассона (2) его дисперсия совпадает с $\langle n \rangle$.

Из формулы (2) следует, что при любой заданной мощности оптического сигнала существует конечная вероятность пропуска сигнала ($P(0) \neq 0$), поскольку даже при отсутствии шумов число фотонов будет флуктуировать около величины $\langle n \rangle$. Эти флуктуации называют **квантовым шумом сигнала**.

В оптическом диапазоне энергия одного кванта велика, поэтому по сравнению с радиодиапазоном, для заданной мощности сигнала P_0 , число $\langle n \rangle$ относительно мало.

7.8.5. Помехоустойчивость ВОСП в условиях гауссовых шумов

При большом $\langle n \rangle$ случайные флуктуации сигнала на выходе главного усилителя ПРОМ, подчиняющиеся распределению Пуассона (2), переходит в распределение Гаусса:

$$P(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}}$$

где σ - дисперсия шума. В данных условиях Для обеспечения минимального значения BER, порог срабатывания компаратора S нужно выбирать из условия:

$$Q = \frac{S - u_0}{\sigma_0} = \frac{u_1 - S}{\sigma_1} = \frac{u_1 - u_0}{\sigma_1 + \sigma_0}, \quad (1)$$

Последнее соотношение определяет важный показатель помехоустойчивости ЦСП- так называемый **параметр Персонака**, **параметр поезозащищенности** или просто **Q - фактор**. Значения всех входящих в (1) величин можно определить из анализа осциллограммы (**глаз-диаграммы** (ГД)) цифрового сигнала на входе порогового устройства, которая представляет собой результат наложения всех возможных комбинаций принимаемых последовательностей в течении нескольких тайм-слотов (рис.1)

С помощью (1) уровень BER легко выражается через Q - фактор [1-3]:

$$BER = \frac{1}{2} \left[1 - \frac{2}{\sqrt{\pi}} \int_0^{Q/\sqrt{2}} e^{-x^2} dx \right], \quad (2)$$

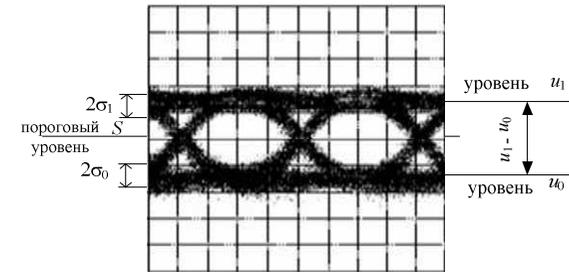


Рис.1. Глаз-диаграмма оптического сигнала

Для практических расчетов зависимости $BER(Q)$ широко используется аппроксимация:

$$BER \approx \frac{1}{\sqrt{2\pi}} \cdot \frac{\exp(-Q^2/2)}{Q} \approx 0.65 \exp[-0.443 \cdot (Q + 0.75)^2], \quad (3)$$

7.9. Оптические системы передачи со спектральным уплотнением

ВОСП со **спектральным уплотнением** (wavelength division multiplexing –WDM) по одному оптическому волокну одновременно передается несколько спектрально разнесенных оптических несущих, каждая из которых модулируется, как правило, многоканальным цифровым сигналом. Это одноволоконные многополосные одноканальные ВОСП. Возможность построения таких систем основывается на сравнительно слабой зависимости коэффициента затухания ОК в пределах используемого спектрального диапазона от частоты (или длины волны) оптической несущей. Поэтому, применяя метод частотного разделения, по одному ОБ, подобно многоствольным радиорелейным системам передачи, можно организовать несколько широкополосных оптических каналов, увеличив тем самым результирующую скорость пере-

дачи информации. Это позволяет обойти ограничение на скорость передачи по одному ОВ, связанное с дисперсионными искажениями. Структурная схема системы передачи со спектральным разделением оптических каналов показана на рисунке 9.

На передающей станции имеется n систем передачи, сигналы которых передаются на n оптических передатчиков ОПР, получающих оптические несущие с длинами волн $\lambda_1; \lambda_2; \dots; \lambda_n$. С помощью устройств спектрального объединения (УСО) осуществляется ввод различных оптических несущих в одно ОВ. На приемной стороне в устройстве спектрального разделения (УСР) оптические несущие разделяются в пространстве и поступают на оптические приемники ОПР. Таким образом, по одному ОВ организуется n спектрально разделенных оптических каналов, т. е. коэффициент использования пропускной способности увеличивается в n раз по сравнению с традиционным построением линейных трактов оптических систем.

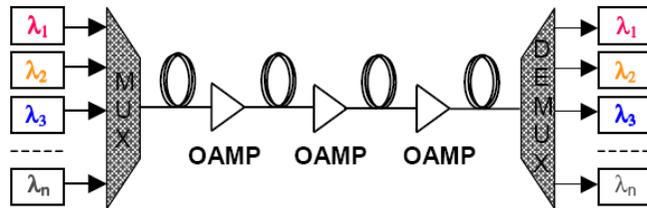


Рис. 1. Структурная схема ВОСП со спектральным уплотнением

Основными компонентами ВОСП со спектральным разделением являются мультиплексоры и демультиплексоры. Мультиплексоры подразделяются на спектрально- нечувствительные и спектрально-чувствительные. К первым относятся волоконные направленные ответвители, соединительные устройства, а также устройства на градиентных линзах. Вторые, называемые *мультиплексорами*, содержат элементы, характеристики которых зависят от оптической длины волны. К ним относятся дифракционные решетки, призмы, фильтры. Конструктивно мультиплексоры разделяются на объемные многоэлементные, объемные интегральные, планарные, волоконные, гибридные и др. На рисунке 2 показаны мультиплексоры: четырехканальный планарный на дифракционной решетке (рис. 4.10, а) и четырехканальный на интерференционных фильтрах (рис. 4.10, б).

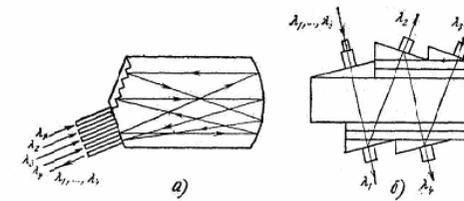


Рис.2. Мультиплексоры: а – четырехканальный планарный на дифракционной решетке; б – четырехканальный на интерференционных фильтрах

Демультиплексоры, входящие в состав систем со спектральным уплотнением, имеют структуру, аналогичную структуре спектрально-чувствительных мультиплексоров при обратном направлении распространения излучения. В настоящее время реализованы мультиплексоры и демультиплексоры с затуханием не более 3 дБ, спектральной шириной канала 5 нм и взаимным влиянием между каналами минус 25... 30 дБ.

WDM системы в настоящее время начали поистине триумфальное шествие в технику ВОСП. На их основе уже внедрены системы, обеспечивающие скорость передачи в одном волокне более 600 Гбит/с. В состоянии разработки находятся системы со скоростью передачи на одно волокно в несколько Тбит/с. При этом используются системы с очень плотным расположением спектральных каналов (системы DWDM – Dense – Wavelength division multiplexing). В таких системах спектральное расстояние между каналами составляет 0.8 нм или даже 0.4 нм, а число каналов может достигать 80. Если учесть, что каждый канал имеет скорость передачи не менее 2.5 Гбит/с, а в лучших разработках 10 Гбит/с и более, то отсюда ясны огромные скорости передачи, которые достигаются в системах DWDM. Естественно, что реализация таких возможностей связана с использованием самых передовых технологий, способных удовлетворить высоким требованиям предъявляемым к параметрам волокна, оптическим усилителям, к системам разделения и коммутации каналов и к другим компонентам, как оптической, так и электронной подсистемы. Но современные ведущие фирмы мира, такие как Lucent Technologies, Lightwave Microsystem (США), NEL Photonics Devices (Япония), OPTOCOM (Франция), NOCIA (Финляндия), IRE-POLUS (Германия – Россия) и другие успешно преодолевают эти трудности и DWDM-системы и их компоненты уже на рынке и в практике. Нет сомнения, что будущее телекоммуникаций именно за этими системами. Всем известен закон Мура: производительность электронных микросхем удваивается без изменения их цены каждые 18 месяцев. Но уже установлен закон Боско: производительность оптических устройств и систем удваивается также, но за 9 месяцев. Именно это обстоятельство позволили компании Lucent Technologies в мае 2000 года объявить о начале эры “Фотонной Долины” пришедшей на смену “Силиконовой долины”.

Отметим, что с целью повышения эффективности уже построенных оптических волноводов, сейчас активно разрабатывается еще один вариант систем со спектральным уплотнением – системы CWDM (Coarse Wavelength division multiplexing), которые по смыслу прямо противоположны DWDM-системам и предполагают большие спектральные расстояния (порядка 20 нм и более) между каналами.

8. КОМПЬЮТЕРНЫЕ СЕТИ

8.1. Принцип построения компьютерных сетей

Компьютерная сеть – это совокупность компьютеров и различных устройств, обеспечивающих информационный обмен между компьютерами в сети без использования каких-либо промежуточных носителей информации.

Все многообразие компьютерных сетей можно классифицировать по группе признаков:

- 1) Территориальная распространенность;
- 2) Ведомственная принадлежность;
- 3) Скорость передачи информации;
- 4) Тип среды передачи;

По территориальной распространенности сети могут быть локальными, глобальными, и региональными. Локальные – это сети, перекрывающие территорию не более 10 м², региональные – расположенные на территории города или области, глобальные на территории государства или группы государств, например, всемирная сеть Internet.

По принадлежности различают ведомственные и государственные сети. Ведомственные принадлежат одной организации и располагаются на ее территории. Государственные сети – сети, используемые в государственных структурах.

По скорости передачи информации компьютерные сети делятся на низко-, средне- и высокоскоростные.

По типу среды передачи разделяются на сети коаксиальные, на витой паре, оптоволоконные, с передачей информации по радиоканалам, в инфракрасном диапазоне.

Компьютеры могут соединяться кабелями, образуя различную топологию сети (звездная, шинная, кольцевая и др.).

Следует различать компьютерные сети и сети терминалов (терминальные сети). Компьютерные сети связывают компьютеры, каждый из которых может работать и автономно. Терминальные сети обычно связывают мощные компьютеры (майнфреймы), а в отдельных случаях и ПК с устройствами (терминалами), которые могут быть достаточно сложны, но вне сети их работа или невозможна, или вообще теряет смысл. Например, сеть банкоматов или касс по продаже авиабилетов. Строятся они на совершенно иных, чем компьютерные сети,

принципах и даже на другой вычислительной технике.

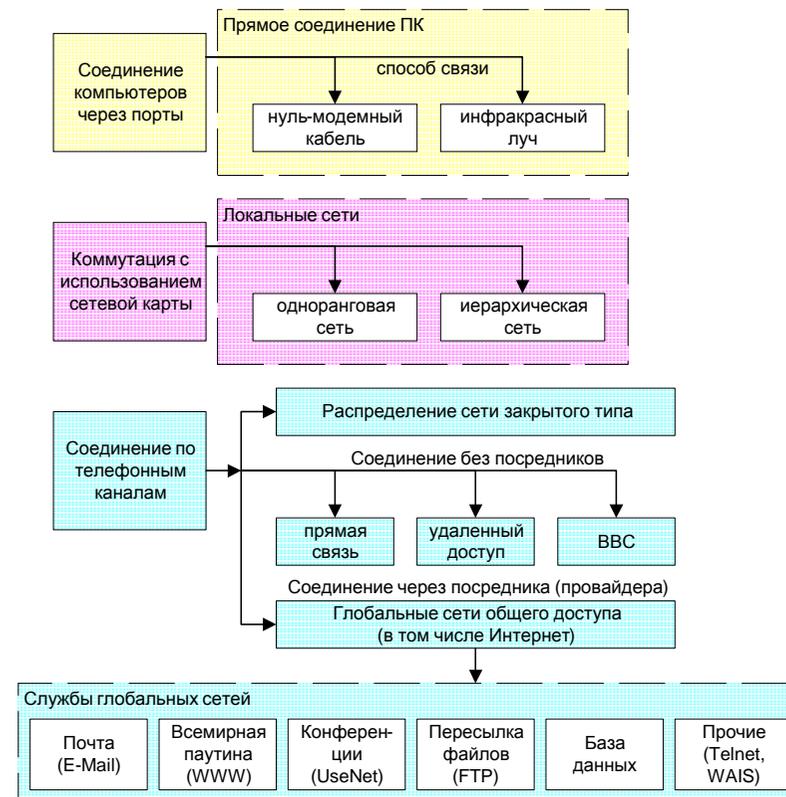


Рис. 1. Способы коммутации компьютеров и виды сетей.

В классификации сетей существует два основных термина: LAN и WAN.

LAN (Local Area Network) – локальные сети, имеющие замкнутую инфраструктуру до выхода на поставщиков услуг. Термин «LAN» может описывать и маленькую офисную сеть, и сеть уровня большого завода, занимающего несколько сотен гектаров. Зарубежные источники дают даже близкую оценку – около шести миль (10 км) в радиусе; использование высокоскоростных каналов.

WAN (Wide Area Network) – глобальная сеть, покрывающая большие географические регионы, включающие в себя как локальные сети, так и прочие телекоммуникационные сети и устройства. Пример WAN – сети с коммутацией пакетов (Frame Relay), через которую могут «разговаривать» между собой различные компьютерные сети.

Термин «корпоративная сеть» также используется в литературе для обозначения объединения нескольких сетей, каждая из которых может быть построена на различных техниче-

ских, программных и информационных принципах.

Рассмотренные выше виды сетей являются сетями закрытого типа, доступ к ним разрешен только ограниченному кругу пользователей, для которых работа в такой сети непосредственно связана с их профессиональной деятельностью. Глобальные сети ориентированы на обслуживание любых пользователей.

На рисунке 1, рассмотрим способы коммутации компьютеров и виды сетей.

8.2. Локальные вычислительные сети (ЛВС)

8.2.1. Основные типы конфигурации ЛВС

Локальной – ЛВС называется вычислительная сеть, системы которой расположены на небольшом расстоянии друг от друга. При таких расстояниях наиболее эффективное средство

связи между системами, объединенными в сеть, - последовательный интерфейс [9]. В качестве передающей среды используются витая пара проводников, коаксиальные и волоконно-оптические кабели, при этом в большинстве случаев достаточно иметь один канал – моноканал, обслуживающий все системы ЛВС в режиме мультиплексирования [17]. На основе моноканала строятся две основные конфигурации (рис. 3.1, а и б):

- магистральная (шинная);
- кольцевая.

В ЛВС могут использоваться и другие конфигурации связей между системами – звездообразная и произвольная.

При звездообразной в качестве центрального узла могут выступать ЭВМ или узел связи, осуществляющий коммутацию данных между системами сети.

Произвольная конфигурация обычно представляет собой объединение нескольких магистральных или кольцевых структур. Для объединения систем могут использоваться узлы связи, осуществляющие коммутацию сообщений, как в глобальных сетях.

Организация ЛВС базируется на принципах многоуровневого управления процессами (рис. 2). Уровни и протоколы ЛВС интерпретируются в терминах семиуровневой архитектуры открытых систем.

Протокол управления физическим каналом (УФК) определяет форму представления и порядок передачи данных через физический канал. Он вводит интерфейс, устанавливающий стандартный для всех систем порядок взаимодействия с каналом.

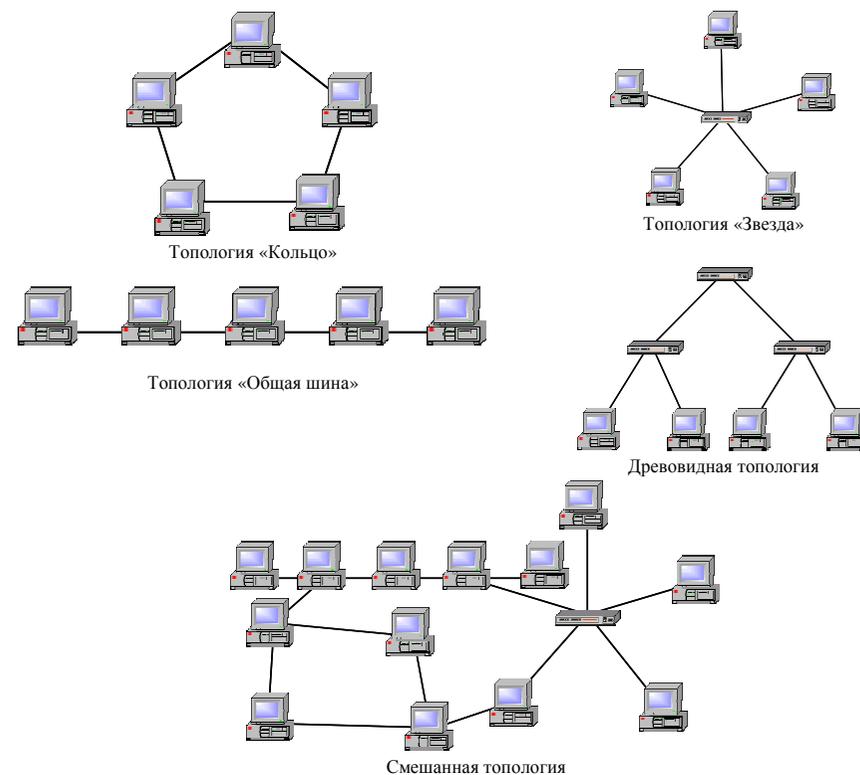


Рис.1. Основные виды концентрации моноканала

Канальный уровень разделен на два подуровня(1 и 2):

- управление доступом к каналу (УДК);
- управление информационным каналом (УИК).

Порядок их функционирования определяется соответствующими протоколами.

Протокол УДК устанавливает процедуру передачи данных через канал и процедуру селекции данных, принимаемых из канала.

Протокол УИК устанавливает порядок обеспечения достоверности данных при передаче данных через физический канал.

В большинстве ЛВС отсутствует необходимость в сетевом уровне управления, поэтому этот уровень отсутствует.

Протокол управления передачей (УП) обеспечивает единый транспортный интерфейс для процессов высокого уровня.

Протоколы высокого уровня - управления сеансами (УС), представлением данных (УПД) и прикладными процессами (УПП) по своим функциям аналогичен соответствующим прото-

колам глобальных сетей.



Рис. 2. Многоуровневая система управления ЛВС

При использовании моноканала сложность процедур управления на физическом, канальном и транспортном уровне невелика. Поэтому эти уровни управления достаточно эффективно реализуются техническими средствами – специальным контроллером, называемым сетевым адаптером СА ЛВС (рис. 3). Адаптер сопрягает внутреннюю магистраль или интерфейс ввода-вывода ЭВМ с последовательным интерфейсом. Адаптер в совокупности с физическим каналом образует информационный моноканал, к которому подключаются системы сети.

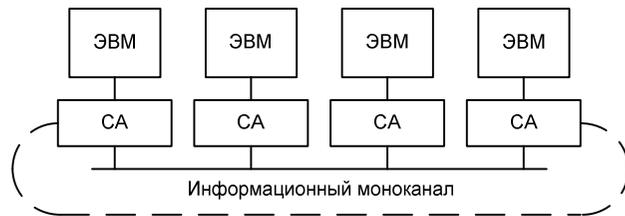


Рис. 3. Сетевые адаптеры и информационный моноканал ЛВС

По принципу предоставления ресурсов в совместное пользование сети разделяют на три типа: одноранговые, на основе сервера и комбинированные.

В одноранговой сети все компьютеры равноправны: нет иерархии и нет выделенного сервера. Каждый компьютер функционирует и как клиент и как сервер. Все пользователи решают, какие данные на своем компьютере сделать общедоступными. Одноранговые сети иногда называют рабочей группой, в которой участвуют не более 10 компьютеров одновременно.

Сети на основе сервера имеют выделенный сервер, который функционирует только как сервер. Он специально оптимизирован для быстрой обработки запросов от сетевых клиентов и для управления защитой файлов и каталогов. В сети могут использоваться несколько серверов:

- файл-серверы и принт-серверы управляют доступом пользователей соответственно к

файлам и принтерам;

- серверы приложений выполняют прикладные части клиент-серверных приложений, в них также находятся данные доступные клиентам;
- почтовые серверы управляют передачей электронных сообщений между пользователями сети;
- коммутационные серверы, управляют потоком почтовых сообщений между сетью и другими сетями.

Комбинированные сети совмещают в себе свойства одноранговых сетей и сетей на основе сервера. Операционные системы для сетей на основе сервера, в этом случае отвечают за совместное использование основных приложений и данных. На компьютерах клиентов применяются операционные системы, управляющие доступом к ресурсам сервера и в тоже время предоставляющие в совместное использование свои ресурсы.

8.2.2. Методы доступа к моноканалу

Моноканал ЛВС представляет собой разделяемый ресурс, использование которого основано на различных методах доступа. В литературе [32] все методы доступа к разделяемому каналу передачи данных классифицируются с позиций детерминированности и случайности доступа к нему. (рис.3.4).

В зависимости от способа предоставления права на пользование каналом методы первой группы могут быть централизованными и децентрализованными. В первом случае санкцию выдает центральный узел в соответствии с запросами рабочих станций, во втором каждая рабочая станция определяет право на пользование каналом самостоятельно. Аналогично и методы второй группы способов могут быть централизованные и децентрализованные.

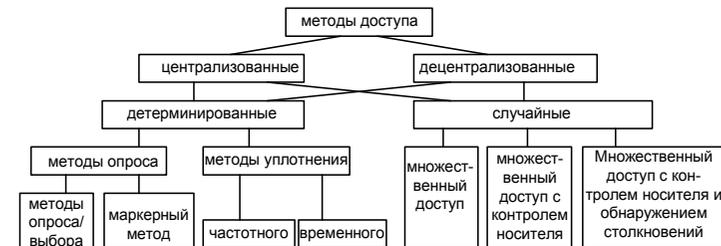


Рис. 4. Классификация методов доступа

Методы опроса является самым простым методом передачи пакетов. При этом первичный узел периодически опрашивает по списку вторичные узлы последовательно один за другим на предмет наличия запроса на передачу.

Одним из способов реализации этого метода может служить работа системы с опро-

сом/выбором. С помощью команды Опрос осуществляется передача данных первичному узлу. Назначение команды Выбор противоположное - передавать данные из первичного узла во вторичный. Первичный узел осуществляет контроль ошибок и посылает положительный ответ АСК при отсутствии ошибок и отрицательный ответ NAC при обнаружении их. Передача АСК/NAC может происходить до тех пор, пока данные не будут полностью переданы и не будет послано уведомление EOT(End Of Transmission) о том, что передача закончена. Система опроса/выбора может быть использована в режиме соединения «точка-многоточие» для разделения пропускной способности канала между двумя вторичными станциями. Недостатком системы с опросом/выбором является наличие неоднократных реакций на опросы, что приводит к непроизводительному потреблению дорогостоящих ресурсов канала.

Система передачи маркера обычно реализуется в сетях с кольцевой топологией или с общей шиной. (Иногда такие системы называются системами с эстафетным доступом). При включении в систему генерируется эстафета - специальная кодовая последовательность, которая передается по кольцу. Если узел имеет данные для передачи, он вынужден ждать, пока предыдущий узел не передаст ему эстафету.

Эстафетный доступ в моноканальных структурах реализуется по следующей схеме (рис. 5)

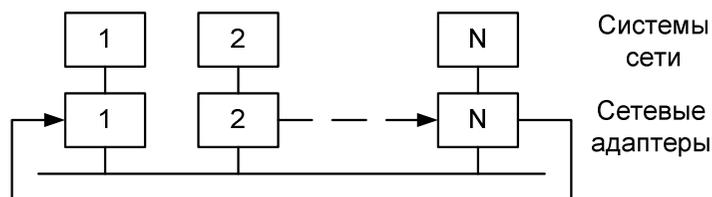


Рис. 5. Эстафетный доступ в ЛВС

Сетевые адаптеры связаны кольцевой цепью, по которой передается эстафета-сигнал, разрешающий доступ к моноканалу.

Если в адаптере пакет на передачу отсутствует, этот адаптер передает эстафету следующему, если пакет есть, то эстафета хранится до окончания передачи пакета.

В кольцевых структурах эстафетный доступ к каналу реализуется с использованием эстафеты-маркера, последовательно передаваемого по кольцу от одной системы к другой. В качестве маркера используется бит, выделенный в последовательности начала кадра. Если в систему приходит маркер, она изымает его из кольца и посылает в кольцо кадр, который последовательно проходит все системы. Каждая система сравнивает принятый адрес с собственным и, если адрес совпадает, принимает его и транслирует дальше, установив в признаке конца кадра в 1 бит приема. Передающая система находится в передающем состоянии до тех пор, пока не

получит обратно свой кадр. После его анализа она пересылает маркер следующей системе, которая по его получении имеет право на передачу своего кадра. Если маркер потерян, то по прошествии тайм-аута он генерируется вновь специально выделенной для этого системой.

К недостаткам эстафетных систем можно отнести затраты времени на перемещение эстафеты, предоставляющей право на пользование шиной.

Методы уплотнения разделяются на временные и частотные. Временное уплотнение реализуется при помощи мультиплексора, предоставляющего каждому из подключенных к общему каналу низкоскоростных устройств один временной такт, в течение которого это устройство получает в свое монопольное пользование быстродействующий канал. Применение метода временного уплотнения, при котором используются фиксированные временные такты эффективно, если каждое из устройств постоянно передает или принимает информацию. Такое на практике случается не часто. Для того чтобы избежать непроизводительных простоев используется метод статистического временного уплотнения, основанный на том, что временные такты предоставляются устройству лишь тогда, когда оно в них действительно нуждается. Статистический мультиплексор должен быть в достаточной степени «интеллектуальным» для того, чтобы выполнить функцию уплотнения в том случае, когда несколько устройств требуют выхода в общий канал.

При использовании частотного уплотнения широкая полоса пропускания некоторой среды передачи разделяется на некоторое число индивидуальных каналов. Частотное уплотнение предоставляет возможность установления связи между несколькими устройствами в данной полосе частот независимо от взаимодействия в других диапазонах.

Методы случайного доступа реализуются в логической топологии общего типа. Простейшей системой такого доступа является множественный доступ. Если у некоторого узла имеется такой, готовый к отправке, то он передает этот пакет, не обращая внимания на то, занят канал в данный момент или нет. По завершению передачи пакета запускается таймер. Если по истечении определенного времени не получено подтверждение о приеме пакета адресатом, то считается, что произошло столкновение, и терминал повторяет передачу того же пакета. Для уменьшения вероятности повторного конфликта между теми же пакетами, интервал, через который производится передача пакета, задается случайным образом. Преимущество такой системы состоит в простоте ее реализации, а недостаток - в очень низком коэффициенте использования канала. Одним из способов повышения производительности сети является тактирование. При этом центральный узел формирует серию последовательных временных тактов (слот-тайм), и передача пакета осуществляется только в начале каждого такта. Подобный прием позволяет почти в двое увеличить коэффициент использования канала.

Множественный доступ с контролем носителя (несущий, МНДК) позволяет уменьшить вероятность столкновения пакетов. Посылка пакета начинается после освобождения среды

передачи от носителя информации. Столкновения в системе МНДК возможны лишь в том случае, когда два или более узла одновременно пытаются переслать пакет сразу после освобождения канала. Поэтому существуют различные способы начала передачи пакета. При этом системы МНДК подразделяются на системы I-, N- и P-типа.

В системах I-типа передача пакета начинается сразу после освобождения тракта передачи. Вероятность столкновения достаточно высока.

В системах N-типа, если канал оказывается занятым, передача пакета откладывается на более поздний момент, чем освобождение такта передачи. Вероятность возникновения столкновений в такой системе незначительна, однако, существенно возрастает вероятность простоя канала, а коэффициент использования тракта передачи остается в целом невысоким.

В системе P-типа после освобождения канала посылка пакета начинается с вероятностью p . Если известна зависимость между вероятностью появления запроса на передачу пакета и длительностью передачи, то можно определить оптимальное значение вероятности p .

Система МДКН I-типа, в которой дополнительно предусмотрено обнаружение столкновений называется множественным доступом с контролем носителя и обнаружением столкновений - МДКН/ОС (CSMA/CD). Принцип обнаружений столкновений поясняется рисунком 3.6. Метод оказался очень эффективным на практике, в реальных сетях, реализующих этот метод, не требуется специальных подтверждений приема отдельных пакетов.

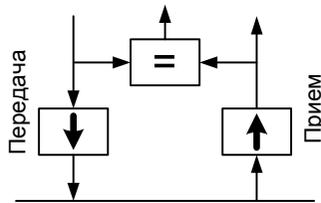


Рис. 6. Обнаружение столкновений пакетов

Информационный канал строится на основе физического канала, дополняя последний средствами обеспечения достоверности данных, передаваемых в форме пакетов между узлами ЛВС. Порядок функционирования информационного канала задается протоколом, который относится ко второму подуровню уровня 2 и определяет формат пакетов, средства контроля данных и исправления ошибок, вносимых в данные при передаче пакета по каналу. В ЛВС используется дейтограммный способ передачи данных, т.е. каждый пакет рассматривается как независимый объект, передаваемый между узлами сети. Пропускная способность моноканала, как правило, превышает потребную для сети, и поэтому нет необходимости в уменьшении служебных полей, что упрощает процедуры и средства формирования и приема пакетов.

Структура кадра представлена на рис. 7.

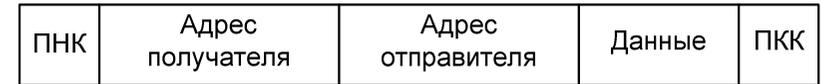


Рис. 3.7. Структура кадра информационного канала

ПНК состоит, как правило, из 2-16 бит, используемых в качестве флага начала пакета и поля управлением доступом к моноканалу. Поле данных имеет различную форму - обычно от 4 до 128 (512) байт. Корректность данных, содержащихся в пакете, может контролироваться разными способами: по четности байтов или с использованием 16- и 32-разрядных циклических кадров. В последовательности конца кадра выделяются поля для кода циклической суммы, признака доставки пакета, признака продолжения в следующем пакете и др.

Система, принимающая пакет, проверяет его корректность с помощью средств контроля достоверности данных. При обнаружении ошибки производится повторная передача пакета в порядке, определенном протоколом, например с использованием квитанции и тайм-аута. В кольцевых сетях пакет возвращается отправителю с отметкой о приеме.

8.2.3. Адаптеры и приемопередатчики ЛВС

Сетевые адаптеры обеспечивают сопряжение узлов ЛВС с моноканалом и реализуют протоколы канального уровня: управления физическим каналом, доступом к каналу и информационным каналом [18]. Структура адаптера имеет следующий вид (см. рис.3.8).

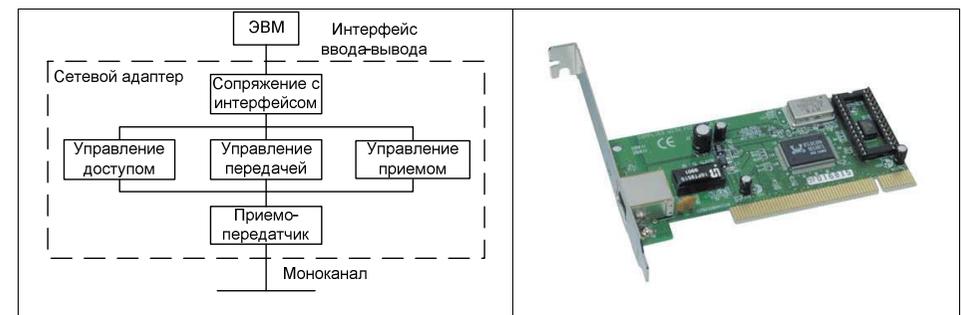


Рис.8. Структура адаптера ЛВС

Приемопередатчик согласует логические сигналы, формируемые в адаптере, с физическими сигналами в моноканале. Блок управления доступом выполняет протокол доступа к моноканалу. Блок управления передачей обеспечивает вывод на приемопередатчик последовательности битов, соответствующей пакету. Блок управления приемом анализирует пакеты, пе-

редаваемые через моноканал, и выделяет пакеты, адресованные узлу. Адаптер связывается с ЭВМ с помощью блока сопряжения через соответствующий интерфейс.

При использовании в моноканале витой пары или коаксиального кабеля надежность канала существенно зависит от организации электропитания и заземления приемопередатчиков. Радикальный способ разделения элементов сети по питанию - установка в соответствующих цепях развязывающих элементов - трансформаторов или оптронных пар (рис. 9.). В магистральных ЛВС развязывающие элементы (РЭ) устанавливаются между приемопередатчиком и остальной частью адаптера, а в кольцевых - в конце каждого сегмента кабеля, соединяющего соседние адаптеры.

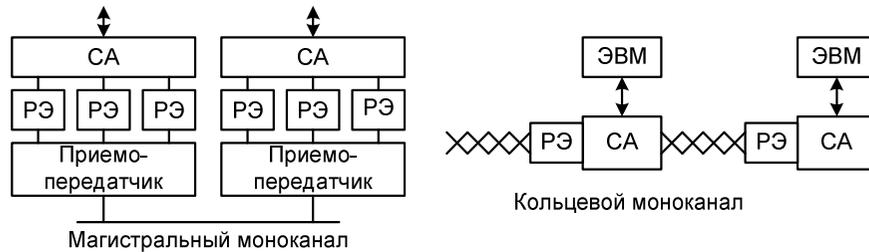


Рис. 9. Развязывающие элементы в ЛВС

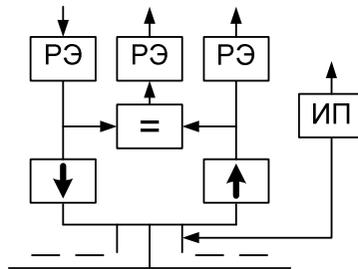


Рис. 10. Развязывающие элементы на оптронах

Если в качестве развязки используются оптроны, то приемопередатчик имеет собственный источник питания ИП, изолированный от земли (см. рис. 10). Нулевая шина приемопередатчика соединяется с экраном коаксиального кабеля, что существенно снижает помехи, возникающие на протяженной шине. Приемопередатчики магистральных сетей должны иметь высокое сопротивление по отношению к каналу чтобы изменение числа систем (адаптеров), подключенных к каналу, не вызывало значительного изменения нагрузки.

Пример приемопередатчика для кольцевой сети приведен на рис.11. Сегмент канала, выполненный из скрученной пары, соединяет передатчик одного адаптера с приемником сле-

дующего. На выходе приемника формируются логические сигналы, представляющие принимаемые биты.

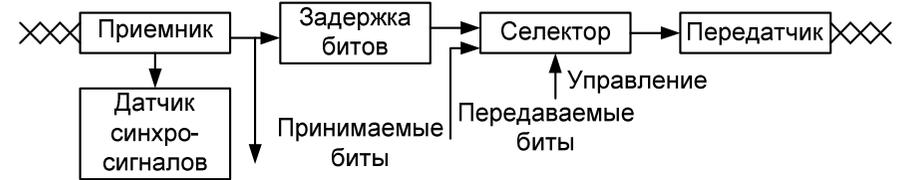


Рис. 11. Приемопередатчик кольцевой ЛВС

Кроме того, из сигналов, поступающих на приемник, формируются синхросигналы. Для обработки поступающей информации (например, установки бита приема в единицу) в цепь ретрансляции введена задержка. Селектор служит для управления работой передатчика, через него могут подаваться либо ретранслируемые, либо передаваемые биты. В кольцевой сети каждый адаптер, т.е. его приемопередатчик, должен обеспечивать трансляцию сигналов даже при выключенной ЭВМ. Поэтому в сети имеются специальные источники питания, к которым подключаются только приемопередатчики сетевых адаптеров.

8.2.4. Управление передачей кадров

Наиболее существенное влияние на организацию адаптеров оказывает способ обмена данными между адаптером и ЭВМ. Могут использоваться два способа:

- без буферизации;
- с буферизацией пакетов.

По первому способу пакет, подлежащий передаче, размещается в оперативной памяти ЭВМ, откуда передается в адаптер в виде последовательности слов. Адаптер преобразует каждое поступившее слово в последовательность битов. После передачи битов слова в моноканал адаптер через интерфейс обращается к оперативной памяти ЭВМ за следующим словом. При приеме пакета в адаптере из поступающих битов формируются слова, которые последовательно отсылаются в заданную область оперативной памяти ЭВМ в темпе формирования слов.

По второму способу в адаптере организуется буферное запоминающее устройство для хранения передаваемых и принимаемых пакетов. Передаваемый или принимаемый пакеты накапливаются сначала в буферной памяти адаптера.

Управление передачей пакетов реализуется следующим образом (рис. 12). Из ЭВМ в блок сопряжения с интерфейсом через канал передачи кадра заносятся адрес начала области оперативной памяти (в которой хранится пакет), а также длина пакета в битах. Блок сопряжения считывает из оперативной памяти первое слово пакета, которое передается в буфер вывода и затем в преобразователь параллельного кода в последовательный. Передача кадра начинается

с посылки в передатчик последовательности начала кадра (ПНК). Затем через схему бит-стаффинга выводится последовательность битов, составляющих пакет.

По окончании передачи хранимого в преобразователе параллельного кода в последовательный слова, в него загружаются новые слова. Процесс продолжается до тех пор, пока не будет передано заданное число битов, что отмечается сигналом U4, формируемым счетчиком длины. При передаче данных определяется контрольная сумма, которая передается вслед за данными, после чего выводится последовательность конца кадра (ПКК).

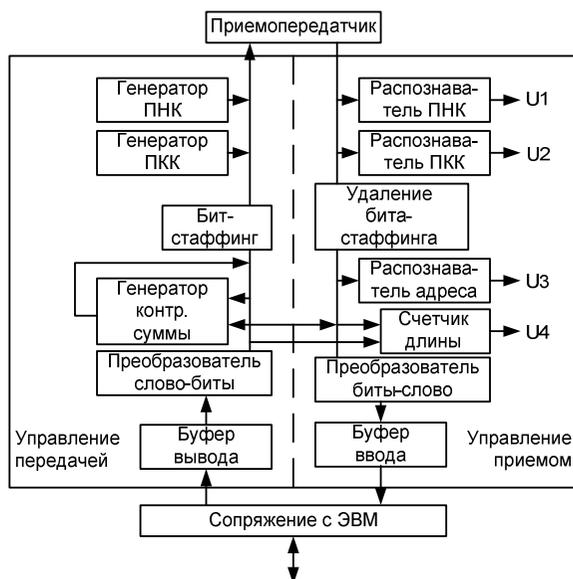


Рис. 12. Управление передачей кадров

Готовность адаптера к приему кадра обеспечивается передачей адреса оперативной памяти, выделенной для размещения принимаемого пакета из ЭВМ в блок сопряжения. Распознаватель последовательности начала кадра формирует сигнал U1, отличающий начало пакета, передаваемого в кадре. Последующие биты обрабатываются схемой удаления бит-стаффинга и поступают на схему распознавания адреса. Совпадения адресов отмечаются сигналом U3. Принимаемые биты поступают на преобразователь последовательного кода в параллельный. Сформированные слова передаются в буфер вывода. Биты, составляющие пакет, формируются на счетчике длины и обрабатываются генератором контрольной суммы. Счетчик фиксирует длину принятого пакета и формирует сигнал U4, если длина пакета превосходит допустимую.

Прием кадра и передача содержащегося в нем пакета заканчиваются при поступлении последовательности конца кадра, что отмечается сигналом U2. Сигнал U4 используется для пре-

ращения приема пакета, имеющего недопустимую длину. При передаче данных определяется контрольная сумма, которая передается вслед за данными, после чего выводится последовательность конца кадра (ПКК).

Готовность адаптера к приему кадра обеспечивается передачей адреса оперативной памяти, выделенной для размещения принимаемого пакета из ЭВМ в блок сопряжения. Распознаватель последовательности начала кадра формирует сигнал U1, отличающий начало пакета, передаваемого в кадре. Последующие биты обрабатываются схемой удаления бит-стаффинга и поступают на схему распознавания адреса. Совпадения адресов отмечаются сигналом U3. Принимаемые биты поступают на преобразователь последовательного кода в параллельный. Сформированные слова передаются в буфер вывода. Биты, составляющие пакет, формируются на счетчике длины и обрабатываются генератором контрольной суммы. Счетчик фиксирует длину принятого пакета и формирует сигнал U4, если длина пакета превосходит допустимую.

Прием кадра и передача содержащегося в нем пакета заканчиваются при поступлении последовательности конца кадра, что отмечается сигналом U2. Сигнал U4 используется для прекращения приема пакета, имеющего недопустимую длину.

8.2.5. Стандарты ЛВС

Разработка ЛВС началась в 70-х годах. К настоящему времени создано большое число их вариантов, отличающихся типом передающей среды, пропускной способностью моноканала, структурой и назначением. Типичными областями применения ЛВС являются: системы административного управления (офисные системы), автоматизация проектирования и технологической подготовки производства, автоматизация научных исследований, промышленные и бортовые системы. К последним предъявляются требования повышенной надежности при работе в условиях сильных помех и, возможно, в широком диапазоне температур окружающей среды.

Разнообразие структурных и технических решений насущно требовали приведения их к единому стандарту [19]. В 1981 году специалисты Института инженеров по электротехнике и радиоэлектронике (IEEE) США подготовили проект стандарта локальных сетей.

Стандартизацией охватываются два нижних уровня управления физическим и информационным каналом [20].

В основу управления информационным каналом положен протокол HDLC - стандартный протокол второго уровня сетей при передаче данных с коммутацией пакетов. Формат кадра имеет вид рис. 3.13.

Адреса назначения и источника кодируются одним или несколькими байтами (до 7). Первый разряд байта называется битом расширения адреса. Если он равен 1, то байт последний, если 0, то адрес продолжается в следующем байте. Второй бит адреса назначения L/G - идентификатор типа адреса, L/G=0 - индивидуальный адрес, L/G=1 - групповой. Если в адресе все

единицы - глобальный адрес. Бит C/R адреса источника используется для идентификации команды или ответа: C/R=0 - команда; C/R=1 - ответ. Байт управления используется для обозначения типа кадра.



Рис. 13. Формат кадра стандарта ЛВС

Применяется три типа кадров: I - информационный, C - командный, U - нумерованный. Тип кадра определяется значением одного или двух первых битов поля управления. N(S) - порядковый номер передаваемого кадра по модулю 8, N(K) - порядковый номер следующего принимаемого кадра (нумерация может быть как нормальной, так и расширенной - в этом случае поле управления является 2-х байтовым), P/F - бит запроса/ответа.

Управляющие кадры типа C предназначены для передачи подтверждений и запросов на повторную передачу. Поле CC указывает функцию управления. Нумерованные кадры используются для реализации дополнительных функций управления. Область данных может содержать любое целое число байтов, а в управляющих и нумерованных может отсутствовать. КЦС - код циклической суммы, служит для проверки корректности кадра. Проект позволяет осуществлять два способа доступа к моноканалам: свободный с проверкой несущей и столкновений и эстафетный доступ.

В нумерованных типа U кадрах выполняемая команда определяется полем MM.

Предусматривается 4 скорости передачи данных: 1, 5, 10 и 20 мегабит в секунду. В качестве передающей среды допускается использовать витые пары, коаксиальные или волоконно-оптические кабели.

Более 250 западных фирм выпускают в настоящее время различные марки ЛВС, в той или иной степени соответствующие данному стандарту. В последующих разделах рассмотрим основные особенности наиболее популярных типов ЛВС, получивших распространение в России.

8.2.6. Технология Ethernet

Ethernet самый распространенный на сегодняшний день стандарт локальных сетей. В

широком смысле, под этим словом понимается любой из вариантов этой технологии. На основе изначально существовавших вариантов реализации этой технологии был разработан стандарт IEEE 802.3, в котором функции протокола разделены на уровни MAC и LLC. В зависимости от типа физической среды стандарт IEEE 802.3 имеет различные модификации - 10 Base-5, 10Base-2, 10Base-T, 10Base-FL, 10 Base-FB.

В 1995 году был принят стандарт Fast Ethernet, который во многом не является самостоятельным стандартом и описан в специальном разделе стандарта IEEE 802.3, так же как и принятый в 1998 году стандарт Gigabit Ethernet.

Все виды стандартов Ethernet (в том числе Fast Ethernet и Gigabit Ethernet) используют один и тот же метод разделения среды передачи данных - метод CSMA/CD.

Соответствие уровней протоколов стандарта IEEE стандартным уровнем семиуровневой системы ISO/OSI приведено на рис. 14.

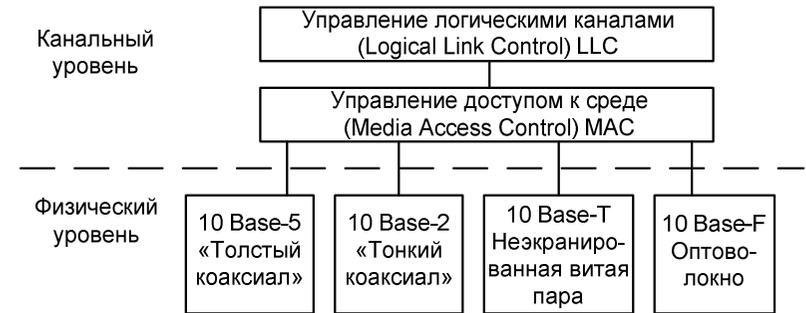


Рис. 14. Соответствие уровней протоколов IEEE и ISO/OSI

Протокол LLC обеспечивает для технологий локальных сетей нужное качество услуг транспортной службы, передавая свои кадры либо дейтограммным способом, либо с помощью процедур с установлением соединения и восстановления кадров. Протокол LLC передает свой кадр вместе с адресной информацией протоколу уровня MAC, который упаковывает его в свой кадр. В основу протокола LLC положен протокол HDLC (High level Data Link Control), являющийся стандартом ISO и послуживший основой для проекта стандарта ЛВС.

На MAC - уровне для идентификации сетевых интерфейсов узлов сети используются регламентированные стандартом IEEE 802.3 уникальные 6-байтовые адреса, называемые MAC - адресами. Обычно MAC - адрес записывают в виде шести пар шестнадцатеричных цифр, разделенных тире или двоеточием например 11-A0-17-3D- BC- 01.

Помимо отдельных интерфейсов, MAC - адрес может определять группу интерфейсов или даже все интерфейсы сети. Младший бит старшего байта адреса назначения является признаком того, является адрес индивидуальным или групповым. Если он равен 0, то адрес

является индивидуальным, а если 1, то это групповой адрес. Групповой адрес связан только с теми интерфейсами, которые сконфигурированы как члены группы, номер, которой указан в групповом адресе. Если адрес состоит из всех единиц, он идентифицирует все узлы сети и называется широковещательным адресом.

Второй бит старшего адреса определяет способ назначения адреса - централизованный или локальный. Если этот бит равен 0, то адрес назначен централизованно. Комитет IEEE распределяет между производителями оборудования для сетей так называемые организационно уникальные идентификаторы (Organizationally Unique Identifier, OUI). Каждый производитель помещает выделенный ему идентификатор в три старших байта адреса. Уникальность централизованно распределенных адресов распространяется на все основные технологии сетей. Локально администрируемые адреса назначаются администратором сети, в обязанности которого входит обеспечение их уникальности.

Физические специализации технологии Ethernet по стандарту IEEE 802.3 включает в себя следующие среды передачи данных:

- 10 Base-5 - коаксиальный кабель диаметром 0,5" («толстый» коаксиал) с волновым сопротивлением 50 Ом и максимальной длиной сегмента 500 м;
- 10 Base-2 - коаксиальный кабель диаметром 0,25" («тонкий» коаксиал) с волновым сопротивлением 50 Ом и максимальной длиной сегмента 185 м;
- 10 Base-T - кабель с неэкранированной витой парой VTP, образующий звездообразную топологию на основе концентратора, расстояние между концентратором и конечным узлом не более 100м;
- 10 Base-F - волоконно-оптический кабель с топологией аналогичной стандарту 10 Base-T. Существует несколько вариантов этой спецификации: FOIRL, 10 Base-FL и 10 Base-FB.

Стандарт FIORL (Fiber Optic Inter-Repeater Link)- первый стандарт для использования оптоволоконна в сетях Ethernet. Длина оптоволоконной связи между повторителями до 1 км.

Стандарт 10 Base-FL - улучшение стандарта FIORL. Повышена мощность передатчиков, поэтому максимальное расстояние между узлом и концентратором увеличилось до 200 м.

Стандарт 10 Base-FB предназначен только для соединения повторителей. Повторители, соединенные по стандарту 10 Base-FB, при отсутствии кадров для передачи постоянно обеспечиваются специальными последовательностями сигналов, отличающихся от сигналов кадров данных, для поддержания синхронизации, поэтому повторители вносят минимальные задержки при передаче данных из одного сегмента в другой, что позволяет увеличить их число.

Использование метода доступа CSMA/CD налагает ряд ограничений на устройства и кабельную систему сетей.

В сегменте-домене коллизий не может находиться более 1024 устройств.

Домен коллизий - это часть сети Ethernet, все узлы которого распознают коллизии независимо от того, в какой части этой сети коллизия возникла. Сеть Ethernet, построенная на повторителях, всегда образует один домен коллизий. Мосты, коммутаторы и маршрутизаторы делят сеть на несколько доменов коллизий.

Стандарты 10 Base-5 и 10 Base-2 разрешают использование в сети не более 4 повторителей и, соответственно, не более 5 сегментов кабеля. При максимальной длине сегмента кабеля 500м максимальная длина сети 10 Base-5 составляет 2500м. Для стандарта 10 Base-2 максимальная длина сети равна $5 \cdot 185 = 925$ м. Только 3 сегмента из 5 могут быть нагруженными, т. е. такими, к которым подключаются конечные станции. Правило применения повторителей в сети Ethernet называется «правило 5-4-3»: 5 сегментов, 4 повторителя, 3 нагруженных сегмента.

В стандарте 10 Base-T определено максимальное число концентраторов между любыми двумя станциями сети - 4. Это правило получило название «правило 4 хабов», оно заменяет «правило 5-4-3» применяемое к сетям 10 Base-5 и 10 Base-2. При создании сети 10 Base-T с большим числом станций концентраторы можно соединять друг с другом иерархическим способом, образуя древовидную структуру.

Соблюдение многочисленных ограничений, установленных для различных стандартов физического уровня сетей Ethernet, гарантирует корректную работу сети. Однако если сеть состоит из сегментов различной физической природы, необходимо выполнение следующих условий:

- количество станций в сети - не более 1024;
- максимальная длина физического сегмента - не более стандартной;
- время двойного оборота сигнала (Path Delay Value, PDV) между двумя самыми удаленными станциями - не более 575 битовых интервала;
- сокращение межкадрового интервала (Path Variability Value, PVV) при прохождении кадров через все повторители - не более 49 битовых интервала.

Для того чтобы произвести расчет PDV, можно воспользоваться справочными данными IEEE по значениям задержек распространения сигналов в повторителях, приемопередатчиках и различных физических средах. Данные для расчета значений PDV приведены в таблице 3.1.

Задержки сигналов, приведенные в таблице, состоят из задержки входного усилителя, блока повторителя и выходного усилителя. Все эти задержки представлены одной величиной названной базой сегмента, измеряемой в битовых интервалах bt. Чтобы не нужно было два раза складывать задержки, вносимые кабелям, даются удвоенные величины задержек для каждого типа кабеля. Используются также такие понятия, как левый сегмент, правый сегмент и промежуточный сегмент. С каждым сегментом связана постоянная задержка, на-

званная базой, которая зависит только от типа сегмента и от положения сегмента на пути сигнала. База правого сегмента, в котором возникает коллизия, намного превышает базу левого и промежуточного сегментов.

Таблица 3.1. Данные для расчета PDV

Тип сегмента	База левого сегмента, bt	База промежуточ-	База правого сегмента, bt	Задержка сред- ды на 1м, bt	Максимальная длина сегмен-
10Base-5	11,8	46,5	169,5	0,0866	500
10Base-2	11,8	46,5	169,5	0,1026	185
10 Base-T	15,3	42,0	165,0	0,113	100
10Base-FB		24		0,1	2000
10Base-FL	12,3	33,5	156,5	0,1	2000

Кроме того, с каждым сегментом связана задержка распространения вдоль кабеля сегмента, которая зависит от длины сегмента и вычисляется путем умножения времени распространения сигнала по одному метру кабеля на длину кабеля в метрах.

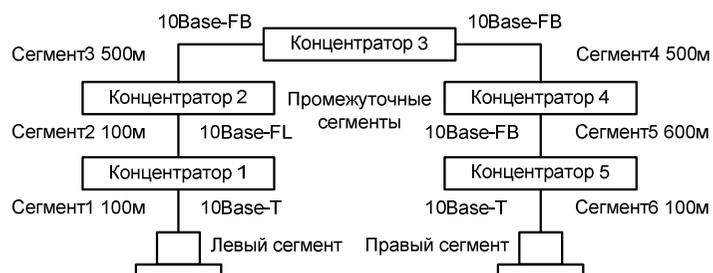


Рис. 15. Пример сети для расчета PDV

Расчет будет состоять из следующих действий:

- для левого сегмента 1; $15,3 + 100 \times 0,113 = 26,6$;
- для промежуточного сегмента 2; $33,5 + 1000 \times 0,1 = 133,5$;
- для промежуточного сегмента 3; $24 + 500 \times 0,1 = 74,0$;
- для промежуточного сегмента 4; $24 + 500 \times 0,1 = 74,0$;
- для промежуточного сегмента 5; $24 + 600 \times 0,1 = 84,0$;
- для правого сегмента 6; $165 + 100 \times 0,113 = 176,3$.

Сумма всех составляющих дает значение равное 568,4, что меньше максимально допустимой величины PDV.

Чтобы признать конфигурацию сети корректной, нужно рассчитать также умень-

шение межкадрового интервала повторителями, т. е. величину PVV.

Для расчета PVV можно воспользоваться значениями максимальных величин уменьшения межкадрового интервала при прохождении повторителей различных физических сред, которые представлены в таблице 3.2.

Сумма величин, определенная для всех сегментов сети и будет расчетным значением PVV, которое не должно превосходить значение 49 битовых интервалов.

Таблица 3.2 Данные для расчета PVV

Тип сегмента	Передающий сегмент, bt	Промежуточный сегмент, bt
10 Base-5, 10 Base-2	16	11
10Base-FB		2
10 Base-FL	10,5	8
10Base-T	10,5	8

Для приведенного выше примера значение PVV можно рассчитать следующим образом:

- для левого сегмента 1; сокращение в 10,5 bt;
- для промежуточного сегмента 2; в 8 bt;
- для промежуточного сегмента 3; в 2 bt;
- для промежуточного сегмента 4; в 2 bt;
- для промежуточного сегмента 5; в 2 bt;
- для правого сегмента 6; в 2 bt.

Сумма этих величин дает значение PVV, равное 24,5, что меньше предельного значения в 49 битовых интервала.

8.3. Глобальные компьютерные сети

8.3.1. Характеристика глобальных компьютерных сетей

Первоначально глобальные сети решали задачу доступа удаленных ЭВМ и терминалов к мощным ЭВМ, которые назывались host-компьютер (часто используют термин сервер). Такие подключения осуществлялись через коммутируемые или некоммутируемые каналы телефонных сетей или через спутниковые выделенные сети передачи данных, например, сети, работающие по протоколу X.25.

Для подключения к таким сетям передачи данных использовались модемы, работающие под управлением специальных телекоммуникационных программ, таких как BITCOM, COMIT, PROCOM, MITEZ и т.д. Эти программы, работая под операционной системой MS-DOS, обеспечивали установление соединения с удаленным компьютером и обмен с ним информацией. С закатом эры MS-DOS их место занимает встроенное в операционные системы

коммуникационное программное обеспечение. Примером могут служить средства Windows95 или удаленный доступ (RAS) в WindowsNT.

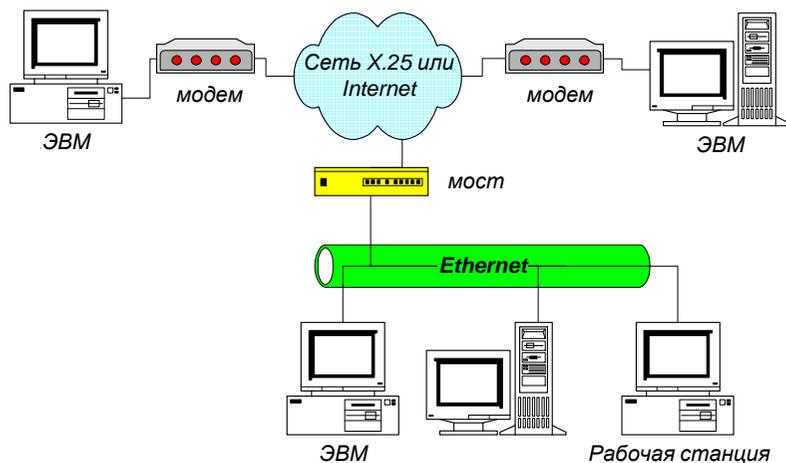


Рис. 1. Принцип объединения компьютеров в глобальных сетях.

В настоящее время все реже используются подключенные к глобальным сетям одиночные компьютеры. Это в основном домашние ПК. В основной массе абонентами компьютерных сетей являются компьютеры, включенные в локальные вычислительные сети (ЛВС), и поэтому часто решается задача организации взаимодействия нескольких удаленных локальных вычислительных сетей. При этом требуется обеспечить удаленному компьютеру связь с любым компьютером удаленной локальной сети, и, наоборот, любому компьютеру ЛВС с удаленным компьютером. Последнее становится весьма актуальным при расширении парка домашних и персональных компьютеров.

В России крупнейшими глобальными сетями считаются Спринт сеть (современное название Global One), сеть Инфотел, сети Роснет и Роспак, работающие по протоколу X.25, а также сети Relcom и Internet, работающие по протоколу TCP/IP.

В качестве сетевого оборудования применяются центры коммутации, которые для сетей X.25 часто исполняются как специализированные устройства фирм-производителей Siemens, Telenet, Alcatel, Ericsson и др., а для сети с TCP/IP используются маршрутизаторы фирм Cisco и Decnis. Структура сетей показана на рисунке 1.

8.2.2. Сеть Internet

Internet является старейшей глобальной сетью. Internet предоставляет различные спосо-

бы взаимодействия удаленных компьютеров и совместного использования распределенных услуг и информационных ресурсов.

Internet работает по протоколу TCP/IP. Основным «продуктом», который вы можете найти в Internet, является информация. Эта информация собрана в файлы, которые хранятся на хост-компьютерах, и она может быть представлена в различных форматах. Формат данных зависит от того, каким сетевым сервисом вы воспользовались, и какие возможности по отображению информации есть на ПК. Любой компьютер, который поддерживает протоколы TCP/IP, может выступать в качестве хост-компьютера.

Ключом к получению информации в Internet являются адреса ресурсов. Вам придется использовать почтовые адреса (mail addresses) при пересылке сообщений по электронной почте своим коллегам и адреса хост-компьютеров (host names) для соединения с ними и для получения файлов с информацией.

Одним из недостатков передачи данных по сети Internet является недостаточная защита информации.

8.2.3. Услуги Internet.

1. Передача файлов по протоколу FTP. Информационный сервис, основанный на передаче файлов с использованием протокола FTP (протокол передачи файлов).
2. Поиск файлов с помощью системы Archie. Archie – первая поисковая система необходима для нахождения нужной информации, разбросанной по Internet.
3. Электронная почта. ЭП – это вид сетевого сервиса. ЭП предусматривает передачу сообщений от одного пользователя, имеющего определенный компьютерный адрес, к другому. Она позволяет быстро связаться друг с другом.
4. Списки рассылки. Список рассылки – это средство, предоставляющее возможность вести дискуссию группе пользователей, имеющих общие интересы.
5. Телеконференции. Телеконференции в Internet предоставляют возможность вести дискуссии (при помощи сообщений) по тысячам размещенных тем.

8.2.4. Возможности сети Internet.

Интернет представляет собой глобальную компьютерную сеть, содержащую гигантский объем информации по любой тематике, доступной на коммерческой основе для всех желающих, и предоставляющую большой спектр информационных услуг. В настоящее время Интернет представляет собой объединение более 40 000 различных локальных сетей, за что она получила название сеть сетей. Каждая локальная сеть называется узлом или сайтом, а юридическое лицо, обеспечивающее работу сайта – провайдером. Сайт состоит из нескольких компьютеров – серверов, каждый из которых предназначен для хранения информации определенного типа и в определенном формате. Каждый сайт и сервер на сайте имеют уникальные имена, по-

средствам которых они идентифицируются в Интернет.

Для подключения в Интернет пользователь должен заключить контракт на обслуживание с одним из провайдеров в его регионе.

8.2.4. Доступ к информационным ресурсам.

Имеется несколько видов информационных ресурсов в Интернет, различающихся характером информации, способом ее организации, методами работы с ней. Каждый вид информации хранится на сервере соответствующего типа, называемых по типу хранимой информации. Для каждой информационной системы существуют свои средства поиска необходимой информации во всей сети Интернет по ключевым словам. В Интернет работают следующие информационные системы:

- World Wide Web (WWW) – Всемирная информационная паутина. Эта система в настоящее время является наиболее популярной и динамично развивающейся. Информация в WWW состоит из страниц (документов). Страницы могут содержать графику, сопровождаться анимацией изображений и звуком, воспроизводимым непосредственно в процессе поступления информации на экран пользователя. Информация в WWW организована в форме гипертекста. Это означает, что в документе существуют специальные элементы – текст или рисунки, называемые гипертекстовыми ссылками (или просто ссылками), щелчок мышью на которых выводит на экран другой документ, на который указывает данная ссылка. При этом новый документ может храниться на совершенно другом сайте, возможно, расположенном в другом конце земного шара.
- Gopher-система. Эта система является предшественником WWW и сейчас утрачивает свое значение, хотя пока и поддерживается в Интернет. Просмотр информации на Gopher-сервере организуется с помощью древовидного меню, аналогичного меню в приложениях Windows или аналогично дереву каталогов (папок) файловой системы. Меню верхнего уровня состоит из перечня крупных тем, например, экономика, культура, медицина и др. Меню следующих уровней детализируют выбранный элемент меню предыдущего уровня. Конечным пунктом движения вниз по дереву (листом дерева) служит документ аналогично тому, как конечным элементом в дереве каталогов является файл.
- FTP (File Transfer Protocol) – система, служащая для пересылки файлов. Работа с системой аналогична работе с системой NC. Файлы становятся доступными для работы (чтение, исполнения) только после копирования на собственный компьютер. Хотя пересылка файлов может быть выполнена с помощью WWW, FTP-системы продолжают оставаться весьма популярными ввиду их быстрой работы и простоты использования.

8.2.5. Адресация и протоколы в Интернет.

Компьютер, подключенный к Интернет, и использующий для связи с другими компью-

терами сети специальный протокол TCP/IP, называется хостом. Для идентификации каждого хоста в сети имеются следующие два способа адресации, всегда действующие совместно.

Первый способ адресации, называемый IP-адресом, аналогичен телефонному номеру. IP-адрес хоста назначается провайдером, состоит из четырех групп десятичных цифр (четыре байта), разделенных точками, заканчивается точкой.

Аналогично телефонам, каждый компьютер в Интернет должен иметь уникальный IP-адрес. Обычно пользователь свой IP-адрес не использует. Неудобство IP-адреса состоит в его безликости, отсутствии смысловой характеристики хоста и потому трудной запоминаемости.

Второй способ идентификации компьютеров называется системой доменных имен, именуемой DNS (Domain Naming System).

DNS-имена назначаются провайдером и, например, имеет вид: **win.smtp.dol.ru**.

Приведенное выше доменное имя состоит из четырех, разделенных точками, простых доменов (или просто доменов). Число простых доменов в полном доменном имени может быть произвольным. Каждый из простых доменов характеризует некоторое множество компьютеров. Домены в имени вложены друг в друга, так что любой домен (кроме последнего) представляет собой подмножество домена, следующего за ним справа. Так, в приведенном примере DNS-имени домены имеют следующий смысл:

ru – домены страны, в данном случае обозначает все домены в России;

dol – домен провайдера, в данном случае обозначает компьютеры, локальной сети российской фирмы Demos;

smtp – домен группы серверов Demos, обслуживающих систему электронной почты;

win – имя конкретного компьютера из группы smtp.

Таким образом, по всей организации и внутренней структуре DNS-система напоминает полный путь к конкретному файлу в дереве каталогов и файлов. Одно из различий состоит в том, что домен более высокого уровня в DNS-имени находится правее. Так же, как и IP-адрес, DNS-имя должно однозначно идентифицировать компьютер в Интернет. Полное доменное имя должно заканчиваться точкой.

8.2.6. Протокол Frame Relay (FR).

Frame Relay – это протокол, который описывает интерфейс доступа к сетям быстрой коммутации пакетов. Он позволяет эффективно передавать крайне неравномерно распределенный по времени трафик и обеспечивает высокие скорости прохождения информации через сеть, малые времена задержек и рациональное использование полосы пропускания.

По сетям FR возможна передача не только собственно данных, но и также оцифрованного голоса.

Согласно семиуровневой модели взаимодействия открытых систем OSI, FR – протокол

второго уровня. Однако он не выполняет некоторых функций, обязательных для протоколов этого уровня, но выполняет функции протоколов сетевого уровня. В то же время FR позволяет устанавливать соединение через сеть, что в соответствии с OSI, относится к функции протоколов третьего уровня.

9. АНТЕННЫ

9.1. Назначение и классификация антенн

9.1.1. Назначение антенны.

Антенна является необходимой частью любой радиотехнической системы. Антенны соединяются с радиопередатчиком через фидеры (линии передачи), а в некоторых случаях — непосредственно. Совокупность антенны и всех звеньев фидера называется *антенно-фидерным устройством* или *антенно-фидерным трактом*, а если они очень сложны по выполняемым функциям и структуре, то их по праву называют *антенно-фидерными системами*.

Передающая часть радиолинии состоит из радиопередатчика, передающего фидера и передающей антенны. Передатчик предназначен для получения достаточно мощного радиосигнала в виде тока высокой (несущей) частоты, модулированного управляющим сигналом. Радиосигнал поступает в фидер и распространяется по нему в форме направляемых электромагнитных волн. Передающая антенна преобразует их в свободные волны, т. е. излучает в свободное пространство.

На приемном конце радиолинии происходит обратный процесс: свободно распространяемые волны попадают в зону действия приемной антенны, и принятый радиосигнал передается по приемному фидеру в форме направляемых радиоволн; на входе приемника эти волны вызывают модулированный ток, который не отличается по частоте и форме от тока на выходе передатчика, приемник усиливает и демодулирует (детектирует) поступивший радиосигнал, в результате выделяется управляющий сигнал; содержащаяся в управляющем сигнале информация воспроизводится оконечным прибором (телефоном, громкоговорителем, электронно-лучевой трубкой и т. п.).

Итак, *передающий антенно-фидерный тракт предназначен для преобразования энергии модулированных токов высокой частоты в энергию свободных электромагнитных волн, а приемный - для обратного преобразования энергии, и при этом важно избежать искажения передаваемой информации.*

9.1.2. Классификация антенн.

Антенны классифицируют по многим признакам и параметрам. Назовем основные из них.

1. *Направленность излучения и приема.* По этому признаку различают: *слабонаправленные антенны*, линейные размеры которых либо много меньше длины волны излучаемых (принимаемых) волн, либо соизмеримы с ней; *умеренно направленные антенны*, размеры которых порядка единиц длин волн; *остронаправленные антенны*, размеры которых порядка десятков длин волн.

2. *Принцип действия и конструктивное выполнение антенны.* Любая антенна содержит вполне определенные конструктивные элементы (проводники, отверстия, щели, замедляющие структуры и др.), которые излучают электромагнитные волны при условии, что эти элементы обтекаются токами проводимости или смещения достаточно высокой частоты. В связи с этим различают:

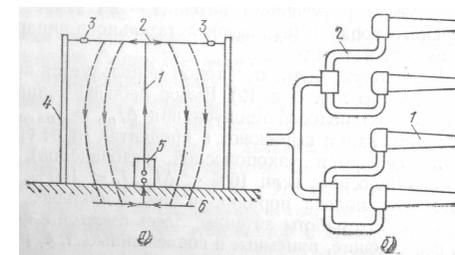


Рис. 1. Антенны: а – линейная проволочная; б – поверхностных волн

Проволочные и штыревые антенны, применяемые в основном на километровых, гектометровых и декаметровых волнах. В линейной проволочной антенне, показанной на рис. 1.1, а, излучающим элементом служит вертикальный провод (7) с присоединенным к нему горизонтальным проводом (2), который через изоляторы (3) подвешивается к мачтам (4). Ток, вызываемый высокочастотным генератором передатчика (5), замыкается по цепи: провода (1, 2), воздушная среда (ток смещения, показанный штриховыми линиями), почва и заземление (6), выполненное в виде системы проводов, зарытых на уровне грунтовых вод. Заземление уменьшает электрическое сопротивление цепи и этим способствует увеличению тока антенны и мощности излучаемых волн.

Щелевые антенны (рис. 1.16) состоят из щелей, прорезанных в металлических экранах или в стенках волноводов, и используются на дециметровых и сантиметровых волнах. Щели излучают при пересечении их токами смещения.

Антенны поверхностных (или вытекающих) волн, где излучение во внешнее пространство происходит в результате замедленного или ускоренного распространения волны по поверхности антенны. Примером может служить диэлектрический стержень (1 на рис. 1.1, б), обтекаемый бегущей волной высокочастотного тока, который поступает от передатчика через линию передачи (2).

Апертурные антенны — это антенны, в которых излучение происходит с большой, по сравнению с квадратом длины волны, площади (апертуры). Это может быть открытое отверстие металлического волновода (рис. 8.1), раскрыв металлических рупоров (рис. 8.5), параболического зеркала (рис. 8.8) или линзы (рис. 8.19). Волноводные и рупорные антенны являются аналогами излучателей акустического типа, а зеркальные и линзовые — оптического.

Многоэлементные антенны — *антенные решетки*. Излучающими элементами здесь служат слабонаправленные антенны. Простейшая антенная решетка из диэлектрических стержней показана на рис. 1.1, б,

3. *Характер сканирования, т. е. перемещения луча в пространстве*. В несканирующих антеннах луч занимает фиксированное положение, а в сканирующих можно управлять его пространственным положением. Сканирующие антенны, как правило, остронаправленные. Они применяются в радиолокации, космической радиосвязи и радиоастрономии. По мере развития этих и других перспективных направлений современной радиотехники области применения сканирующих антенн все более расширяются, схемы и конструкции их непрерывно совершенствуются и становятся все более разнообразными.

Возможно механическое, электромеханическое и электрическое (фазовое и частотное) сканирование.

Механическое сканирование осуществляется вращением всей антенны. *Электромеханическое сканирование* производится перемещением какой-то части антенны, а остальная (более массивная) часть ее остается неподвижной. При *электрическом сканировании* вся антенна неподвижна, а луч перемещается в результате изменения амплитудно-фазового распределения токов или полей возбуждения излучателей антенны.

Фазовое сканирование успешно реализуется при помощи *фазированных антенных решеток (ФАР)* с фазовращателями в качестве элементов управления. *Частотное сканирование* основано на дисперсии волн в системе возбуждения элементов антенны, оно требует применения перестраиваемых по частоте передатчика и приемника и такой фидерной системы, по которой волны передаются с замедлением, резко зависящим от частоты.

К числу последних достижений антенной техники относятся *активные фазированные антенные решетки (АФАР)*. Элементами АФАР являются не одиночные излучатели, а модули, в состав которых кроме излучателя входят генератор, приемник и устройство развязки режимов прием—передача. Все эти устройства связаны не только конструктивно, но и функционально, благодаря чему возможности АФАР значительно шире возможностей пассивных решеток.

Угловая скорость луча при механическом, электромеханическом и электрическом сканировании составляет соответственно доли, десятки — сотни и десятки тысяч радиан в секунду.

4. *Второстепенные признаки классификации антенн:*

— область применения — по этому признаку антенны разделяются на связные, радиовещательные, телевизионные, радиолокационные, радионавигационные, для радиотелескопов и т. д.;

— место размещения антенны — на земле (наземные), на борту морского корабля (судовые), летательного аппарата (самолетные, космические) и т. д.;

— диапазон волн, в котором используется антенна (эти диапазоны перечислены в ч. I, с. 12). Выбор рабочей волны связан, в частности, с требуемой полосой пропускания $\Delta f_{np} = f_{max} - f_{min}$. Обычно полосу пропускания выражают в процентах от ее средней частоты f_0 . Антенна считается узкополосной (резонансной), если $\Delta f_{np} / f_0 < 10\%$; широкополосной, если $10\% < \Delta f_{np} / f_0 < 100\%$, и сверхширокополосной, если полоса порядка октавы и больше;

— режим работы антенны. Здесь имеется в виду разделение антенн на передающие, приемные и совмещенные, т. е. работающие попеременно: то как передающие, то как приемные. Это возможно потому, что для антенн справедлив принцип взаимности (при передаче они трансформируют направляемые электромагнитные волны в свободные, а при приеме — наоборот, причем параметры и характеристики антенны в обоих режимах одинаковы). В соответствии с этим принципом достаточно изучать только передающие антенны, а при необходимости отмечать особенности приемных антенн.

Можно и на других примерах показать, что перечисленные признаки - второстепенные. Скажем, проволочные линейные антенны используются преимущественно на низких радиочастотах, но они применяются и на более высоких частотах, вплоть до сантиметровых волн. Или, к самолетным антеннам относятся и линейные, и проволочные, и щелевые, и антенны поверхностных волн, и апертурные, и антенные решетки. В дальнейшем будем придерживаться в основном наиболее характерного признака классификации антенн — принципа действия и геометрии излучающей структуры.

9.2. *Первичные параметры и характеристики передающих антенн*

Характеристики и параметры антенн можно делить на первичные и вторичные. Зависимость тех и других от частоты определяет частотный диапазон антенны. Рассмотрим все эти вопросы, опираясь на приведенные в ч. I, гл. 13 сведения о теории электромагнитного поля, диполе Герца и проволочном симметричном вибраторе.

Начнем с первичных параметров антенны. Из них, заметим, сопротивления излучения и потерь и выраженный через них КПД применяются только при исследовании слабонаправленных антенн: проволочных, штыревых и щелевых.

9.2.1. Сопротивление излучения.

Сопротивление излучения по смыслу связано с мощностью излучения антенны P_{Σ} , т. е. со средним значением потока электромагнитной энергии, проходящего в единицу времени сквозь сферу, окружающую антенну. Так как мощность излучения активная (она не возвращается в передающую антенну и передатчик), то ее можно выразить через активное сопротивление, называемое *сопротивлением излучения*,

$$R_{\Sigma} = P_{\Sigma} / I^2 \quad (1.)$$

где I — действующее значение тока в антенне.

9.2.2. Поле излучения диполя Герца.

Применим понятие «сопротивление излучения» к диполю Герца. Напомним, что диполь Герца имеет длину l , значительно меньшую длины волны λ , и поэтому по всей длине обтекается током с одинаковой амплитудой I_m . Пространство, окружающее диполь, мысленно делится на три зоны: ближнюю радиусом $r \ll \lambda$, промежуточную и дальнюю радиусом $r > \lambda$ (рис. 1.2). Электромагнитное поле ближней зоны имеет преимущественно реактивный характер, т. е. в нем поля E и H сдвинуты по фазе на 90° . В дальней зоне поля E и H синфазные, и они совместно свободно распространяются от диполя. При этом вектор E находится в меридиональной, т. е. в проходящей через ось диполя, плоскости, а вектор H — в экваториальной, т. е. в перпендикулярной оси диполя, плоскости. Построенные по уже известным векторам E и H векторы Пойнтинга Π указывают, что *диполь Герца излучает бегущую сферическую волну, которую вдали от диполя можно рассматривать как плоскую*.

Математический анализ приводит к следующим формулам амплитуд напряженностей магнитного и электрического полей диполя Герца:

$$H_m = \frac{I_m l}{2r\lambda} \sin \theta, \quad (2)$$

$$E_m = Z_0 H_m = 60\pi \frac{I_m l}{r\lambda} \sin \theta, \quad (3)$$

где $Z_0 = 120\pi = 377$ Ом — волновое (характеристическое) сопротивление свободного пространства; θ — зенитный угол, образованный осью вертикального вибратора и радиусом r сферы, проходящим через точку наблюдения.

Рассмотрим свойства диполя Герца, вытекающие из формул (2) и (3).

1. Амплитуды напряженностей полей вибратора прямо пропорциональны амплитуде тока в нем; это объясняется тем, что ток вибратора определяет подводимую к нему мощность, а следовательно, и мощность излучения.

2. Напряженности полей прямо пропорциональны отношению l/λ , что соответствует известному положению об усилении излучения электромагнитных волн с увеличением отно-

сительной длины излучателя l/λ (это справедливо в известных пределах).

3. Напряженности полей обратно пропорциональны расстоянию r от вибратора; такая закономерность характерна, как известно, для сферической волны.

4. Амплитуды напряженностей полей изменяются по закону $\sin \theta$ и, следовательно, возрастают от нуля до максимума с приближением направления излучения к экваториальной плоскости вибратора. В связи с этим густота силовых линий и длина векторов E , H , Π (рис. 1.2) минимальны при $\theta = 90^\circ, 270^\circ$ и спадают до нуля к оси вибратора.

9.2.3. Входное сопротивление антенны

Входное сопротивление — это сопротивление нагрузки для генератора, питающего антенну. Генератор расходует мощность на создание электромагнитного поля излучения (P_{Σ}) и компенсацию потерь в антенне (P_u). Кроме того, когда нет резонанса в антенне, в ее ближней зоне энергия электрического поля превышает энергию магнитного или наоборот, и избыточное поле соответствует реактивной мощности генератора P_x .

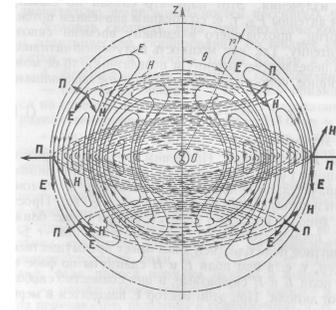


Рис. 2. Поля излучения диполя Герца

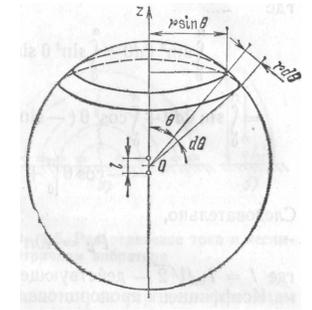


Рис. 3. К расчету сопротивления излучения диполя Герца

Значит, в общем случае входное сопротивление антенны комплексное $Z_{вхА}$. Оно определяется как отношение напряжения \dot{U}_A на зажимах антенны к протекающему через них току I_A или как частное от деления на I_A^2 мощности \dot{P}_A , подводимой к антенне. Так как эта мощность складывается из активных мощностей излучения P_{Σ} и потерь P_u и мощности реактивных полей антенны P_x [ч. I, § 91], то

$$\dot{Z}_{вхА} = P_A / I_A^2 = P_{\Sigma} / I_A^2 + P_u / I_A^2 + jP_x / I_A^2 = R_{\Sigma 0} + R_{u0} + jX_{вхА} = R_{вхА} + jX_{вхА},$$

где $R_{вхА} = R_{\Sigma 0} + R_{u0}$ — активная составляющая входного сопротивления, равная сумме сопротивлений излучения и потерь, отнесенных к входным клеммам антенны; $X_{вхА}$ — реак-

тивная составляющая входного сопротивления, соответствующая мощности реактивных полей вокруг антенны P_x . При резонансе $P_x = 0$, а следовательно, $X_{вх А} = P_x / I_A^2 = 0$ и входное сопротивление антенны активное: $Z_{вхА} = R_{вхА} + jX_{вхА} = R_{вхА}$.

9.2.4. Характеристика направленности антенны.

Точечный изотропный излучатель является абсолютно ненаправленным. Но это абстракция: любая реальная антенна хотя бы немного направленная.

Характеристикой направленности антенны называется зависимость комплексной амплитуды напряженности поля излучения E_m от угловых координат точки наблюдения φ, θ при условии измерения этого поля на одинаковом расстоянии r от антенны. Комплексная амплитуда \dot{E}_m имеет модуль E_m и аргумент ψ ; кроме того, волна характеризуется поляризацией. Соответственно характеристики направленности разделяются на амплитудные, фазовые и поляризационные. Чаще всего используются амплитудные (в таких случаях слово амплитудный опускается).

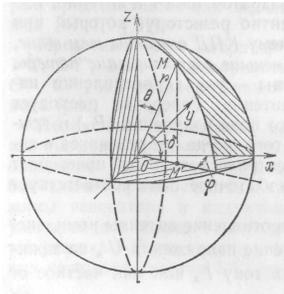


Рис. 6. Сферические координаты точки наблюдения

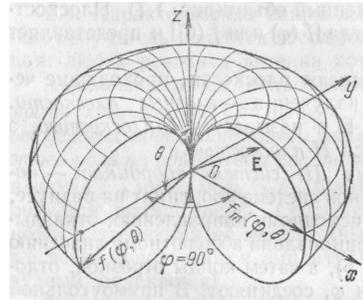


Рис 7. Пространственная (объемная) ДН диполя Герца

Функция, описывающая характеристику направленности, называется *функцией направленности*, а ее графическое изображение - *диаграммой направленности (ДН)*.

Исследуя направленные свойства антенны, полагают, что антенна находится в центре O сферы, а точка наблюдения M расположена на поверхности этой сферы (рис. 6). Сферические координаты точек наблюдения: дальность (радиус сферы) r , азимут φ , зенитный угол θ или угол места $\delta = 90^\circ - \theta$.

9.2.5. Диаграмма направленности (ДН) антенны

Кроме деления на амплитудные, фазовые и поляризационные, ДН классифицируют по

следующим признакам.

1. *По полю и мощности.* ДН полю выражает зависимость амплитудного E_m или действующего E значения напряженности поля от направления (углов φ, θ), а ДН по мощности — от плотности потока мощности Π . Так как плотность потока мощности пропорциональна квадрату напряженности поля ($\Pi = E^2 / 120\pi = 120\pi H^2$), то если ДН по полю выражается функцией $f(\varphi, \theta)$, то ДН по мощности функцией $f^2(\varphi, \theta)$.

Согласно (1.3) амплитуда напряженности электрического поля диполя Герца в направлении зенитного угла θ

$$E_m = (60\pi I_m l \sin \theta) / r\lambda = E_{mm} \sin \theta, \quad (.8)$$

где $E_{mm} = 60\pi I_m l / r\lambda$ — максимальная амплитуда поля, соответствующая $\theta = 90, 270^\circ$. Из (8) находим функции направленности диполя Герца по полю и мощности

$$f(\theta) = \sin \theta, \quad (.9)$$

$$f^2(\varphi, \theta) = \sin^2 \theta \quad (.10)$$

2. *Пространственные (объемные) и плоскостные ДН.* Пространственная ДН — функция двух углов φ, θ и представляет собой поверхность $f(\varphi, \theta)$, которая охватывает определенный объем (рис. 1.7). Плоскостная ДН является функцией одного угла [$f(\varphi)$ или $f(\theta)$] и представляет собой плоскую фигуру (рис. 1.8).

Обычно для построения ДН выбирают плоскости, проходящие через направления максимального излучения — *главные плоскости*. Одна из них совпадает с вектором E и называется *E-плоскостью*, а другая — с вектором H и называется *H-плоскостью*.

3. *По выбранной для построения ДН системе координат — полярной или прямоугольной.* В полярной системе координат на радиусе, расположенном под углом θ или φ к исходному направлению, откладывают отрезок, длина которого пропорциональна абсолютному значению $f(\theta)$ (рис. 1.8, а) или $f(\varphi)$ (рис. 1.8, б), а затем концы отрезков, отложенных для различных углов θ или φ , соединяют. В прямоугольной системе координат угол θ или φ откладывают по оси абсцисс, а $f(\theta)$ (рис. 1.8, в) или $f(\varphi)$ (рис. 1.8, д) — по оси ординат и полученные точки также соединяют.

Функции направленности диполя Герца $f(\varphi, \theta) = \sin 90^\circ = 1$ соответствует пространственная ДН в виде тора — тела вращения окружности диаметром $f_m(\varphi, \theta) = \sin 90^\circ = 1$ вокруг оси диполя (рис. 1.7). Сечения тора главными плоскостями - амплитудные ДН по полю, причем ДН в E-плоскости $f(\theta) = \sin(\theta)$ имеет в полярных координатах форму восьмерки (рис. 8, а), а в прямоугольных — полусинусоид положительного знака (рис. 8, в); ДН в φ-плоскости имеет форму окружности в полярных (рис. 1.8, б) и прямой, параллельной оси абсцисс, в прямо-

угольных координатах (рис. 8, д). Отсутствие направленности в Н-плоскости ($f(\varphi) = \text{const}$) согласуется с тем, что в выражение для функции направленности угол φ не входит.

Фазовая ДН (рис. 8, з) отражает скачкообразное изменение фазы поля Ψ на 180° : когда зенитный угол проходит значения $\theta = 180, 360^\circ$, функция $f(\theta) = \sin \theta$ изменяет знак.

ДН по мощности $f^2(\theta) = \sin^2 \theta$, показанная штриховой линией на рис. 1.8, а, в, совпадает с ДН по полю $|f(\theta)| = |\sin \theta|$ в точках, где $\theta = 0, 90, 180, 270, 360^\circ$, а при остальных значениях θ функция $\sin^2 \theta < |\sin \theta|$.

Из расположения векторов Е, Н и П на рис.8, а, б видно, что Е-плоскость азимутальная, а Н-плоскость — экваториальная. Это характерно для антенн с осевой симметрией. Плоскостные ДН как более простые применяются чаще, чем пространственные.

4. По выбранным масштабам отсчета ДН, ДН многих антенн, особенно остронаправленных, имеют многолепестковую структуру (рис. 9): кроме главного лепестка (луча) (1) они содержат боковые (2, 3) и задний (4) лепестки. В таких случаях предпочтительнее изображение ДН в прямоугольных координатах (рис. 9, б), на котором можно выбрать любой масштаб по осям и этим представить слабые боковые и задний лепестки более отчетливо, чем на полярной ДН (рис. 9, а).

Различают ненормированные и нормированные ДН. Нормирование означает, что величины E_m и P в направлении φ, θ отнесены к их максимальным значениям E_{mm} и P . Функции нормированных ДН по полю и мощности обозначаются через

$$F(\varphi, \theta) = f(\varphi, \theta) / f_m(\varphi, \theta) = E_m / E_{mm}, \quad (11)$$

$$F^2(\varphi, \theta) = f^2(\varphi, \theta) / f_m^2(\varphi, \theta) = P / P_m, \quad (12)$$

Нормированные ДН удобно представлять в логарифмическом масштабе, при котором функции направленности выражаются в децибелах [ч. I, п. 18.3]. Децибел — это десятая доля бела, который определяется десятичным логарифмом отношения мощностей, т. е.

$$N[\text{дБ}] = 10 \lg F^2(\varphi, \theta) = 20 \lg F(\varphi, \theta). \quad (13)$$

По данной формуле составлена табл. 1.1, иллюстрирующая пересчет в децибелы N [дБ] нормированной функции направленности по мощности $F^2(\varphi, \theta)$ и по полю $F(\varphi, \theta)$.

Например, по ДН, показанной на рис. 9, уровень первого бокового лепестка (2) ниже уровня луча (1) на 6 дБ, т. е. меньше в два раза по напряженности поля и в четыре раза по мощности.

Заметим еще, что если функция $f(\varphi, \theta)$ имеет максимум $f_m(\varphi, \theta) = 1$, то ненормированная и нормированная ДН совпадают: $f(\varphi, \theta) = F(\varphi, \theta)$, как, например, для диполя Герца (рис. 8).

9.2.6. Предельно пропускаемая мощность

Если мощность P_A , подводимая к антенне, больше допустимой $P_{\text{доп}}$, то напряженность электрического поля в фидере или вблизи антенны превышает критическое значение. Газовый разряд, происходящий при таком перенапряжении, связан с непрерывной затратой энергии на ионизацию и рекомбинацию (восстановление) молекул воздуха. Это способствует пробое диэлектриков и проводов антенны и уменьшает ее КПД.

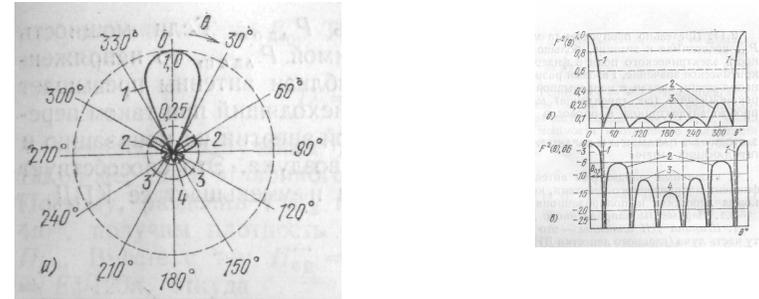


Рис. 9. Нормированные ДН по мощности в линейном (а, б) и логарифмическом (в) масштабах

10. ОПТИЧЕСКАЯ ГОЛОГРАФИЯ

10.1. Физические принципы голографии

Основы голографии были заложены в 1948 г. английским физиком Д.Габором. Желая усовершенствовать электронный микроскоп, Д. Габор предложил регистрировать информацию не только об амплитудах, но и о фазах электронных волн путем наложения на одну из них когерентной волны. Однако отсутствие мощных источников когерентного света не позволило ему получить качественные голографические изображения. В 1971 г. за открытие метода голографии профессору Д. Габору была присуждена Нобелевская премия по физике. Второе рождение голография пережила в 1962 и 1963 гг., когда американские физики Э. Лейт и Ю. Упатниек применили в качестве источника света лазер и разработали схему с наклонным опорным пучком. Советский ученый Ю. Н. Денисюк впервые предложил и осуществил запись голограмм в трехмерной среде.

Голография представляет собой способ записи и восстановления волнового поля, основанный на регистрации интерференционной картины (голограммы), которая образована волной (рис. 1, а), отраженной предметом, освещаемым источником света (объектная волна), и когерентной с ней волной, идущей непосредственно от источника света (опорная волна). Голограмма, освещенная опорной волной, создает такое же амплитудно-фазовое пространственное распределение волнового поля, как и при записи объектной волны (рис. 1.1, б).

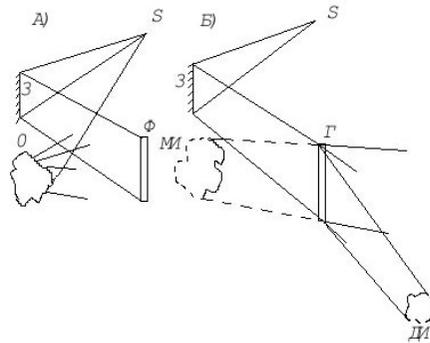


Рис. 1. – Схемы получения голограммы(А) и восстановления волнового фронта(Б) S - источник света; O - объект, φ - фотопластинка; Г - голограмма; З — зеркало; МИ - мнимое изображение; ДИ - действительное изображение

Голографический метод применим ко всем волнам: электронным, рентгеновским, световым, акустическим и сейсмическим микроволнам, при условии, что они когерентны для создания требуемых интерференционных картин. Этот метод наиболее пригоден в оптическом диапазоне электромагнитного спектра. После создания лазеров оптическая голография стала быстро развиваться.

Рассмотрение основ голографии начнем с волновых явлений интерференции и дифракции.

10.1.1.Интерференция.

В голографии обычно имеют дело с интерференцией объектной и опорной волн. Простейший тип синусоидальных волн приведен на рис. 1.2.

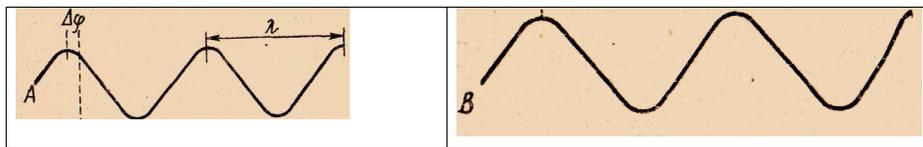


Рис.1.2 – Синусоидальные волны, сдвинутые по фазе.

Расстояние между двумя гребнями волны называют *длиной волны* λ. Для световых волн λ характеризует один цвет. Волны A и B имеют одинаковые длины и амплитуды. Волна B опережает волну A по фазе на Δφ. Если волна B распространяется относительно волны A, то разность фаз этих волн изменяется во времени.

Для голографии интересны два случая взаимодействия волн: 1) волны находятся в фазе; 2) волны находятся в противофазе. При взаимодействии волн A и B получают новую волну, амплитуда которой является результатом сложения амплитуды волны A с амплитудой волны B

в каждой точке. Если волны находятся в фазе, то при их наложении возникает усиливающая интерференция и результирующая волна имеет большую интенсивность, чем каждая из волн в отдельности (рис. 1.3, а). В том случае, когда волны находятся в противофазе, наблюдается ослабляющая интерференция; волны полностью гасят друг друга, амплитуда результирующей волны равна нулю (рис. 1.3, б). Процесс сложения нескольких волн, в результате которого образуется новая волна, называется *интерференцией*.

Рассмотрим интерференцию двух плоских волн, спроецированных на экран (рис. 1.3).

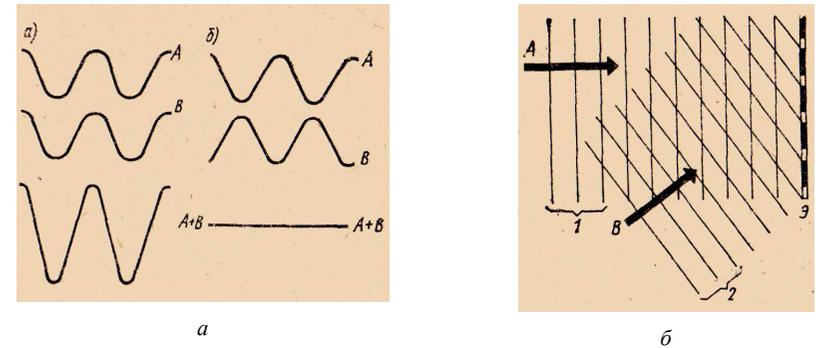


Рис. 3. Интерференция двух плоских волн на экране:
а - сложение двух синусоидальных волн;
б - 1 - гребни волны А; 2 - гребни волны В; Э - белый экран

Интерференция возникает в различных точках вдоль экрана и зависит от разности фаз волн, приходящих в данную точку. В одних точках гребни волны B совпадают с гребнями волны A, в других - гребень волны A встречается с впадиной волны B. Соответственно на экране наблюдаются области большой и малой интенсивности. Пространственное распределение интенсивности излучения, получаемое при этом, называют *интерференционной картиной*. При интерференции двух волн $U_r = U_r \cdot \exp[i\varphi_r(x, y)]$; $U_0 = U_0 \cdot \exp[i\varphi_0(x, y)]$ результирующая интенсивность излучения

$$I = |U_r + U_0|^2 = |U_r|^2 + |U_0|^2 + U_r \cdot U_0 + U_r \cdot U_0^*$$

или

$$I = |U_r|^2 + |U_0|^2 + 2U_r U_0 \cos(\varphi_2 - \varphi_1) \tag{1}$$

Таким образом, интенсивность излучения в любой точке интерференционной картины, образованной двумя волнами, является суммой интенсивностей отдельных волн и интерференционной составляющей, содержащей информацию о разности фаз.

10.1.2. Дифракция.

Дифракция — явление, заключающееся в искривлении световых лучей при встрече с препятствием, оптическое пропускание или отражение которого значительно изменяется на расстояниях, близких к длине волны света. На маленьких отверстиях в экранах дифрагируют все световые лучи, прошедшие через них, на больших — лучи, прошедшие вблизи краев отверстия. Световые лучи, проходящие через центральную часть большого отверстия, не испытывают дифракции.

Если на пути волнового фронта находится несколько небольших препятствий, он в результате дифракции изменяется таким образом, что световые лучи, распространяющиеся за препятствием, имеют качественно новый волновой фронт. Дифракцию можно рассматривать как механизм, посредством которого создают новый волновой фронт света.

Простейшее устройство, которое таким путем формирует новый волновой фронт, называют *дифракционной решеткой*. При падении лазерного пучка на дифракционную решетку ДР (рис. 1.5) часть его проходит через решетку прямо, а часть загибается; в результате формируются два новых пучка, выходящих из решетки под некоторым углом к исходному.

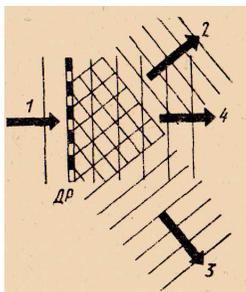


Рис. 5. Получение плоских волн методом дифракции:

1 — опорная волна; 2 — восстановленная плоская волна; 3 — сопряженная восстановленная плоская волна; 4 — часть плоской волны, прошедшая через решетку.

Если исходный лазерный пучок имеет, например, плоский волновой фронт, то и два образованных пучка также обладают плоскими волновыми фронтами. Так как ДР на рис. 5 аналогична картине на рис. 4, то ее можно рассматривать как простейший пример голограммы.

Голограмма сама по себе есть объект, вызывающий дифракцию. На рис. 1.6 представлена в разрезе плоская дифракционная решетка.

Для плоской волны, падающей на решетку, условие синфазности дифрагированных пучков, ведущее к их взаимному усилению, является уравнением решетки:

$$d(\sin i + \sin \delta) = \lambda, \quad (1.2)$$

где d — постоянная решетки; i — угол падения; δ — угол дифракции. Как видно из рис. 1.6, при выполнении условия (1.2) под углом δ образуется главный максимум дифрагированной плоской волны. Кроме того, возможны отрицательные и более высокие порядки дифракции.

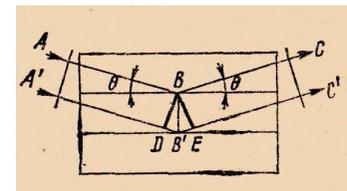


Рис. 7. К пояснению дифракции света в объемной дифракционной решетке

На рис. 7 показана (в разрезе) объемная дифракционная решетка, содержащая периодически расположенные рассеивающие плоскости, освещенные плоской волной. Здесь справедлив тот же самый принцип: интенсивность излучения максимальна в направлении, в котором происходит синфазное сложение световых волн, рассеянных последовательными плоскостями. Условие образования главного максимума дифрагированной плоской волны

$$2d \sin \Theta = \lambda \quad (3)$$

называют *законом Брэгга*. Английский проф. У. Брэгг предположил, что дифракция рентгеновских лучей в кристалле обусловлена отражением падающей волны от кристаллических плоскостей. Максимум дифракции возникает, когда углы, образованные падающим и отраженным лучами с кристаллической плоскостью, равны (рис. 2.7), причем угол θ удовлетворяет условию (3).

Сравнение (2) и (3) обнаруживает, что последнее выражение накладывает более жесткие условия на наблюдение максимума дифракции. Для объемной решетки выбор угла падения определяет длину волны и угол дифракции. Для плоских решеток допускается произвольный выбор этих величин.

10.2. Формирование голограмм.

Для того чтобы получить голограмму, когерентный свет лазера необходимо разделить на две волны, одна из которых освещает объект, а другая является опорной (см. рис. 1, а). Опорная волна направляется таким образом, чтобы она пересекалась со световой волной, прошедшей через объект или отраженной от него. Если обе волны абсолютно когерентны, то интерференционная картина образуется во всем объеме, в котором они перекрываются. Фотопластинка, помещенная в область перекрытия, регистрирует интерференционную картину. После соответствующей химической обработки фотопластинки получают голограмму.

10.3. Восстановление волнового фронта.

Если голограмма освещается опорной волной, то часть дифрагировавшего на ней света вновь воссоздает волновой фронт, который при регистрации шел от объекта. Восстановленная волна исходит из голограммы точно так же, как первоначальная объектная волна. Наблюдатель, видящий волну, идентичную объектной волне, воспринимает ее как бы исходящей от мнимого изображения предмета, расположенного там, где ранее находился предмет (см. рис. 1.1, б). В двумерной голограмме одновременно восстанавливается сопряженная волна, образующая искаженное действительное изображение предмета.

Таким образом, голографический процесс включает в себя процесс получения голограммы и ее восстановления.

Что же дает этот трудоемкий процесс, кроме копии волнового фронта, которой располагали с самого начала? Ответы на этот вопрос будут получены при рассмотрении применения голографии в различных областях науки и производства. В большинстве случаев интерес представляет не восстановление волнового фронта, а некоторый объект (возможно трехмерный), сквозь который волна проходит или от которого она отражается, распространяясь по направлению к плоскости записи. Возможность записи информации об объекте в виде голограммы, а не прямого изображения дает множество преимуществ, особенно в смысле характера и качества информации, которая может быть восстановлена.

10.4. Основное уравнение голографии

Идея голографической записи состоит в том, чтобы кроме волны, которая нас интересует (объектная волна), регистрировать дополнительно опорную волну, когерентную с объектной волной, таким образом, чтобы в результате их интерференции осуществлялось кодирование фазы ср. Введем следующие обозначения: $U_o = U_o \cdot \exp[i\varphi_1 \cdot (x, y)]$ — объектная волна, $U_0 = U_0 \cdot \exp[i\varphi_2 \cdot (x, y)]$ — опорная волна. Возьмем для записи пластинку с линейным откликом на интенсивность падающей волны. Она регистрирует распределение интенсивности, описываемое выражением $I = |U_r + U_0|^2 = |U_r|^2 + |U_0|^2 + U_r * U_0 + U_r^* U_0$. Ее пропускание после обычной фотохимической обработки будет пропорционально интенсивности I .

Осветим эту пластинку опорной волной. Волна, прошедшая сквозь пластинку,

$$U_r I = (|U_r|^2 + |U_0|^2)U_r + |U_r|^2 U_0 + U_r^* U_0^* \quad (4)$$

Это выражение основное для голографии. В нем три слагаемых:

1) $(|U_r|^2 + |U_0|^2)U_r$ — опорная волна U_r , амплитуда которой модулирована коэффициентом $(|U_r|^2 + |U_0|^2)$;

2) $|U_r|^2 U_0 = U_r^2 U_0 \cdot \exp(i\varphi_0)$ — объектная волна U_0 или модулированная коэффициентом U_r^2 ;

3) $U_0 * U_r^2 = U_r^2 U_0 \cdot \exp(i(2\varphi_2 - \varphi_0))$ — волна, комплексно-сопряженная с объектной. Это слагаемое появляется из-за квадратичного детектирования. Оно несет информацию, близкую к информации об объекте, но отличается от объектной волны обратной фазой.

Таким образом, получение фазы и амплитуды объектной волны на этапе восстановления голограммы сопровождается восстановлением двух паразитных волн: опорной, которую ослабили, но не подавили совсем, и сопряженной.

10.5. Основные типы голограмм

10.5.1. Двумерная и объемная голограммы.

Важность третьего измерения, т. е. глубины голографической записи, установил еще на раннем этапе развития голографии Ю. Н. Денисюк. За большой вклад в становление и развитие голографии Ю. Н. Денисюк был удостоен Ленинской (1970) и Государственной (1982) премий СССР, а также премии им. Д. Габора (1983).

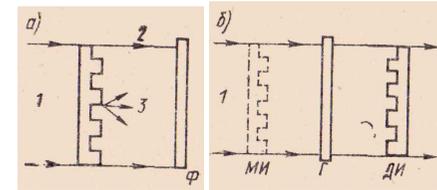


Рис. 8. Схемы получения (а) и восстановления (б) голограммы Габора: / - освещающий пучок; 2 - опорная волна; Я - объектная волна; Ф - фотопластинка; МИ - мнимое восстановленное изображение объекта; Г - голограмма; ДИ - действительное комплексно-сопряженное изображение объекта

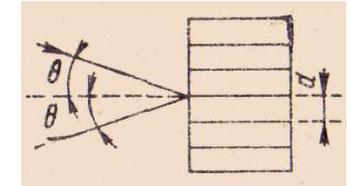


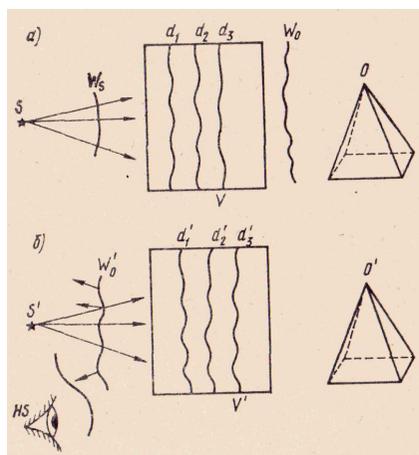
Рис. 9. Интерференционная картина, зарегистрированная в объеме светочувствительной эмульсии.

Голограмму можно рассматривать как тонкую дифракционную структуру лишь в том случае, когда ее оптическая толщина меньше длины волны. На практике, однако, на характеристики голограммы влияет не отношение ее толщины к длине волны, а соотношение между толщиной и периодом самых тонких полос, записанных на голограмме. Если период тончайших полос больше толщины, то голограмма является двумерной, если же период полос меньше толщины, то - трехмерной.

К двумерным относятся голограммы Габора, Лейта и Упатниекса, имеющие период полос d дифракционной структуры больше ее толщины h ($d > h$).

Д. Габор, в распоряжении которого были только источники света с очень малой длиной когерентности, смешал объектную и опорную волны (рис. 8, а). Голограммы Габора имеют существенный недостаток: паразитные волны (опорная и сопряженная) восстанавливаются в том же направлении, что и объектная. В результате очень яркий однородный фон и наличие сопряженного расфокусированного изображения объекта мешают наблюдателю видеть объект (рис. 1.8, б). Устранить этот недостаток позволила схема записи голограмм Лейта, использующая наклонную опорную волну (см. рис. 1.1). При восстановлении волнового фронта наклонной опорной волной наблюдатель может видеть мнимое изображение объекта.

К трехмерным относятся голограммы Денисюка, имеющие период полос дифракционной структуры d меньше ее толщины h ($d < h$).



10. Схемы получения (а) и восстановления (б) трехмерной голограммы: S - источник излучения со сплошным спектром; V' - трехмерная голограмма; W₀' - волна излучения, восстановленного трехмерной голограммой; O' — восстановленное голограммой объемное изображение объекта, с помощью которого была получена голограмма во время записи.

Рис. 9 поясняет образование полос в объеме светочувствительной эмульсии дифракционной решетки для плоских объектной и опорной волн. Если такую решетку осветить исходной опорной волной, то каждый луч до выхода из эмульсии последовательно рассеивается от большого числа периодически расположенных поверхностей максимальной плотности. Чтобы амплитуда результирующей дифрагированной волны была максимальной, волны, рассеянные последовательными слоями, должны быть синфазны. Для этого необходимо, чтобы выполнялось определенное соотношение между длиной волны λ , углом Θ , который составляет освещающая голограмму волна с рассеивающими поверхностями, и расстоянием между этими поверхностями. Эту взаимосвязь устанавливает закон Брэгга (1.3), который можно записать в виде:

$$2d \cdot \sin \Theta = \lambda_0 / n, \quad (5)$$

Где λ_0 - длина волны в воздухе; n - средний показатель преломления светочувствительной среды; Θ - угол, который освещающая и дифрагированная волны составляют с рассеивающими слоями. Закон Брэгга определяет угол падения, если длина волны и расстояние между слоями заданы. Если же угол падения и постоянная решетки выбираются независимо, то закон Брэгга определяет длину волны. Таким образом, трехмерные голограммы, свойства которых описываются законом Брэгга, являются селективными относительно освещающего их излучения.

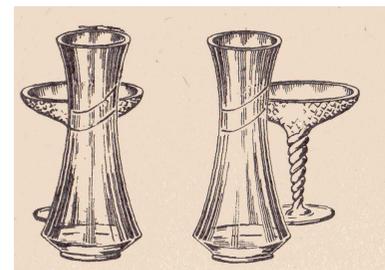


Рис. 11. Восстановленное изображение объемного трехмерного объекта под разными углами зрения

На рис. 10 а приведена схема записи трехмерных голограмм Денисюка. При записи когерентное излучение W_s источника S, прошедшее через объем V, заполненный прозрачным светочувствительным материалом, падает на объект O. В результате интерференции отраженной от объекта волны W_0 и опорной волны W_s в пространстве перед объектом возникает стоячая волна, которую в некотором приближении можно представить в виде системы поверхностей пучностей d_1, d_2, d_3 и т. д., которым соответствует максимальная интенсивность поля. Воздействуя на светочувствительный материал, поле стоячей волны по-разному засвечивает различные части заполненного им объема. В этом объеме после экспозиции и соответствующего проявления образуется трехмерная голограмма, подобная трехмерной дифракционной решетке. Зафиксированные голограммой поверхности пучностей стоячей волны d_1', d_2', d_3' и т. д. (рис. 1.10,б) можно рассматривать как систему зеркал, поверхности которых имеют сложную форму. Трехмерную голограмму можно восстанавливать источником света, имеющим сплошной спектр излучения, так как голограмма является селективной по отношению к освещающему ее излучению (см. (1.5)). При восстановлении точно воспроизводятся практически все параметры зарегистрированного на ней волнового поля объекта - амплитуда, фаза и спектральный состав. В частности, из сплошного спектра источника S' выбирается и отражается излучение той длины волны, которая совпадает с длиной волны излучения, с помощью которого была получена голограмма во время записи.

Наблюдатель HS , воспринимающий восстановленную волну, не может отличить ее от первоначальной объектной и, следовательно, видит объемное изображение O' объекта в цвете, соответствующем длине волны излучения, освещавшего объект при записи (рис. 1.10, б).

Голографическое изображение обладает эффектами параллакса, благодаря чему удается увидеть объект в различных положениях (рис. 1.11).

11. СПИСОК ЛИТЕРАТУРЫ.

1. Гоноровский И.С. Радиотехнические цепи и сигналы. М.: Радио и связь, 1986, -511с.
2. Баскаков СИ. Радиотехнические цепи и сигналы. М.: Высшая школа, 2003, -462 с.
3. Баскаков СИ. Радиотехнические цепи и сигналы. Руководство к решению задач: Учебное пособие. М.: Высшая школа, 2002, - 214 с.
4. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. - СПб.: Питер, 1999. - 672 с.
5. Новиков Ю.В., Кондратенко СВ. - Локальные сети: архитектура, алгоритмы, проектирование. М.: Издательство ЭКОМ, 2001. - 428 с.
6. Keiser G. Optical Fiber Communications. New York: McGraw-Hill Inc., 1991. -461 p.
7. G.P. Agrawal, *Fiber Optic Communication Systems*, 3rd Ed., NY: Wiley, 2002
8. E.Voges, K.Petermann. Optische Kommunikationstechnik. Handbuch für Wissenschaft und Industrie. Springer-Verlag:Berlin, 2002, 1110 s.
9. Крук Б.И., Попантонопуло В.Н., Шувалов В.П.
- 10.
- 11.
- 12.
- 13.
- 14.
- 15.
- 16.
- 17.
- 18.
- 19.
- 20.
- 21.
- 22.
- 23.

- 24.
- 25.
- 26.
- 27.
- 28.
- 29.
- 30.
- 31.
- 32.
33. Компьютерные системы и сети: Учеб.пособие/ В.П.Косарев и др./Под ред. В.П.Косарева и Л.В.Еремина-М.:Финансы и статистика,1999.
34. М.Пайк. Internet в подлиннике :Пер.с англ.-СПб.:ВНУ-Санкт-Петербург,1996.
35. Шварц М.Сети связи:протоколы, моделирование и анализ:в 2-х ч., ч.II: Пер.с англ.- М.:Наука-Гл.ред.физ.-мат.лит.,1992.
36. Гребенников О. Ф., Тихомирова Г.В. Основы записи и воспроизведения информации (в аудиовизуальной технике): Учебное пособие. СПб.: Изд. СПбГУКиТ,2002.-712с.
37. Спортак Марк, Паппас Френк и др. - Компьютерные сети и сетевые технологии. К.: ООО "ТИД "ДС", 2002. - 479 с.
38. Оокоси Т. Оптоэлектроника и оптическая связь. М.: Мир, 1978.
39. Гауэр Дж. Оптические системы связи. М.: Радио и связь, 1989.
40. Унгер Х.-Г. Планарные и волоконные оптические волноводы. М.: Мир, 1980.
41. Сайт проф. Давыдова А.В. Сигналы и линейные системы <http://prodav.exponenta.ru/signals>