

А.В. ДАРЬЯН

ОБЩАЯ ТЕОРИЯ СВЯЗИ

Часть первая

Ереван–2017

ՀԱՅ-ՌՈՒՍԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

Ա.Վ. ԴԱՐՅԱՆ

ԿԱՊԻ ԸՆԴՀԱՆՈՒՐ ՏԵՍՈՒԹՅՈՒՆ

Դասախոսություններ

Մաս առաջին

Երևան

ՀՌՀ հրատարակչություն

2017

РОССИЙСКО-АРМЯНСКИЙ УНИВЕРСИТЕТ

А.В. ДАРЬЯН

ОБЩАЯ ТЕОРИЯ СВЯЗИ

Курс лекций
Часть первая

Ереван
Издательство РАУ
2017

УДК 621.391 (042.4)

ББК 32.811я7

Д 207

Печатается по решению РИС и НТС РАУ

Дарьян А.В.

Д 207 Общая теория связи: Курс лекций / А.В. Дарьян. – Ер.: РАУ, 2016. – 130 с.

Курс лекций соответствует учебному плану подготовки бакалавров по дисциплине «Общая теория связи» согласно государственному образовательному стандарту.

Цель первой части – подготовить читателя к изучению современной теории связи, которое требует привлечения новых понятий и методов, в том числе математических. В первую часть вошли те разделы курса, которые не требуют привлечения развитого аппарата теории вероятностей и математической статистики. Здесь обсуждаются общие вопросы построения звеньев систем связи, для понимания которых достаточны знания, основанные на теории цепей и радиотехнике. Понятие вероятности если и используется, то в той степени, в которой оно может быть применено без сложных разъяснений.

В первой части курса излагаются основные закономерности и методы передачи сообщений по каналам связи. Излагаются способы математического представления сообщений, сигналов и помех. Рассматриваются методы аналоговой и дискретной модуляции, передача дискретных сигналов, принципы многоканальной связи и помехоустойчивого кодирования.

Пособие предназначено для студентов высших учебных заведений, обучающихся по направлению «Радиотехника» и «Инфокоммуникационные технологии и системы связи».

УДК 621.391 (042.4)

ББК 32.811я7

Д 207

ISBN 978-9939-67-176-5

©Издательство РАУ, 2017

СОДЕРЖАНИЕ

Предисловие.....	7
1. Введение.....	9
2. Основные сведения из теории сигналов и цепей.....	20
2.1.Сигнал	20
2.2.Гармонический сигнал.....	22
2.3.Периодический сигнал. Ряд Фурье	25
2.4.Непериодический сигнал. Преобразование Фурье	29
2.5.Линейный четырехполюсник.....	35
2.5.1. Спектральное описание линейных цепей.....	37
2.5.2. Временной подход к описанию линейных цепей.....	38
2.6.Временная функция корреляции детерминированных сигналов.....	41
2.7.Квазигармоническое представление узкополосных сигналов.....	43
3. Аналоговая модуляция гармонического несущего сигнала.....	47
3.1. Амплитудная модуляция (АМ).....	48
3.2. Угловая модуляция (ФМ и ЧМ).....	52
3.3. Частотное уплотнение линий связи (Частотное разделение каналов)	56
4. Цифровое представление аналоговых сигналов.....	59
4.1. Амплитудно-импульсная Модуляция (АИМ)	59
4.2. Теорема Котельникова.....	62
4.3. Квантование.....	66
4.4. Импульсно-кодовая модуляция. Цифровой сигнал	69
4.5. Временное уплотнение линии (Временное разделение каналов)	72
4.5.1. Временное уплотнение сигналами АИМ	72
4.5.2. Временное уплотнение в системах с ИКМ	75
4.6. Регенерация дискретных сигналов.....	76
5. Передача дискретных сигналов в основной (первичной)полосе	80
5.1.Передача импульсов	80
5.2.Тактовая (символьная) синхронизация.....	85
5.2.1.Асинхронная передача	86

5.2.2. Синхронная передача	87
5.3. Передача последовательностей	90
5.4. Кодирование для линии	93
5.5. Частотная полоса	99
6. Дискретная модуляция гармонического сигнала (полосовая передача)	101
6.1. Двоичные методы модуляции	101
6.2. Недвоичная модуляция	103
7. Основы помехоустойчивого кодирования	108
7.1. Постановка задачи помехоустойчивого кодирования	108
7.2. Математический аппарат	111
7.3. Линейные коды	114
7.4. Простейшие линейные коды	118
7.4.1. Код с одной проверкой на чётность	118
7.4.2. Код Хемминга	119
7.5. Прочие методы построения корректирующих кодов	122
Заключение	124
Список рекомендуемой литературы по разделам	126
Список литературы	129

Предисловие

Телекоммуникации (связь), как явление, существует независимо от человека, как и все явления природы. Роль телекоммуникаций в природе еще плохо изучена, однако ее роль в живой природе хорошо видна каждому даже без какой-либо теоретической подготовки. Растения и животные, в том числе и человек, общаются (обмениваются информацией) самыми разнообразными способами. Тут задействованы механизмы различной природы: и химические, и механические (в том числе звуковые), и электромагнитные (в том числе оптические), и есть подозрения, что в живой природе действуют также методы, **неоснованные** на известных физических явлениях. Однако, несмотря на столь всеобъемлющую роль, связь (телекоммуникации) как техническая отрасль и как наука возникла сравнительно недавно. Началом современной техники связи можно считать создание первого практически пригодного телеграфного аппарата Шиллинга (1832г.). Что касается теории связи, то ее рождение связывают с работами Шеннона (1948г.), в которых были заложены основы математического подхода к проблемам передачи информации.

Предпосылкой к развитию информационных технологий следует считать формирование языкового общения между людьми. С созданием письменности появляется возможность хранения информации в таком виде, как мы это представляем в настоящее время. Далее возникает почтовая связь как средство передачи информации. И, хотя почтовая связь весьма далека от нашего нынешнего представления о связи (телекоммуникациях), именно в ней заложены принципы, используемые в современных сетях связи.

Предмет «Общая теория связи» рассматривает общие принципы передачи информации посредством физических носителей. Физические носители в технике связи – это электромагнитные волны, в том числе и световые. Круг вопросов, рассматриваемых в рамках данной теории, – это получение информации, формирование сиг-

налов, построение каналов связи, передача сигналов по каналам и выделение информации для предоставления ее получателю. Центральной задачей теории связи является обеспечение верности передачи, то есть соответствия полученной информации переданной. Это связано с тем, что сигналы при прохождении по системе связи подвергаются различным случайным воздействиям, которые невозможно ни предсказать, ни предотвратить. В результате, переданная информация частично теряется, и сигнал на приёмном конце оказывается искаженным. Отсюда возникает главная идея помехоустойчивости – передача большего количества информации чем требуется, таким образом, чтобы из информации, достигшей приемного конца, было возможно извлечь полезную информацию. Именно с целью осуществления этой идеи была разработана математическая теория Шеннона. Математический аппарат теории связи применяется также в технике хранения информации, которую можно рассматривать как «передачу во времени», при которой информация также может непредсказуемым образом искажаться.

При построении первой части курса «Общая теория связи» предполагалось, что студентами усвоены знания, приобретенные в рамках курсов «Основы теории цепей» и «Основы радиотехники». Тем не менее, в данное пособие введены разделы, включающие основные сведения из пройденных курсов, знание которых играет определяющую роль в усвоении теории связи.

Автор выражает благодарность рецензентам, члену-корреспонденту НАН РА, д.ф.н. А.А. Ахумяну и заведующему кафедрой «Радиоустройства» Национального Политехнического университета Армении, к.т.н., доценту О.А. Гомцяну, а также доценту кафедры «Системы связи» Национального Политехнического университета Армении, к.ф.-м.н.О.В. Багдасаряну за полезные обсуждения и критические замечания.

1. Введение

Основными объектами изучения в теории связи являются **источник, система связи и получатель** [1–§§1.1,1.2]. Система связи предназначена для передачи сообщения от источника получателю (Рис.1–1). Источник информации вырабатывает **сообщение**, в котором и заключается информация, подлежащая передаче. Сообщение может быть функцией нескольких переменных, в том числе времени. Примером такого сложного сообщения является движущееся изображение, для передачи которого предназначена телевизионная система передачи. Более простым является звуковое сообщение, которое представляет собой функцию времени (изменение во времени звукового давления на мембрану микрофона). Текстовое сообщение, если его представить как последовательность букв (а не как картинку), имеет самую простую структуру, так что его можно передавать с помощью примитивного телеграфного аппарата. Тоже относится к дискретным и цифровым данным (последовательностям нулей «0» и единиц «1»). Однако в настоящее время тексты и данные передаются не с помощью телеграфных аппаратов, поскольку последние не отвечают современным требованиям по скорости передачи и удобству в эксплуатации.

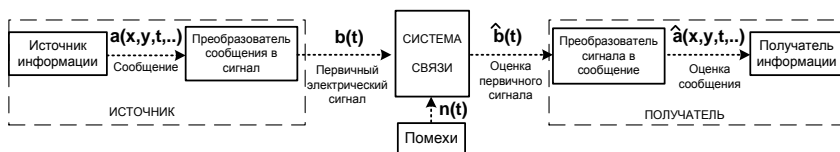


Рис. 1–1. Упрощенная схема передачи информации.

В систему связи поступает не сообщение, а его электрический образ, **первичный электрический сигнал**. Это, например, напряжение от микрофона, видеосигнал от видеокамеры, код букв или двоичных данных в виде импульсов и т.д. Первичный электрический сигнал, в отличие от сообщения, всегда есть функция одной

переменной – времени. Назначением системы связи является передача сигналов. Хотя основной задачей теории связи (телекоммуникаций) является построение систем связи, ее понятия и методы помогают решать задачи, относящиеся к источникам и получателям.

Поскольку сигнал является объектом транспортировки (передачи) по системе связи, целесообразно определить физические параметры сигнала, которые являются определяющими с точки зрения возможности его транспортировки [1–§ 1.1]. Такими параметрами являются **длительность сигнала** $T_{СИГ}$, **ширина частотного спектра** $F_{СИГ}$ и **динамический диапазон** $D_{СИГ}$, выраженный в децибелах:

$$D_{СИГ} = 10 \lg \frac{P_{СИГ.МАХ}}{P_{СИГ.МИН}}, \quad (1-1)$$

где $P_{СИГ.МАХ}$ и $P_{СИГ.МИН}$, соответственно, максимальная и минимальная мощности сигнала.

Параметром, обобщающим введённые три понятия, является **объем сигнала**:

$$V_{СИГ} = T_{СИГ} F_{СИГ} D_{СИГ}. \quad (1-2)$$

Объем сигнала дает общее представление о его способности содержать информацию и о возможности его передачи по системе связи. В рамках теории информации доказывается, что чем больше объем сигнала, тем большее количество информации можно в него «вложить» и тем труднее осуществить его передачу (транспортировку).

В системе связи действуют **помехи (шум)**. Помехи являются одним из основных объектов исследования теории связи [1–§ 1.3]. Под помехами мы понимаем все случайные изменения, которым

подвергается сообщение при прохождении через систему связи. В результате действия помех информация частично теряется, и на приемном конце бывает невозможно точно воспроизвести переданное сообщение. В результате, принятое сообщение $\hat{a}(x, y, t, \dots)$ отличается от переданного сообщения $a(x, y, t, \dots)$.

Не следует путать помехи (шум) с **искажениями**. Искажения зависят от параметров цепей и звеньев системы связи, так что они могут быть измерены, и их влияние может быть учтено при передаче сообщения и его восстановлении на приемном конце. В этом смысле искажения обычно не приводят к потерям информации. Более того, часто сам процесс передачи предполагает преднамеренное искажение сигналов при **передаче с** последующим восстановлением его на приемном конце.

Представленная на Рис.1–2. схема системы связи также упрощенная, но уже позволяет обсудить общие вопросы, касающиеся строения современных систем связи[1-§1.4].



Рис. 1–2. Более подробная схема системы связи.

Первичный электрический сигнал обычно кодируется (с помощью кодера), то есть представляется в некоторой форме, которую требует система передачи. Отчетливо это видно в цифровой системе передачи аналоговых сигналов, где аналоговый сигнал трансформируется в цифровую последовательность в двоичной форме. Более того, данная последовательность может быть дополнительно

перекодирована для удобства ее дальнейшей передачи. Далее следует модулятор, роль которого привести сигнал в такую форму, чтобы было возможно передать его по линии связи. Например, в радиосистемах модулятор переносит спектр сигнала в область высоких частот. Тем самым, сигнал размещается в нужной спектральной области и одновременно создается возможность его излучения в виде радиоволн посредством антенны.

Передачик предназначен для передачи сигнала в линию связи. Сигналы на выходе модулятора и передачика обозначены одинаково, так как с точки зрения структуры они одинаковы. Передачик просто технически осуществляет процесс передачи.

Следующее звено, изображенное на схеме, **это – линия связи –** физическая среда, в которой происходит передача сигнала в пространстве (свободное пространство, проводная или волноводная линия, в том числе оптическая линия, и т.п.) [1-§1.4]. В линии связи на сигнал воздействуют помехи, которые, как уже было отмечено, представляют собой случайные изменения, которым подвергается сигнал.

В общем случае сигнал на выходе линии можно записать в виде:

$$Z(t) = K \cdot S(t) + n(t), \quad (1-3)$$

где $S(t)$ – сигнал на входе линии, K – коэффициент передачи линии, а $n(t)$ – шум, действующий в линии, который называется аддитивным, поскольку прибавляется к сигналу.

То, что на схеме шум изображен только в линии связи, является упрощением, так как помехи могут действовать и в других звеньях системы связи. Однако, именно в линии связи сигнал наиболее подвержен действию помех.

Особым понятием в теории и технике связи является понятие канала связи [1-§1.4]. **Канал связи** формально определяется как «совокупность технических средств, обеспечивающих передачу сигнала между некоторыми двумя точками системы». Например, каналом связи является линия связи. Свойства линии связи как канала определяются физическими свойствами линии, которые, тем самым, определяют требования к устройствам, подсоединяемым на передающем и приёмном концах линии. В частности, как уже было отмечено, свойства линии могут привести к необходимости модуляции.

Если же рассматривать канал между входом модулятора и выходом демодулятора, то свойства этого канала определяются характеристиками модулятора и демодулятора. Для использования такого канала пользователь строит кодер и декодер в соответствии с входными и выходными параметрами канала. Процесс передачи по линии тут заложен в самой конструкции канала. Примером для сравнения может служить телефонный канал и физическая телефонная линия. Подключение интернета типа “dial-up” использует **телефонный канал**, то есть телефонную линию с оконечными устройствами телефонных станций. Это накладывает ограничения на возможности обеспечения высокой скорости интернета. Услуга же HDSL использует **физическую линию** телефонного канала для предоставления интернета, и поэтому использует все возможности линии, которые шире и позволяют повысить скорость передачи благодаря построению канала (оконечных устройств) в соответствии со своими требованиями.

В канале связи на сигнал воздействуют помехи, которые действуют как в линии связи, так и во всех звеньях канала [1-§§1.4,1.5]. Как уже было отмечено, помехи (шум) представляют собой случайные изменения, которым подвергается сигнал. Как и

в случае линии связи, в общем случае сигнал на выходе канала можно записать в виде

$$Z(t) = K \cdot S(t) + n(t), \quad (1-3)$$

где $S(t)$ – сигнал на входе канала, K – коэффициент передачи канала, а $n(t)$ – аддитивный шум, действующий в канале. Коэффициент передачи канала может быть или постоянным, (постоянный канал), или функцией времени (переменный канал), или случайным процессом (случайный канал). В последнем случае случайное изменение коэффициента передачи канала рассматривается как шум, который называется мультипликативным (мультипликативная помеха). Мультипликативный шум, так же как и аддитивный, приводит к потерям информации. Величина полезного сигнала на приемном конце (на входе приёмника) равна $S_{ПРМ} = K \cdot S(t)$, поэтому можно записать:

$$Z(t) = S_{ПРМ}(t) + n(t), \quad (1-4)$$

Предположим, что в канале действует только аддитивный шум. В этом случае определяющей характеристикой влияния шума на процесс приема сигнала является так называемое «отношение сигнал/шум» (ОСШ, signal-to-noise ratio, SNR) на входе приемника

$$ОСШ = P_{СИГ} / P_{Ш}, \quad (1-5)$$

где $P_{СИГ}$ и $P_{Ш}$ – соответственно, средние мощности сигнала и шума.

Все преобразования, производимые с сигналом перед передачей (на нашей схеме – кодирование и модуляция) должны быть обратимыми, что является показателем того, что они не приводят к потерям информации. Соответственно, производя на приемном конце соответствующие обратные преобразования (демодуляция и декодирование), мы получим переданный сигнал. Однако из-за действия помех очевидное преобразование, следующее из (1-3),

$$\hat{S}(t) = \frac{Z(t) - n(t)}{K}, \quad (1-6)$$

оказывается невозможным. А именно: нам неизвестна конкретная реализация аддитивного шума $n(t)$. И, тем более, если есть мультипликативная помеха, то нам неизвестен и коэффициент передачи K . И то и другое случайно, и не может быть заранее измерено. Мы можем измерить только статистические (вероятностные) характеристики помех, и на основе этого знания дать **статистическую оценку** $\hat{S}(t)$ переданного сигнала $S(t)$. В любом случае, помеха приводит к потерям информации, вследствие чего принятый сигнал отличается от переданного сигнала.

Естественная мера, которую нужно предпринять для того, чтобы потери информации не приводили к неправильному приему, является формирование и передача сигнала, содержащего большее количество информации, чем содержится в сообщении. Задача заключается в том, как именно привносить избыточную информацию в передаваемый сигнал, чтобы оказалось возможным на приёмном конце восстановить полезную информацию, несмотря на потери вследствие действия помех. И вторая задача: как именно восстанавливать нужную информацию. На интуитивном уровне данная идея применялась во все времена. Например, метод повторения, когда мы одно и то же повторяем несколько раз, иногда применяем синонимы и даже произносим слово по отдельным буквам. Однако несложный математический расчёт показал, что применение метода повторения в технике связи не позволяет сколько угодно уменьшить вероятность ошибочного приёма, так как при этом неограниченно возрастает время передачи (скорость передачи стремится к нулю). Поэтому долгое время существовало принципиальное заблуждение, что помехи неизбежно приводят к неправильному приему. В современной теории связи доказывается, что помехи ограничивают не точность передачи, а скорость передачи. А именно: сколь удобная точность достижима при ненулевой скорости передачи. Этот принципиальный вывод послужил толчком к бурному развитию теории и техники связи с начала пятидесятих годов XX века.

Для канала основными параметрами, определяющими его способность транспортировать сигналы, являются **длительность канала** $T_{КАН}$, **ширина полосы пропускания** $F_{КАН}$ и **динамический диапазон**, выраженный в децибелах $D_{КАН}$ [1-§1.2]. Здесь следует уточнить, что под динамическим диапазоном канала понимается отношение максимальной мощности сигнала к минимальной мощности сигнала, различимого на фоне шума канала. Принято считать, что минимальная мощность сигнала, различимого на фоне шума, равна мощности этого шума, и, соответственно, динамический диапазон, выраженный в децибелах

$$D_{КАН} = 10 \lg \frac{P_{МАХ}}{P_{МИН}} = 10 \lg \frac{P_{МАХ}}{P_{Ш}} \quad (1-7)$$

Параметром, обобщающим введённые три понятия, является **объем канала**:

$$V_{КАН} = T_{КАН} F_{КАН} D_{КАН}. \quad (1-8)$$

Объем канала дает общее представление о его способности передавать информацию. Чем больше объем канала, тем большее количество информации можно через него передать. В рамках теории информации доказывается, что **в каналах со слабым шумом объем канала численно равен максимальному количеству информации, которое можно передать через данный канал.**

Для того чтобы сигнал мог быть передан по каналу, очевидно, необходимо выполнение условия

$$V_{КАН} \geq V_{СИГ}. \quad (1-9)$$

Если при этом три условия $T_{КАН} \geq T_{СИГ}$, $F_{КАН} \geq F_{СИГ}$, $D_{КАН} \geq D_{СИГ}$ не выполняются одновременно, то сигнал можно преобразовать так, что эти условия выполняются. Такое преобразование называется обменом параметров сигнала. Пусть, например, $D_{КАН} = D_{СИГ}$, но $T_{КАН} = 2T_{СИГ}$, $F_{КАН} = 0,5F_{СИГ}$. При этом,

$V_{KAN} = V_{СИГ}$ и передача теоретически возможна. В этом случае необходимо растянуть сигнал во времени вдвое, при этом его спектр сузится вдвое и сигнал «поместится» в канале по всем трем измерениям. Так мы осуществляем обмен длительности на ширину полосы. В технике связи большой интерес представляет обмен полосы частот на динамический диапазон. В современных системах связи применяются различные методы расширения полосы сигнала, позволяющие работать при больших уровнях шума или с более слабыми сигналами (то есть с меньшим динамическим диапазоном) без ухудшения качества передачи. Исторически, первым шагом в этом направлении явилось применение частотной модуляции (FM) в радиосвязи. Именно эффективный обмен полосы частот на динамический диапазон позволяет реализовать потенциальные возможности систем связи, определяемые в рамках теории информации.

Основными характеристиками системы связи, как и канала, являются **скорость передачи и верность передачи**. Эти величины для дискретных и аналоговых систем определяются по-разному[1-§1.7].

Дискретные системы передают дискретные последовательности, то есть последовательности, элементы которых могут принимать дискретное множество значений, число которых будем обозначать m . В частности, при $m=2$ имеем двоичные сигналы, то есть последовательности из нулей и единиц (битов). Для передачи каждого элемента последовательности выделяется определенный промежуток времени T (тактовый интервал). **Скорость передачи** ν определяется в «Бодах» как число символов в секунду. Соответственно,

$$\nu(\text{Бод}) = \frac{1}{T}. \quad (1-10)$$

Существует также «**скорость телеграфирования**» R , показывающая, какой скорости передачи **двоичных** символов в секунду соответствует данная скорость передачи $\nu(\text{Бод})$. Поскольку при

перекодировке в двоичную форму каждому элементу m -значной последовательности требуется $k = \log_2 m$ двоичных символов (обычно число m равно степени двойки), то скорость телеграфирования будет:

$$R\left(\frac{\text{бит}}{с}\right) = \frac{\log_2 m}{T}. \quad (1-11)$$

В двоичном случае эти две скорости совпадают: $v = R$.

Двоичный источник выдает двоичные символы с некоторой скоростью R , означающей число двоичных символов (битов) в секунду. Не следует путать «бит», используемый здесь как «двоичный символ» с битом как единицей количества информации. Соотношение (1.11) определяет «техническую» скорость источника в отличие от **производительности источника**, определяемой в рамках теории информации как количество информации, выдаваемое источником в секунду.

Из-за действия шумов форма сигнала, представляющего данный дискретный символ, искажается, в результате чего на приемном конце есть вероятность того, что переданный символ будет неверно распознан. **Верность передачи в дискретной системе** характеризуется вероятностью ошибки $P_{\text{ош}}$. На практике определяется относительная доля неверно принятых символов от общего числа принятых символов:

$$BER = \frac{N_{\text{ош}}}{N_{\text{общ}}} = \frac{v_{\text{ош}}}{v}, \quad (1-12)$$

то есть относительная скорость ошибок (BER – bit error rate). Эта величина является статистической оценкой вероятности ошибки $P_{\text{ош}}$. Важным теоретическим результатом является то, что вероятность ошибки зависит от отношения сигнал/шум (1.5).

В отличие от дискретных сигналов, которые могут быть приняты или верно или неверно, аналоговый сигнал на приемном конце **всегда** отличается от переданного аналогового сигнала. Поэтому **показателем верности приема аналогового сигнала** является сте-

пень отличия принятого сигнала $\hat{S}(t)$ от переданного $S(t)$. В качестве математического критерия такого отличия чаще всего используется среднеквадратичное отклонение. Используя обозначения (1.4), можем записать среднеквадратичное отклонение в виде:

$$\varepsilon^2 = \frac{1}{T} \int_0^T [Z(t) - S_{\text{ПРМ}}(t)]^2 dt. \quad (1.13)$$

Предположим, что в канале действует только аддитивный шум $n(t)$. Тогда сигнал на выходе канала равен $Z(t) = S_{\text{ПРМ}}(t) + n(t)$. Среднеквадратичное отклонение будет равно:

$$\begin{aligned} \varepsilon^2 &= \frac{1}{T} \int_0^T [Z(t) - S_{\text{ПРМ}}(t)]^2 dt = \\ &= \frac{1}{T} \int_0^T [S_{\text{ПРМ}}(t) + n(t) - S_{\text{ПРМ}}(t)]^2 dt = \frac{1}{T} \int_0^T n(t)^2 dt = \bar{P}_{\text{ш}}, \end{aligned} \quad (1.14)$$

то есть равно средней мощности шума. Относительная ошибка при этом равна отношению мощности сигнала к мощности шума, так называемому **отношению сигнал/шум (SNR signal-to-noise ratio)**. Таким образом, отношение сигнал/шум на выходе канала в аналоговых системах также является важнейшей характеристикой, определяющей качество связи.

Что касается скорости передачи, то в аналоговых системах четкого определения нет. Очевидно лишь то, что чем шире полоса канала, тем бóльшую скорость передачи информации можно осуществить.

2. Основные сведения из теории сигналов и цепей

2.1. Сигнал

Как уже было отмечено, **сигнал**, содержащий информацию и подлежащий передаче, – это некоторая физическая (электрическая) величина, изменение которой несёт в себе информацию [1-§2.1, 2-Гл 1]. Сигнал математически может быть описан функцией одной переменной – времени $S(t)$. Это может быть, например, напряжение, однако в теории связи мы не конкретизируем данную физическую величину, это задача теории цепей. Поэтому в теории сигналов мощность сигнала $S(t)$ формально определяется как:

$$P(t) = S^2(t). \quad (2.1-1)$$

Можно сказать, что если $S(t)$ – напряжение или ток, то $P(t)$ равно мгновенной мощности, выделяемой на сопротивлении в 1 Ом.

Средняя мощность сигнала на конечном интервале равна:

$$P_{CP} = \frac{1}{T} \int_0^T S^2(t) dt. \quad (2.1-2)$$

Средняя мощность сигнала на неограниченном интервале равна

$$P_{CP} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T S^2(t) dt. \quad (2.1-3)$$

Если сигнал периодический, то среднюю мощность сигнала на неограниченном интервале можно вычислять по формуле (2.1-2), где T – период сигнала, то есть усреднять по одному (любому) периоду.

Энергия сигнала, определенного на интервале $[0, T]$, равна

$$E = \int_0^T S^2(t) dt. \quad (2.1-4)$$

Среднее значение (**постоянная составляющая**) сигнала, соответственно, определяется как:

$$S_{CP} = \frac{1}{T} \int_0^T S(t) dt \text{ или } S_{CP} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T S(t) dt . \quad (2.1-5)$$

Как и в случае мощности, для периодического сигнала постоянную составляющую можно вычислять усреднением по одному периоду.

В технике связи мощности и амплитуды принято измерять в относительных единицах децибелах (дБ). Децибел – это единица для описания отношения двух однородных физических величин. Мощность, выраженная в дБ, называется уровнем сигнала.

Если два сигнала имеют мощности P_0 и P , то второй сигнал имеет мощность по отношению к первому, выраженную в децибелах, равную:

$$\frac{P}{P_0} (\text{дБ}) = 10 \lg \frac{P}{P_0} . \quad (2.1-6)$$

Если условиться измерять мощность в дБ относительно 1Вт, то есть $P_0=1\text{Вт}$, то уровень сигнала с мощностью $P(\text{Вт})$ будет:

$$P(\text{дБВт}) = 10 \lg (P(\text{Вт})) . \quad (2.1-7)$$

Если измерять мощность относительно 1мВт, то будем иметь единицу дБм, то есть

$$P(\text{дБм}) = 10 \lg (P(\text{мВт})) . \quad (2.1-8)$$

Отношение амплитуд тока или напряжения также можно выражать в дБ, но при этом

$$\frac{U}{U_0} (\text{дБ}) = 20 \lg \frac{U}{U_0} . \quad (2.1-9)$$

Удвоение коэффициента связано с тем, что мощность пропорциональна квадрату напряжения (или тока), так что отношение амплитуд и отношение мощностей в децибелах выражается одним и тем же числом. Тут следует уточнить, что последнее имеет

место, если речь идёт об одной точке цепи или о разных цепях, но имеющих равные сопротивления.

Если рассматривать сигнал как бесконечномерный вектор (компоненты которого – его значения $S(t)$ в координатах t), то, по аналогии с конечномерными векторами, скалярное произведение двух сигналов $S_1(t)$ и $S_2(t)$ определится как:

$$(S_1(t) \cdot S_2(t)) = \int_0^T S_1(t) \cdot S_2(t) \cdot dt \quad (2.1-10)$$

Соответственно, два сигнала являются ортогональными, если их скалярное произведение равно нулю.

2.2. Гармонический сигнал

Гармонический сигнал [2-Гл1] описывается выражением

$$S(t) = A \cos(\omega t + \varphi), \quad (2.2-1)$$

где A – амплитуда, $\omega = 2\pi f$ – угловая частота, f – частота, а φ – фаза (начальная фаза). Частота f связана с периодом T соотношением $f = 1/T$. Выражение $\Phi(t) = \omega t + \varphi$, стоящее под знаком косинуса, называется обобщенной фазой сигнала.

Сигнал $S_1(t) = A \sin(\omega t + \varphi) = A \cos(\omega t + \varphi - \pi/2)$ также является гармоническим, и он отличается от сигнала (2.2-1) сдвигом фазы на $\pi/2=90^\circ$, то есть ортогонален ему. Соответственно, сигнал (2.2-1) можно записать в виде $S(t) = A \sin(\omega t + \varphi + \pi/2)$. Ортогональность этих двух сигналов соответствует определению (2.1-10).

Из (2.2-1) можно получить так называемое квадратурное представление гармонического сигнала:

$$\begin{aligned} S(t) &= A \cos(\omega t + \varphi) = A \cos \varphi \cos \omega t - \\ &- A \sin \varphi \sin \omega t = A_{c\varphi} \cos \omega t + A_{KB} \sin \omega t. \end{aligned} \quad (2.2-2)$$

Так мы представляем гармонический сигнал с произвольной фазой φ в виде суммы двух ортогональных составляющих – «син-фазной» (с нулевой начальной фазой) и «квадратурной» (ортогональной, с фазой 90°) и с амплитудами, соответственно, $A_{c\varphi} = A \cdot \cos \varphi$ и $A_{KB} = -A \cdot \sin \varphi$. Такое представление даёт способ практического построения сигнала с произвольной фазой путем сложения сигналов $\cos \omega t$ и $\sin \omega t$ с соответствующими коэффициентами.

Иногда полезным бывает наглядное представление гармонического сигнала (2.2–1) в виде вращающегося вектора (векторная диаграмма)[2-Гл1]. Длина вектора равна амплитуде сигнала A , угловая скорость вращения равна ω , и вектор начинает вращение из начального положения под углом φ к горизонтальной оси. За положительное направление вращения принимается направление против часовой стрелки (вращение по часовой стрелке означает $\omega < 0$). При таком представлении сигнал (2.2–1) является проекцией вектора на горизонтальную ось (Рис.2.2–1).

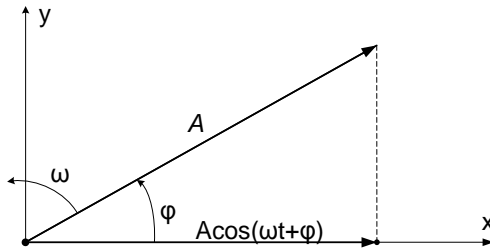


Рис.2.2–1. Векторная диаграмма сигнала (2.2–1).

Этот вращающийся вектор можно представить в комплексной форме, если принять, что вертикальная ось мнимая:

$$\begin{aligned} \dot{S}(t) &= A \cdot e^{i(\omega t + \varphi)} = \dot{A}(t) \cdot e^{i\omega t} = \\ &= A(\cos(\omega t + \varphi) + i \cdot \sin(\omega t + \varphi)), \end{aligned} \tag{2.2–3}$$

где $\dot{A}(t) = A \cdot e^{i\varphi}$ – комплексная амплитуда, представляющая амплитуду и фазу сигнала. Сигнал $\dot{S}(t)$ называется **комплексным**

сигналом, соответствующим вещественному сигналу $S(t)$. При этом действительный сигнал выражается через соответствующий ему комплексный сигнал как

$$S(t) = \operatorname{Re} \dot{S}(t) = \frac{1}{2} (\dot{S}(t) + \dot{S}^*(t)), \quad (2.2-4)$$

где звездочка означает комплексное сопряжение.

Из (2.2-4) можно непосредственно получить представление сигнала (2.2-1) в **комплексной форме**, а именно:

$$A \cos(\omega t + \varphi) \equiv \frac{A}{2} (e^{i(\omega t + \varphi)} - e^{-i(\omega t + \varphi)}). \quad (2.2-5)$$

Векторная диаграмма, соответствующая представлению (2.2-5), дана на Рис.2.2-2. Действительный вектор $A \cos(\omega t + \varphi)$ равен сумме двух комплексных векторов, то есть векторов, вращающихся в противоположных направлениях – с частотами ω и $-\omega$ и амплитудами $A/2$.

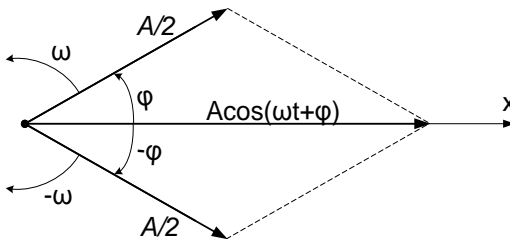


Рис.2.2-2. Векторная диаграмма сигнала (2.2-1) в комплексной форме.

Мощность гармонического сигнала получается из формулы (2.1-2) путем усреднения по периоду $T=2\pi/\omega$:

$$\begin{aligned} P_{CP} &= \frac{1}{T} \int_0^T S^2(t) dt = \\ &= \frac{1}{T} \int_0^T A^2 \cdot \cos^2(\omega t + \varphi) dt = \frac{A^2}{2} = A_D^2, \end{aligned} \quad (2.2-6)$$

где $A_d = \frac{A}{\sqrt{2}}$ – действующее значение гармонического сигнала (тока или напряжения).

Среднее значение гармонического сигнала (постоянная составляющая) равно нулю (формула (2.1–5)).

Гармонический сигнал обладает важным свойством, что, проходя через линейные цепи (линейный четырехполюсник) он остается гармоническим сигналом с той же частотой, измениться может только его амплитуда и фаза. Это свойство, а также выполнение принципа суперпозиции (наложения) в линейных цепях делает спектральное описание сигналов эффективным инструментом анализа линейных цепей.

В нелинейных цепях форма сигнала меняется, и гармонический сигнал перестает быть гармоническим. Если, например, нелинейная цепь осуществляет квадратичное преобразование, то на выходе будем иметь постоянную составляющую и «вторую гармонику» – гармонический сигнал с частотой, равной удвоенной частоте входного сигнала. Математически это выглядит следующим образом: $(\sin \omega t)^2 = \frac{1}{2} - \frac{1}{2} \cos 2\omega t$.

2.3. Периодический сигнал. Ряд Фурье

Периодом сигнала $S(t)$ называется наименьшее число T такое, что

$$S(t) = S(T + nT), (-\infty < t < +\infty, n = 1, 2, \dots). \quad (2.3-1)$$

Периодический сигнал можно разложить в ряд Фурье [2-Гл2], то есть представить в виде суммы гармонических составляющих:

$$S(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} A_n \cos(n\omega_1 t + \varphi_n), \quad \omega_1 = \frac{2\pi}{T} = 2\pi f_1 \quad (2.3-2)$$

Частота f_1 (или ω_1) называется **основной частотой** данного периодического сигнала, а гармоническая составляющая

$A_1 \cos(\omega_1 t + \varphi_1)$ называется **первой (основной) гармоникой** сигнала. Коэффициенты A_n образуют **амплитудный спектр**, фазы φ_n — **фазовый спектр**, а частоты $n\omega_1$ — **частотный спектр**. Как мы видим, частотный спектр **периодического** сигнала содержит только **частоты, кратные основной частоте** ω_1 . Величина $\frac{a_0}{2}$ является постоянной составляющей сигнала и, соответственно, вычисляется по формуле:

$$\frac{a_0}{2} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} S(t) \cdot dt. \quad (2.3-3)$$

Для вычисления амплитудного и фазового спектров представим гармоники ряда Фурье в квадратурной форме:

$$S(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} (a_n \cos n\omega_1 t + b_n \sin n\omega_1 t). \quad (2.3-4)$$

Коэффициенты a_n и b_n данного разложения вычисляются по известным формулам:

$$a_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} S(t) \cos n\omega_1 t \cdot dt, \quad b_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} S(t) \sin n\omega_1 t \cdot dt, \quad (2.3-5)$$

при этом

$$A_n = \sqrt{a_n^2 + b_n^2}, \quad \varphi_n = \arccos \frac{a_n}{\sqrt{a_n^2 + b_n^2}}. \quad (2.3-6)$$

Выражения (2.3-3), (2.3-5) и (2.3-6) применяются для математического определения спектра произвольного периодического сигнала.

Возможность разложения функции $S(t)$ в сумму вида (2.3-4) обусловлено тем, что тригонометрические функции $\cos n\omega_1 t$ и $\sin n\omega_1 t$, $n = 0, 1, 2, \dots, \infty$ образуют полную систему функций, орто-

гональных на интервале T , что математически выражается следующим образом:

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \cos n\omega_1 t \cdot \sin m\omega_1 t \cdot dt = 0, \quad n, m = 1, 2, \dots \infty. \quad (2.3-7a)$$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \cos n\omega_1 t \cdot \cos m\omega_1 t \cdot dt = 0, n \neq m, \quad n, m = 1, 2, \dots \infty \quad (2.3-7б)$$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \sin n\omega_1 t \cdot \sin m\omega_1 t \cdot dt = 0, n \neq m \quad n, m = 1, 2, \dots \infty \quad (2.3-7в)$$

Таким образом, выражение (2.3–3) представляет собой разложение бесконечномерного вектора (функции) $S(t)$ по ортогональной системе тригонометрических функций, аналогичное разложению трехмерного вектора на координаты по трем взаимно-перпендикулярным осям.

Ряд Фурье является частным случаем разложения произвольных функций по ортогональному базису. В дальнейшем будут рассмотрены разложения по некоторым ортогональным системам функций, представляющим интерес для теории связи.

Если в выражение (2.1–2) для средней мощности сигнала подставить его спектральное разложение (2.3–2), то получим выражение, которое можно рассматривать как спектральное разложение мощности сигнала (см. (2.2–6)):

$$P_{cp} = \left(\frac{a_0}{2}\right)^2 + \frac{1}{2} \sum_{n=1}^{+\infty} A_n^2 = \left(\frac{a_0}{2}\right)^2 + \sum_{n=1}^{+\infty} A_{Дn}^2 \quad (2.3-8)$$

Данное выражение показывает, что **мощность периодического сигнала равна сумме мощностей его гармонических составляющих**, а также не зависит от соотношения их фаз. При выводе выражения (2.3–8) существенно используется ортогональность гармонических составляющих на интервале T – оно является

многомерной теоремой Пифагора. В общем случае произвольных сигналов принцип суперпозиции для мощностей не выполняется.

Используя комплексное представление гармоник (2.2–5), получаем комплексную форму ряда Фурье:

$$S(t) = \sum_{n=-\infty}^{+\infty} C_n e^{i(n\omega_1 t + \phi_n)} = \sum_{n=-\infty}^{+\infty} \dot{C}_n e^{in\omega_1 t}, \dot{C}_n = C_n \cdot e^{i\phi_n}, \quad (2.3-9)$$

Здесь C_n – амплитудный спектр (двусторонний), ϕ_n – фазовый спектр, связанный с коэффициентами (2.3-6) соотношениями

$$C_n = \begin{cases} \frac{A_n}{2}, & n > 0 \\ \frac{A_{-n}}{2}, & n < 0 \end{cases}, \quad \phi_n = \begin{cases} \varphi_n, & n > 0 \\ -\varphi_{-n}, & n < 0 \end{cases}, \quad C_0 = \frac{a_0}{2} \quad (2.3-10)$$

Соотношение (2.3–8) в коэффициентах C_n комплексного разложения принимает более компактный вид:

$$P_{cp} = \sum_{n=-\infty}^{+\infty} C_n^2 \quad (2.3-11)$$

Комплексный спектр периодического сигнала состоит из гармоник с положительными и отрицательными частотами (двухсторонний спектр), амплитуды которых равны половине амплитуд вещественного (одностороннего) спектра. Односторонний спектр есть реальный, физический спектр, а комплексный спектр – математическая абстракция. В качестве примера на рисунке (2.3–1) изображена временная диаграмма периодической последовательности прямоугольных импульсов и приведены диаграммы её амплитудных спектров – одностороннего и двухстороннего.

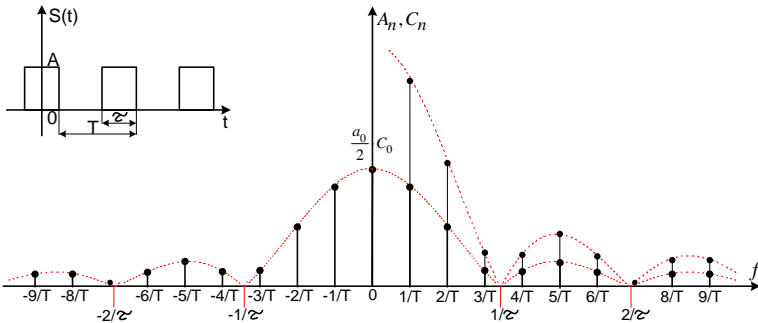


Рис.2.3–1. Временная диаграмма периодической последовательности прямоугольных импульсов и диаграммы её амплитудных спектров – одностороннего и двухстороннего.

2.4. Непериодический сигнал. Преобразование Фурье

В отличие от периодического сигнала, непериодический сигнал разлагается не в сумму гармонических составляющих, а в интеграл [2-Гл2], который называется **обратным преобразованием Фурье**

$$S(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \dot{S}(\omega) e^{i\omega t} d\omega \tag{2.4-1}$$

Комплексная функция $\dot{S}(\omega)$ называется **спектральной плотностью** сигнала $S(t)$ и вычисляется по формуле:

$$\dot{S}(\omega) = \int_{-\infty}^{+\infty} S(t) e^{-i\omega t} dt, \tag{2.4-2}$$

которая является **прямым преобразованием Фурье**. Таким образом, спектр непериодического сигнала непрерывный, содержащий непрерывный ряд частот.

Такое описание математически применимо к классу так называемых абсолютно интегрируемых сигналов, то есть удовлетворяющих условию:

$$\int_{-\infty}^{+\infty} |S(t)| \cdot dt < \infty. \quad (2.4-3)$$

Однако можно наложить на рассматриваемые сигналы более слабое условие интегрируемости в квадрате:

$$\int_{-\infty}^{+\infty} |S(t)|^2 \cdot dt < \infty. \quad (2.4-4)$$

Это условие означает, что сигнал имеет конечную энергию (см. (2.1–4)). На практике для сигналов с конечной энергией данное спектральное описание чаще всего применимо.

Спектральную плотность можно представить в виде

$$\dot{S}(\omega) = S(\omega) \cdot e^{i\varphi(\omega)}, \quad (2.4-5)$$

где $S(\omega) = |\dot{S}(\omega)|$ – амплитудный спектр, $\varphi(\omega)$ – фазовый спектр. Доказывается, что $S(\omega)$ чётная функция, а $\varphi(\omega)$ – нечётная:

$$S(-\omega) = S(\omega), \quad \varphi(-\omega) = -\varphi(\omega). \quad (2.4-6)$$

Поэтому можно определить одностороннюю (физическую) спектральную плотность амплитуды как

$$S_0(\omega) = \begin{cases} 2S(\omega), & \omega \geq 0 \\ 0, & \omega < 0. \end{cases} \quad (2.4-7)$$

Используя соотношение (2.2–2) и учитывая (2.4–5), (2.4–6) и (2.4–7), разложение (2.4–1) можно представить в виде

$$S(t) = \frac{1}{\pi} \int_0^{\infty} [a(\omega) \cos(\omega t) + b(\omega) \sin(\omega t)] d\omega \quad (2.4-8)$$

Аналогично соотношению (2.3–7) для периодических сигналов, для сигналов, описываемых с помощью преобразования Фурье, имеет место соотношение

$$E = \int_{-\infty}^{+\infty} S^2(t) \cdot dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S^2(\omega) \cdot d\omega, \quad (2.4-9)$$

являющееся частным случаем **равенства Парсеваля**. Данное соотношение показывает, что функцию $S^2(\omega)$ можно рассматривать как спектральную плотность энергии сигнала.

Учитывая (2.4-6), получаем;

$$\begin{aligned} E &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} S^2(\omega) \cdot d\omega = \\ &= \frac{1}{2\pi} \int_0^{+\infty} 2S^2(\omega) \cdot d\omega = \frac{1}{2\pi} \int_0^{+\infty} S_D^2(\omega) \cdot d\omega, \end{aligned} \quad (2.4-10)$$

откуда следует, что функцию $\sqrt{2} \cdot S(\omega) = S_0(\omega) / \sqrt{2} = S_D(\omega)$ можно назвать односторонней спектральной плотностью действующих значений. Физический смысл данной функции заключается в том, что величина

$$\frac{1}{2\pi} S_D^2(\omega) \cdot d\omega = E(\omega, \omega + d\omega), \quad (2.4-11)$$

показывает суммарную энергию спектральных составляющих, лежащих в интервале $(\omega, \omega + d\omega)$. Аналогичное объяснение для амплитудного спектра $S(\omega)$ или $S_0(\omega)$ не имеет смысла.

Математические свойства пары преобразований Фурье чрезвычайно важны для приложений, в том числе в теории связи, однако их мы будем обсуждать по мере необходимости. Особо отметим только одно свойство, согласно которому:

– **сигнал с конечной длительностью имеет бесконечную ширину спектра;**

– **сигнал с конечной шириной спектра имеет бесконечную длительность.**

Это свойство особенно тем, что оно на практике не выполняется – сигналы всегда имеют конечную длительность и конеч-

ную ширину спектра. Поэтому приходится вводить понятия эффективной ширины спектра и эффективной длительности сигнала.

Эффективной длительностью сигнала называется длительность интервала, в котором заключена основная энергия сигнала. На практике для импульсных сигналов «основная» энергия принимается равной 90% всей энергии сигнала.

Эффективной шириной спектра сигнала называется ширина спектрального диапазона, в котором заключена основная энергия сигнала. Обычно и в этом случае «основная» означает 90%.

Определение эффективных параметров сигнала произвольно и зависит от конкретной практической задачи. Однако тут можно сформулировать общую закономерность: **эффективная ширина спектра сигнала обратно пропорциональна эффективной длительности сигнала:**

$$\Delta f_{эфф} \approx \frac{1}{\tau_{эфф}} \quad (2.4-12)$$

Данное соотношение является аналогом соотношения неопределенностей в квантовой физике, которое также является следствием свойств преобразования Фурье. Что касается коэффициента пропорциональности, то он зависит от вида самого сигнала и от способа определения эффективных параметров.

Ограничение длительности сигнала с ограниченным спектром или спектра сигнала конечной длительности означает искажение сигнала. Допустимая степень искажения определяется из практических соображений.

В качестве примера спектральной плотности, на Рис.2.4–1 приведены временная и спектральная диаграммы прямоугольного импульса (прямоугольного **видеоимпульса**). Отрицательные значения спектральной плотности означают противоположную фазу, то есть скачок фазы на π в нулях функции. По существу, спектральная диаграмма нарисована так, что она содержит как амплитудный, так и фазовый спектр. Если использовать функцию « $\text{sinc}(x)$ », которая определяется как:

$$\text{sinc}(x) = \frac{\sin x}{x}, \quad (2.4-13)$$

то спектр прямоугольного импульса $S(f) = \tau A \cdot \text{sinc}\left(\frac{2\pi f \tau}{2}\right)$.

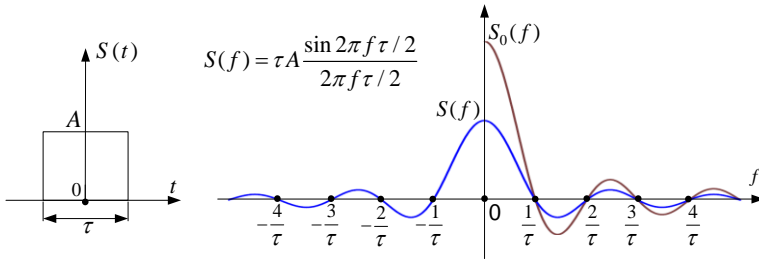


Рис.2.4–1. Прямоугольный импульс и его спектральная плотность.

В «основном лепестке» одностороннего (физического) спектра – в интервале $(0, 1/\tau)$ содержится около 90% энергии импульса. Поэтому за эффективную ширину спектра прямоугольного импульса часто принимается величина $\Delta f_{эфф} = 1/\tau$. Такое ограничение спектра импульса означает его искажение. На Рис.2.4–2 показана временная диаграмма импульса при ограничении спектра величиной $\Delta f_{эфф}$. Допустимая степень искажения прямоугольного импульса, а тем самым и допустимая степень ограничения спектра, определяется из практических соображений.

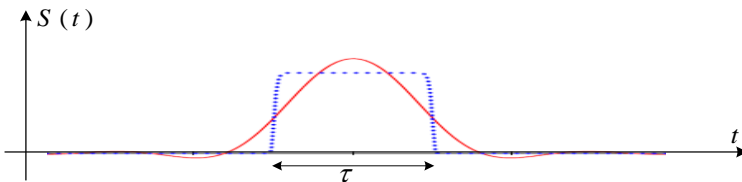


Рис.2.4–2. Искажение прямоугольного импульса при ограничении его спектра величиной $1/\tau$.

На Рис.2.4–3 показаны временная и спектральная диаграммы прямоугольного радиоимпульса. Этот сигнал, фактически, является отрезком синусоиды длительностью τ . «Прямоугольный» означает, что огибающая этой синусоиды – прямоугольный импульс (штриховая линия на графике). Спектр прямоугольного радиоимпульса по виду аналогичен спектру прямоугольного видеоимпульса (Рис.2.4–1), однако сдвинут на величину частоты синусоиды – **высокочастотного наполнения** импульса $f_{\text{НАП}} = 1/T$.

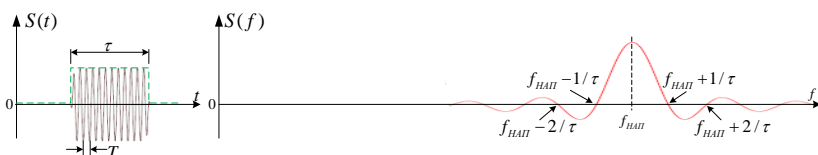


Рис.2.4–3 Прямоугольный радиоимпульс и его спектр.

Ширина спектра радиоимпульса вдвое больше ширины спектра видеоимпульса той же длительности. Если ограничить спектр пределами $f_{\text{НАП}} \pm 1/\tau$, то огибающая радиоимпульса исказится, как показано на Рис.2.4–2, а радиоимпульс примет вид, приведённый на Рис.2.4–4.

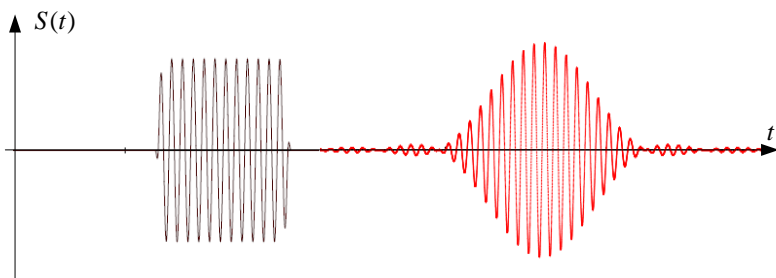


Рис.2.4–4. Искажение прямоугольного радиоимпульса при ограничении спектра пределами $f_{\text{НАП}} \pm 1/\tau$.

На Рис.2.4–5 приведен еще один важный пример – сигнал с равномерным ограниченным спектром, так называемый «идеальный низкочастотный сигнал». Данный сигнал называется также «отсчетной» функцией. Она играет важную роль при рассмотрении дискретизации сигналов.

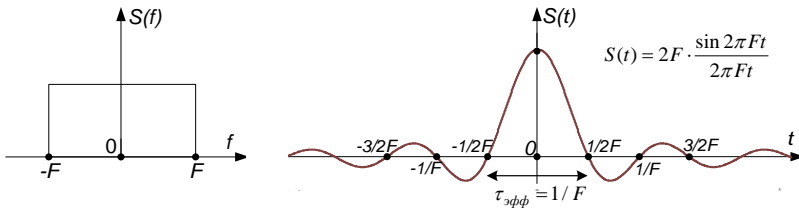


Рис.2.4–5. Спектральная и временная диаграммы «отсчётной функции».

В «основном лепестке» отсчётной функции – в интервале $(-1/2F, 1/2F)$, содержится 90% энергии. Поэтому за эффективную длительность отсчётной функции часто принимается величина $\tau_{эфф} = 1 / F$.

2.5. Линейный четырехполюсник

Линейный четырёхполюсник (и вообще линейная цепь) – это электрическая цепь, состоящая из линейных элементов – резисторов R , конденсаторов C и катушек индуктивности L . Цепи, содержащие нелинейные элементы, например, транзисторы, также можно рассматривать как линейные, если они работают на линейных участках вольтамперной характеристики. **Критерием линейности** элемента является линейная зависимость между **воздействием** и **откликом** [2-§8.2]. Например, линейный резистор представляет собой активное сопротивление, то есть подчиняется закону Ома, а реактивные элементы C и L подчиняются закону Ома для цепи переменного тока:

$$I = \frac{U}{R}, \quad I_C = i\omega C\dot{U}, \quad I_L = \frac{\dot{U}}{i\omega L} \quad (2.5-1)$$

Коэффициенты пропорциональности во всех этих зависимостях между током и напряжением являются характеристиками соответствующих элементов (R, C и L). Эти характеристики постоянны, не зависят от воздействия (напряжения или тока), вследствие чего зависимости (2.5–1) линейны. Поэтому можно дать другое определение линейного элемента: элемент является линейным, если его **характеристики не зависят от воздействия**.

В линейных цепях соблюдается **принцип суперпозиции (наложения)**, согласно которому отклик цепи на сумму воздействий равен сумме откликов на каждое из воздействий. Как следствие, увеличение воздействия в k раз увеличит отклик во столько же раз. Если, например, будем рассматривать линейную цепь как четырех-полосник, то принцип суперпозиции выразится, как изображено на Рис.2.5–1. Здесь воздействие – это входной сигнал $S(t)$, а отклик – соответствующий выходной сигнал $Z(t)$.

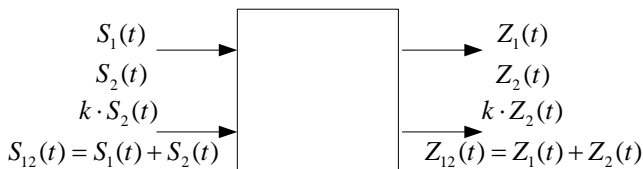


Рис.2.5–1. Суммарное воздействие на линейную цепь вызывает сумму откликов.

Задача определения отклика на произвольное воздействие сама по себе довольно сложна. На принципе суперпозиции основаны общие методы определения **отклика на произвольное воздействие**. А именно, если мы можем представить произвольное воздействие суммой «элементарных» воздействий – воздействий, отклики на которые легко описываются и нам известны, то мы сможем найти суммарный отклик как сумму откликов на «элементарные» воздействия. Один из таких методов – спектральный. В этом методе элементарным воздействием является гармоническая функция. При временном подходе в качестве «элементарного» воздействия выбирается функция Хэвисайда $\sigma(t)$ (единичный скачок или

единичная функция) или ее производная – δ -функция Дирака (единичный импульс), изображённая на Рис. 2.5–2.

2.5.1. Спектральное описание линейных цепей

Практически любой сигнал можно представить в виде суммы гармонических составляющих – в виде ряда или в виде интеграла Фурье. Поскольку гармонический сигнал при прохождении через линейные цепи остаётся гармоническим с той же частотой, то достаточно знать, как изменяется при прохождении через четырех-полосник амплитуда и фаза сигнала данной частоты [2-§8.2]. Таким образом, если на вход подадим сигнал

$$\dot{S}_{BX}(t) = A_{BX}(t) \cdot e^{i(\omega t + \varphi_{BX})}, \quad (2.5-2)$$

то на выходе получим

$$\dot{S}_{BLY}(t) = A_{BLY}(t) \cdot e^{i(\omega t + \varphi_{BLY})}. \quad (2.5-3)$$

Характеристикой линейного четырёхполосника является его **передаточная функция**

$$\dot{K}(\omega) = \frac{\dot{S}_{BLY}(\omega)}{\dot{S}_{BX}(\omega)} = K(\omega) e^{i\varphi(\omega)}, \quad (2.5-4)$$

модуль которой – $K(\omega)$ есть его амплитудно-частотная характеристика (**АЧХ**), которая показывает, каков коэффициент передачи для сигнала с частотой ω , а $\varphi(\omega)$ фазо-частотная характеристика (**ФЧХ**), показывающая сдвиг фазы выходного сигнала относительно входного, для сигнала с **частотой ω** :

$$K(\omega) = \frac{A_{BLY}(\omega)}{A_{BX}(\omega)}, \quad \varphi(\omega) = \varphi_{BLY}(\omega) - \varphi_{BX}(\omega). \quad (2.5-5)$$

Сигнал на входе можно представить в виде интеграла Фурье,

$$S_{BX}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \dot{S}_{BX}(\omega) e^{i\omega t} d\omega. \quad (2.5-6)$$

Из (2.5–4) следует, что спектр выходного сигнала равен:

$$\dot{S}_{\text{ВЫХ}}(\omega) = \dot{S}_{\text{ВХ}}(\omega) \cdot \dot{K}(\omega), \quad (2.5-7)$$

и, следовательно, интеграл Фурье для выходного сигнала принимает вид:

$$\begin{aligned} S_{\text{ВЫХ}}(t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \dot{S}_{\text{ВЫХ}}(\omega) e^{i\omega t} d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \dot{S}_{\text{ВХ}}(\omega) \dot{K}(\omega) e^{i\omega t} d\omega \end{aligned} \quad (2.5-8)$$

Здесь мы интегрируем спектральные (гармонические) составляющие с учетом их преобразования в цепи.

Для того чтобы форма сигнала не искажалась при прохождении через четырёхполосник, требуется, чтобы амплитуды всех частот менялись одинаково, а фаза была линейной функцией частоты:

$$K(\omega) = \text{const.}, \varphi(\omega) = \tau\omega, (2.5-9)$$

где τ – задержка сигнала в цепи.

2.5.2. Временной подход к описанию линейных цепей

При временном подходе [2-§8.2] в качестве «элементарного» воздействия выбирается **единичная функция** $\sigma(t)$ (единичный скачок, или функция Хэвисайда) или её производная – **единичный импульс** (Рис.2.5–2).

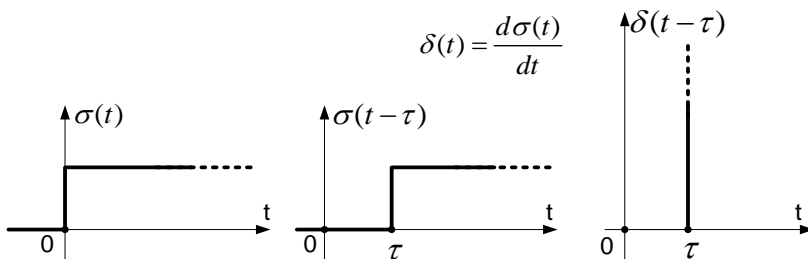


Рис.2.5–2. Определение единичной функции $\sigma(t)$ и единичного импульса $\delta(t)$.

Единичный импульс представляет собой бесконечно высокий и бесконечно узкий импульс и описывается δ -функцией Дирака, которая определяется следующими соотношениями:

$$\delta(t - \tau) = \begin{cases} \infty, t = \tau \\ 0, t \neq \tau \end{cases}, \int_a^b \delta(t - \tau) dt = \begin{cases} 1, \tau \in [a, b] \\ 0, \tau \notin [a, b] \end{cases},$$

$$\int_{-\infty}^{\infty} f(t) \delta(t - \tau) dt = f(\tau). \quad (2.5-10)$$

Переходной характеристикой $h(t)$ называется отклик цепи на единичную функцию. Это значит, что если на вход подать $\sigma(t)$, то на выходе получим $h(t)$.

Импульсной характеристикой или **импульсной реакцией** $g(t)$ называется отклик цепи на единичный импульс. Эти две характеристики связаны соотношением

$$g(t) = \frac{dh(t)}{dt} = h'(t) \quad (2.5-11)$$

На Рис.2.5-3 изображено приближение сигнала наложением (ступенек) разной высоты и с различными временными сдвигами.

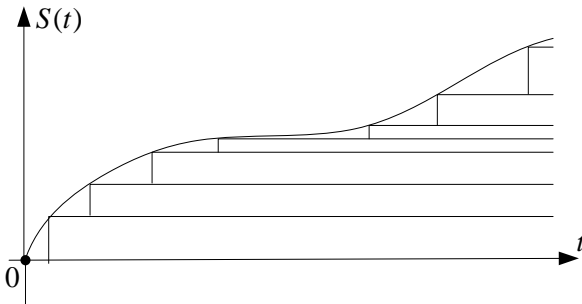


Рис.2.5-3. Приближение сигнала наложением ступенек.

Если перейти к пределу бесконечно малых ступенек, то высота элементарной ступеньки, соответствующая элементу времени $d\tau$, равна $ds = s'(t)d\tau$, где штрих означает знак производной. Тогда

входное воздействие можно выразить через единичную функцию следующим образом:

$$s_{BX}(t) = \int_0^t s'_{BX}(\tau) \sigma(t - \tau) d\tau \quad (2.5-12)$$

Принято, что $t = 0$ соответствует началу входного сигнала.

Отклик цепи на элементарную ступеньку $s'(t)\sigma(t - \tau)d\tau$, по определению, равен, $s'(t)h(t - \tau)d\tau$, так что весь выходной сигнал будет:

$$s_{BbIX}(t) = \int_0^t s'_{BX}(\tau) h(t - \tau) d\tau \quad (2.5-13)$$

Данное выражение является одним из вариантов **интеграла Дюамеля**, выраженным через переходную характеристику. Путем математических преобразований получают ещё три эквивалентные формы интеграла Дюамеля, выражающие выходной сигнал с помощью как переходной $h(t)$, так и импульсной $g(t)$ характеристики:

$$\begin{aligned} s_{BbIX}(t) &= \int_0^t s'_{BX}(t - \tau) h(\tau) d\tau = \\ &= \int_0^t s_{BX}(t - \tau) g(\tau) d\tau = \int_0^t s_{BX}(\tau) g(t - \tau) d\tau \end{aligned} \quad (2.5-14)$$

Передаточная функция и импульсная характеристика линейного четырехполосника связаны парой преобразований Фурье:

$$g(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \dot{K}(\omega) e^{i\omega t} d\omega, \quad \dot{K}(\omega) = \int_0^{\infty} g(t) e^{-i\omega t} dt. \quad (2.5-15)$$

В качестве примера на **Рис.2.5-4** приведены АЧХ (односторонняя) и импульсная характеристика $g(t)$ идеального фильтра нижних частот (ФНЧ). Итак, подавая на вход идеального ФНЧ единичный импульс, на выходе получаем отсчётную функцию (сигнал с равномерным ограниченным спектром) **(Рис.2.5-3)**.

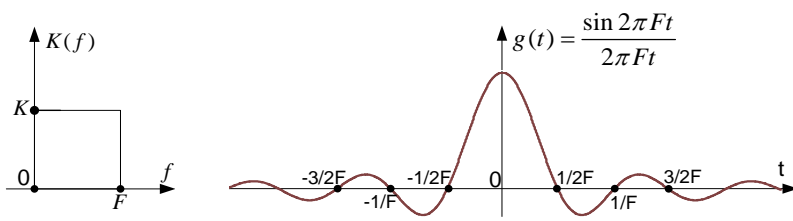


Рис.2.5—4. АЧХ и импульсная характеристика идеального ФНЧ.

2.6. Временная функция корреляции детерминированных сигналов

В данном разделе даются определения и некоторые свойства автокорреляционной функции (АКФ) детерминированного сигнала и взаимокорреляционной функции (ВКФ) двух детерминированных сигналов[2-Гл3]. В обоих случаях имеются в виду **временные функции корреляции**. Уточнение «детерминированных» связано с тем, что понятие корреляции и определение соответствующих функций относится к случайным сигналам. В отношении случайных сигналов корреляционная функция позволяет делать некоторые выводы о зависимости (статистической) между сигналами. Лишь в случае так называемых эргодических случайных сигналов функции корреляции вычисляются как **временны** **е** функции корреляции, которые в настоящем разделе рассматриваются для детерминированных сигналов.

Автокорреляционная функция (АКФ) детерминированного сигнала с ограниченной энергией $S(t)$ определяется как:

$$\psi(\tau) = \int_{-\infty}^{\infty} S(t) \cdot S(t + \tau) \cdot dt = (S(t) \cdot S(t - \tau)) \quad (2.6-1)$$

то есть АКФ в точке τ равна скалярному произведению сигнала и его временного сдвига на величину τ . АКФ обладает двумя очевидными свойствами:

$$\psi(\tau) = \psi(-\tau), \quad (2.6-2a)$$

$$\psi(0) = E, \quad (2.6-2б)$$

где E – энергия сигнала.

Доказывается также, что максимальное значение АКФ достигается при $\tau = 0$, а точнее:

$$\psi(\tau) \leq \psi(0) \quad (2.6-2в)$$

Временная функция взаимной корреляции двух сигналов с ограниченной энергией $S_1(t)$ и $S_2(t)$ определяется как:

$$\psi_{12}(\tau) = \int_{-\infty}^{\infty} S_1(t) \cdot S_2(t - \tau) \cdot dt \quad (2.6-3)$$

Взаимокорреляционная функция свойствами (2.6–2а, б, в) не обладает.

Автокорреляционная функция и энергетический спектр сигнала связаны парой преобразований Фурье:

$$\left| \dot{S}(\omega) \right|^2 = \int_{-\infty}^{\infty} \psi(\tau) \cdot e^{-i\omega\tau} d\tau, \quad \psi(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \dot{S}(\omega) \right|^2 \cdot e^{i\omega\tau} d\omega \quad (2.6-4)$$

По аналогии, преобразование Фурье от взаимокорреляционной функции называется взаимным (энергетическим) спектром:

$$\left| \dot{S}_{12}(\omega) \right|^2 = \int_{-\infty}^{\infty} \psi_{12}(\tau) \cdot e^{-i\omega\tau} d\tau, \\ \psi_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \dot{S}_{12}(\omega) \right|^2 \cdot e^{i\omega\tau} d\omega \quad (2.6-5)$$

Нетрудно доказать, что если энергии сигналов равны, $E_1 = E_2$, то взаимная функция корреляции в точке $\tau = 0$ максимальна при $S_1 \equiv S_2$. Это свойство используется в технике связи для сравнения формы сигналов.

Свойство (2.6–2в) автокорреляционной функции используется для определения точного временного положения сигнала, в част-

ности, в радиолокации. Для достижения высокой точности применяются сигналы с явно выраженным узким максимумом АКФ.

2.7. Квазигармоническое представление узкополосных сигналов

Квазигармоническое представление [2-Гл5]— это представление произвольного сигнала в виде гармонического сигнала со средней частотой ω_0 , амплитуда $A(t)$ и начальная фаза $\varphi(t)$ которого изменяются во времени, то есть являются функцией времени:

$$S(t) = A(t) \cdot \cos[\omega_0 t + \varphi(t)] = A(t) \cdot \cos \psi(t). \quad (2.7-1)$$

То же представление в квадратурной форме имеет вид:

$$S(t) = A(t) \cos(\psi(t)) \cdot \cos \omega_0 t + A(t) \sin(\psi(t)) \cdot \sin \omega_0 t \quad (2.7-2)$$

Функция $\psi(t) = \omega_0 t + \varphi(t)$ является фазой (обобщённой фазой) сигнала $S(t)$. Мгновенная частота определяется как:

$$\omega(t) = \frac{d\psi}{dt}. \quad (2.7-3)$$

Квазигармоническое представление целесообразно применять для **узкополосных сигналов**, то есть сигналов, ширина спектра ΔF которых много меньше нижней частоты сигнала:

$$\Delta F \ll f_{\text{нижняя}} \quad (2.7-4)$$

Данное условие эквивалентно тому, что амплитуда и начальная фаза медленно меняются со временем по сравнению с колебаниями сигнала. Осциллограмма любого узкополосного сигнала выглядит как гармонический сигнал (синусоида), у которого меняется амплитуда и фаза, так что в отношении узкополосного сигнала понятие «квазигармонический» имеет вполне определённый практический смысл.

Для того чтобы произвольный сигнал $S(t)$ представить в квазигармонической форме (2.7-1), необходимо иметь способ

вычисления средней частоты ω_0 , а также функций $A(t)$ и $\psi(t)$. Для этого необходимо применить математический аппарат преобразований Гильберта.

Преобразование Гильберта сигнала $S(t)$ определяется как:

$$S_1(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{S(\tau)}{\tau - t} d\tau. \quad (2.7-5)$$

Сигнал $S_1(t)$ называется сигналом, сопряженным по Гильберту с сигналом $S(t)$. Имеет место обратное преобразование Гильберта:

$$S(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{S_1(\tau)}{\tau - t} d\tau. \quad (2.7-6)$$

Таким образом, $S(t)$ и $S_1(t)$ связаны парой преобразований Гильберта (2.7-5) и (2.7-6).

Используя введенные обозначения, можем определить искомые функции:

$$A(t) = \sqrt{S^2(t) + S_1^2(t)}, \quad \varphi(t) = \arccos \frac{S(t)}{A(t)},$$

$$\omega_0 = \omega_{\text{средняя}} = \lim \frac{1}{T} \int_{-T/2}^{T/2} \omega(t) \cdot dt, \quad (2.7-7a)$$

Непосредственным вычислением можно показать, что:

$$\text{если } S(t) = \sin(\omega t), \text{ то } S_1(t) = -\cos(\omega t) = \sin(\omega t - \pi / 2), \quad (2.7-8a)$$

$$\text{если } S(t) = \cos(\omega t), \text{ то } S_1(t) = \sin(\omega t) = \cos(\omega t - \pi / 2). \quad (2.7-8b)$$

Эти соотношения показывают, что преобразование Гильберта гармонического сигнала означает сдвиг его фазы на $\pi/2$. Данное свойство применяется при аппаратной реализации преобразования Гильберта в системах связи.

Если представить сигнал в виде спектрального разложения (2.4-8)

$$S(t) = \frac{1}{\pi} \int_0^{\infty} [a(\omega) \cos(\omega t) + b(\omega) \sin(\omega t)] d\omega, \quad (2.7-9a)$$

то для сопряженного сигнала сразу можно написать:

$$S_1(t) = \frac{1}{\pi} \int_0^{\infty} [a(\omega) \sin(\omega t) - b(\omega) \cos(\omega t)] d\omega \quad (2.7-9б)$$

Из пары соотношений (2.7-9а) и (2.7-9б) следует, что односторонние амплитудные спектры сигналов $S(t)$ и $S_1(t)$ равны. Для комплексного (двустороннего) спектра $\dot{S}_1(\omega)$ сигнала $S_1(t)$ можно показать, что:

$$\dot{S}_1(t) = \begin{cases} -i \dot{S}(t), \omega > 0 \\ i \dot{S}(t), \omega < 0 \end{cases} \quad (2.7-10)$$

Комплексный сигнал $\dot{Z}(t)$, соответствующий узкополосному сигналу $S(t)$, определяется аналогично тому, как мы его определили в случае гармонического сигнала выражением (2.2-3):

$$\dot{Z}(t) = S(t) + i \cdot S_1(t) = A(t) \cdot e^{i\psi(t)} = \dot{A}(t) \cdot e^{i\omega_0(t)} \quad (2.7-11)$$

Данный сигнал называется **аналитическим сигналом**, соответствующим сигналу $S(t)$. Из выражения (2.7-11) вытекает, что спектральная плотность $\dot{Z}(\omega)$ аналитического сигнала $\dot{Z}(t)$ определяется суммой:

$$\dot{Z}(\omega) = \dot{S}(\omega) + i \dot{S}_1(\omega) \quad (2.7-12)$$

Из соотношений (2.7-10) и (2.7-12) вытекает, что спектр аналитического сигнала односторонний, то есть содержит только положительные частоты, а именно:

$$\dot{Z}(\omega) = \begin{cases} 2\dot{S}(\omega), \omega > 0 \\ 0, \omega < 0 \end{cases} \quad (2.7-13)$$

3. Аналоговая модуляция гармонического несущего сигнала

Модуляция – это процесс управления каким-либо параметром несущего сигнала $S_{HEC}(t)$ в соответствии с передаваемым (управляющим, модулирующим, информационным) сигналом $E(t)$. Модуляция применяется в том случае, если непосредственная передача информационного сигнала по каналу (или линии) нецелесообразна или невозможна [1-§3.2,3.3, 2-§4.1,4.2]. В качестве несущего сигнала выбирается сигнал, который подходит для передачи по линии. В результате модуляции несущий сигнал «берет» на себя информационный сигнал и, тем самым, формируется сигнал, пригодный для передачи и вместе с тем «несущий» на себе информационный сигнал. На приемном конце выделяется изменение параметра модуляции, которое представляет собой переданный сигнал. Этот процесс называется **демодуляцией** или **детектированием**.

В настоящем разделе рассматривается модуляция гармонического несущего сигнала $S_{HEC}(t) = A \sin(\omega_{HEC}t + \varphi)$ **аналоговым информационным сигналом $E(t)$** .

Гармонический несущий сигнал имеет три параметра, которые можно использовать для модуляции – амплитуду, фазу и частоту. Частота и фаза – взаимосвязанные параметры, так что частотная и фазовая модуляции всегда сопутствуют друг другу. Они различаются по способу воздействия на фазу.

Модуляция гармонического несущего сигнала осуществляет перенос спектра передаваемого сигнала в область высоких частот. Одновременно он становится узкополосным (см. 2.7–4). Эти два обстоятельства необходимы в радиосвязи, так как дают возможность эффективного излучения радиоволн. Те же соображения позволяют осуществить **частотное (спектральное) разделение каналов** путем распределения сигналов отдельных каналов в различных интервалах частотного диапазона.

3.1. Амплитудная модуляция (АМ)

Амплитудная модуляция (АМ) – это управление амплитудой гармонического несущего сигнала $S_{HEC}(t) = A \sin(\omega_{HEC}t)$ в соответствии с управляющим сигналом $E(t)$ [1-§3.2, 2-§4.1]. Аналитическое выражение АМ сигнала в случае гармонического управляющего сигнала $E(t) = E \cos(\Omega t)$ имеет вид:

$$S(t) = A \cdot [1 + M \cos(\Omega t)] \sin(\omega_{HEC}t), \quad (3.1-1)$$

где A – амплитуда несущего сигнала, ω_{HEC} – частота несущего сигнала, Ω – частота управляющего сигнала.

Параметр M – глубина модуляции, показывающий степень изменения амплитуды несущего гармонического сигнала относительно ее среднего значения A . На графиках Рис.3.1–1 даны временные диаграммы АМ сигнала при различных значениях M . На третьем графике приведен случай «перемодуляции» ($M > 1$).

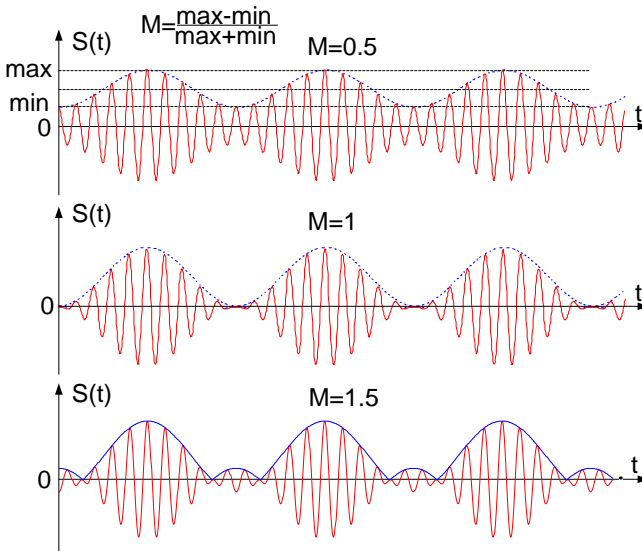


Рис. 3.1–1. Временные диаграммы АМ сигнала при различных глубинах модуляции.

Как видно из рисунка, при перемодуляции ($M > 1$) огибающая АМ сигнала, то есть изменение его амплитуды не совпадает с модулирующим сигналом. Это означает, что если демодулятор (детектор) определяет изменение амплитуды, то он выдаст искаженный сигнал, не совпадающий с модулирующим сигналом. Поэтому при модуляции необходимо обеспечить условие $M < 1$. Однако при более внимательном рассмотрении можно заметить, что в точках излома огибающей фаза сигнала совершает скачок на π . Это означает, что детектор, учитывающий фазу несущего сигнала, может правильно демодулировать сигнал даже при $M > 1$. Таким детектором является **синхронный детектор**.

Из выражения (3.1–1) легко получается разложение сигнала АМ на гармонические составляющие:

$$S(t) = A \cdot \sin(\omega_{HEC} t) + \frac{A \cdot M}{2} [\cos(\omega_{HEC} + \Omega)t + \cos(\omega_{HEC} - \Omega)t]. \quad (3.1-2)$$

Как мы видим, при модуляции гармонического несущего сигнала гармоническим сигналом получается спектр, состоящий из трех частот, как показано на Рис.3.1–2.

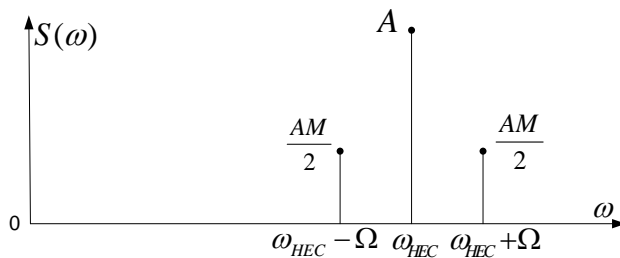


Рис. 3.1–2. Спектральная диаграмма АМ сигнала.

Если управляющий (модулирующий) сигнал $E(t)$ не гармонический, а имеет спектр $E(f)$, то аналитически можно аналогично получить, что в результате модуляции получается сигнал, состоя-

щий из несущего сигнала f_{HEC} и двух боковых полос, как показано на Рис.3.1–3.

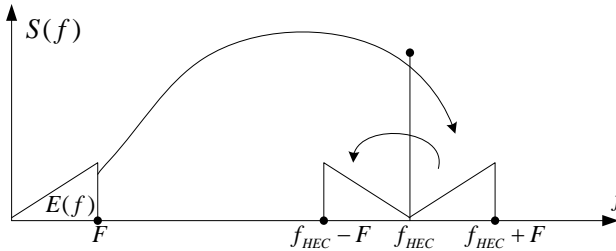


Рис. 3.1–3. Трансформация спектра при амплитудной модуляции

Легко увидеть, что спектр АМ сигнала вдвое шире спектра управляющего, то есть информационного сигнала $E(f)$, подлежащего передаче:

$$\Delta F_{AM} = 2 \cdot F. \quad (3.1-3)$$

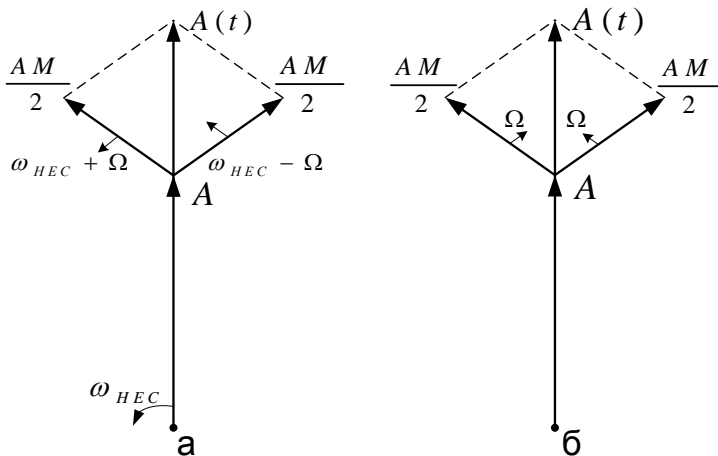


Рис.3.1-4. Векторная диаграмма АМ сигнала.

На Рис.3.1–4а приведена векторная диаграмма АМ сигнала. Сам сигнал – это вектор $A(t)$, являющийся суммой вектора

несущего сигнала A , вращающегося с частотой $\omega_{\text{НЕС}}$ и векторов боковых частот, скорости вращения которых, соответственно, $\omega_{\text{НЕС}} - \Omega$ и $\omega_{\text{НЕС}} + \Omega$. Таким образом, один боковой отстаёт от вектора несущего сигнала, а другой его опережает. Во вращающейся системе координат, связанной с вектором несущего сигнала (Рис.3.1–4б) видно, что боковые векторы вращаются относительно несущего так, что их сумма всегда направлена вдоль последнего, то есть мы имеем вектор с изменяющейся длиной, но равномерной скоростью вращения. Это и есть АМ в чистом виде. Векторная диаграмма наглядно показывает, что при всех преобразованиях АМ сигнала необходимо сохранять симметрию боковых векторов, иначе может исказиться амплитудная модуляция, а также возникнуть «паразитная» модуляция фазы.

Как видно из спектральной диаграммы АМ сигнала (Рис.3.1–2 и 3.1–3), значительная доля мощности сосредоточена на несущей частоте, не содержащей информацию. Если убрать эту частоту из спектра АМ сигнала, мы получим так называемую балансную модуляцию (БАМ), которая экономичнее обычной АМ (Рис.3.1–5а). Можно пойти ещё дальше, исключив одну из боковых полос (Рис.3.1–5б). Мы получим модуляцию с одной боковой полосой – ОБП (SSB). Однако тут возникают проблемы с демодуляцией. Поэтому в спектре сохраняют «немного» несущего колебания (Рис.3.1–5в), так называемый пилот-сигнал. ОБП сигнал позволяет уменьшить мощность передатчика почти вдвое по сравнению с БАМ и, что очень важно, позволяет вдвое сузить частотную полосу передачи. Как модуляция, так и демодуляция каждой разновидности амплитудной модуляции имеет свои особенности и требует своей схемной реализации.

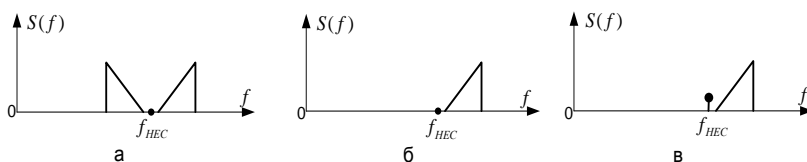


Рис.3.1–5. Спектры сигналов БАМ, ОБП и ОБП с пилот-сигналом.

3.2. Угловая модуляция (ФМ и ЧМ)

При использовании в качестве несущего сигнала гармонический сигнал в процессе модуляции можно изменять не амплитуду сигнала, а его фазу [1-§3.3, 2-§4.2]. Такая модуляция называется угловой модуляцией (УМ). Понятие угловой модуляции включает фазовую модуляцию (ФМ) и частотную модуляцию (ЧМ).

Сигнал с модулированной фазой имеет вид:

$$S_{HEC}(t) = A \cos(\omega_{HEC}t + \varphi(t)) = A \cos \Phi(t), \quad (3.2-1)$$

где $\Phi(t) = \omega_{HEC}t + \varphi(t)$ – обобщенная фаза сигнала.

Мгновенная частота сигнала, в общем случае определяется как:

$$\omega(t) = \frac{d\Phi(t)}{dt} = \omega_{HEC} + \frac{d\varphi(t)}{dt}. \quad (3.2-2)$$

Для гармонического сигнала начальная фаза $\varphi(t) = \varphi_0$ и частота ω_{HEC} постоянны, а обобщенная фаза является линейной функцией времени: $\Phi(t) = \omega_{HEC}t + \varphi_0$.

Мы видим из (3.2-2), что мгновенная частота и фаза взаимосвязаны, так что модуляция фазы означает также модуляцию частоты и наоборот. Поэтому частотная и фазовая модуляции объединены под названием угловой модуляции (УМ).

При фазовой модуляции (ФМ) управляющий сигнал $e(t)$ воздействует непосредственно на начальную фазу сигнала:

$$S(t) = A \cos(\omega_{HEC}t + \varphi(t)) = A \cos(\omega_{HEC}t + k \cdot e(t)), \quad (3.2-3)$$

При частотной модуляции (ЧМ) изменение начальной фазы сигнала пропорционально интегралу управляющего сигнала:

$$S(t) = A \cos \left(\omega_{HEC}t + k \cdot \int_0^t e(\tau) d\tau \right), \quad (3.2-4)$$

Мгновенная частота в этом случае $\omega = \frac{d\Phi(t)}{dt} = \omega_{HEC} + k \cdot e(t)$, поэтому такая модуляция называется частотной. Коэффициент

пропорциональности k в обоих случаях называется коэффициентом модуляции, так как показывает степень влияния управляющего сигнала соответственно на фазу или частоту несущего сигнала.

Индекс модуляции m – это максимальное изменение фазы при модуляции.

Девияция частоты $\Delta\omega$ (или $\Delta f = \Delta\omega / 2\pi$) – это максимальное изменение частоты при модуляции.

Пусть управляющий сигнал гармонический с частотой $\Omega=2\pi F$:

$$e(t) = E \cos(\Omega \cdot t). \quad (3.2-5)$$

Выражение (3.2-3) для ФМ в этом случае сигнала принимает вид:

$$S(t) = A \cos(\omega_{HEC}t + kE \cos(\Omega t)), \quad (3.2-6)$$

откуда видно, что индекс модуляции при ФМ $m=kE$. Используя (3.2-2), легко получить, что девияция частоты при ФМ $\delta\omega = kE\Omega$.

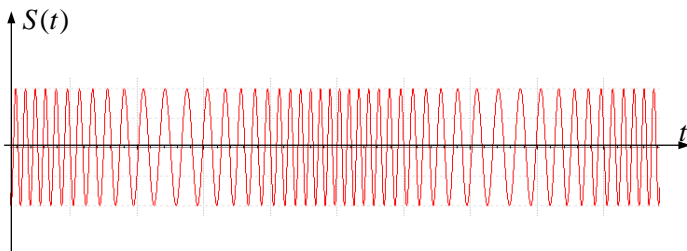
Выражение (3.2-4) для ЧМ в случае управляющего сигнала (3.2-5) принимает вид:

$$S(t) = A \cos\left(\omega_{HEC}t + \frac{kE}{\Omega} \cdot \sin \Omega t\right), \quad (3.2-7)$$

откуда получается, индекс модуляции при ЧМ $m = kE / \Omega$, а девияция частоты $\Delta\omega = k \cdot E$

Мы видим, что индекс модуляции и девияция частоты связаны соотношением $\Delta\omega = m\Omega$.

На Рис.3.2-1 приведена временная диаграмма ЧМ сигнала. Как видно из рисунка, период, а, следовательно, и частота сигнала меняется со временем.



**Рис.3.2–1. Временная диаграмма ЧМ сигнала с $f_{\text{НЕС}} = 25\text{кГц}$,
 $F = 1\text{кГц}$, $\Delta f = 10\text{кГц}$.**

Спектр сигнала УМ неограничен, однако его ширину можно приблизительно принять равной

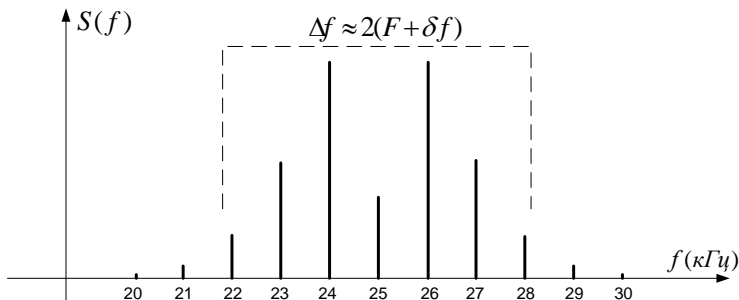
$$\Delta F_{\text{УМ}} \approx 2(F + \Delta f), \quad (3.2-8)$$

так как за пределами этой полосы спектр резко убывает. Здесь Δf – девиация частоты, а F – ширина полосы модулирующего сигнала.

При малых индексах модуляции

$$\Delta F_{\text{УМ}} \approx 2F, \quad (3.2-9)$$

то есть ширина спектра совпадает с шириной спектра АМ сигнала.



**Рис.3.2–2. Спектральная диаграмма ЧМ сигнала с $f_{\text{НЕС}} = 25\text{кГц}$,
частотой модулирующего сигнала $F = 1\text{кГц}$, девиацией
частоты $\Delta f = 2\text{кГц}$.**

На Рис.3.2–2 приведена спектральная диаграмма ЧМ сигнала. Из рисунка видно, что, несмотря на то, что и несущий и модулирующий сигналы гармонические, спектр состоит из большого числа гармоник, расстояние между которыми равно частоте модулирующего сигнала F . Прямоугольником выделена частотная область, содержащая основную мощность сигнала, то есть определяющая ширину спектра.

В случае малого индекса модуляции $m \ll 1$, спектр сигнала УМ непосредственно вычисляется. Если вспомним, что при $x \ll 1$ имеем $\sin x \approx x, \cos x \approx 1$, то получим:

$$S(t) = A \cos(\omega_{HEC}t + m \cdot \cos \Omega t) = A \cos \omega_{HEC}t + \frac{mA}{2} \cos(\omega_{HEC} + \Omega)t - \frac{mA}{2} \cos(\omega_{HEC} - \Omega)t. \quad (3.2-10)$$

Спектральная диаграмма сигнала для этого случая представлена на Рис.3.2–3. Знак минус у нижней боковой полосы означает противоположную фазу в отличие от аналогичного спектра АМ сигнала (формула (3.1–2) и Рис.3.1–2). Это учтено и на векторной диаграмме (Рис.3.2–4). Благодаря такой фазировке, суммарный вектор меняет только направление (фазу), но не амплитуду (если учесть, что $\delta\varphi \ll 1$).

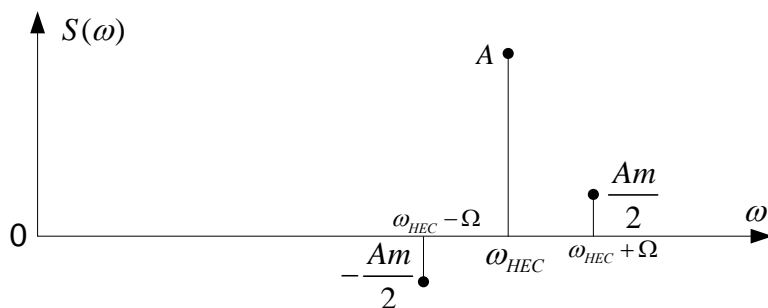


Рис.3.2–3. Спектральная диаграмма УМ в случае малого индекса модуляции.

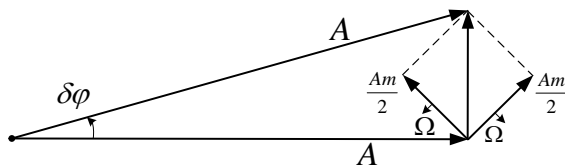


Рис.3.2–4. Векторная диаграмма УМ в случае малого индекса модуляции.

Как следует из (3.2–8), спектр сигнала УМ в общем случае шире, чем у сигнала АМ, что является нежелательным с точки зрения эффективности использования частотного диапазона. Однако информация в сигнале УМ менее подвержена влиянию помех, вследствие чего улучшается верность передачи. Это – пример того, как надлежащее расширение полосы сигнала повышает качество системы связи. Поэтому высококачественное вещание в FM и УКВ диапазонах ведется с частотной модуляцией. Дальнейшее улучшение качества вещания путем расширения спектра связано с переходом на цифровое вещание.

3.3. Частотное уплотнение линий связи.

Частотное разделение каналов

Уплотнение линии связи означает использование **одной линии связи** для организации **более чем одного канала связи**, то есть построение системы **многоканальной связи**. Многоканальная связь обеспечивает передачу информации между более чем одним источником и получателем. Естественно, при уплотнении линии необходимо исключить взаимное влияние между каналами, то есть обеспечить **разделение каналов**. Существуют различные схемы построения многоканальных систем, в зависимости от практической задачи. Соответственно, существует много способов разделения каналов. Одна из схем – это радиовещание, когда в эфире работает одновременно несколько радиопередатчиков в непересекающихся частотных полосах. При этом получатель (радиослушатель) сам выбирает, которым каналом воспользоваться.

Тут каналы различаются по используемой частотной полосе, то есть применяется **частотное разделение каналов**[1-§9.2].

На Рис.3.3–1 приведена достаточно общая схема многоканальной системы связи с частотным разделением каналов. Здесь модуляторы служат для смещения частотной полосы передаваемых сигналов в различные частотные области путем соответствующего выбора несущей частоты. Затем все сигналы поступают в линию связи. В линии сигналы складываются. Этот суммарный сигнал называется **групповым сигналом**. Объединение канальных сигналов в один групповой сигнал называется **мультиплексированием**. В данном случае мультиплексирование есть наложение этих сигналов. Благодаря тому, что частотные полосы каналов **не пересекаются**, сигналы отдельных каналов **взаимно ортогональны**, что является достаточным условием возможности их разделения на приемном конце. Это разделение в случае частотного уплотнения обеспечивается частотной фильтрацией. На приемном конце полосовые фильтры отфильтровывают сигнал, относящийся к данному каналу. После демодуляции сигнал предоставляется соответствующему получателю.

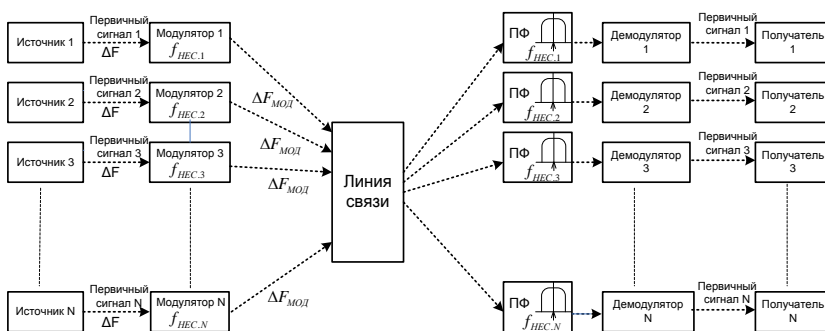


Рис.3.3–1. Схема многоканальной системы связи с частотным разделением каналов.

На Рис.3.3–2 дана спектральная диаграмма, описывающая работу данной многоканальной системы в предположении, что используется АМ. Предполагается, что спектры всех первичных сигналов занимают одну и ту же частотную полосу. Для того чтобы

полосы сигналов после модуляции не перекрывались, промежуток между соседними несущими частотами Δf должен быть не меньше ширины полосы модулированного сигнала $\Delta F_{\text{МОД}}$. На практике $\Delta f > \Delta F_{\text{МОД}}$, так как между передаваемыми полосами оставляется так называемый **защитный промежуток**. Защитный промежуток необходим, чтобы не произошло наложение каналов из-за нестабильности несущей частоты или не идеальности полосового фильтра ПФ. В результате наложения в канал может попасть сигнал от соседних каналов, что является помехой для данного канала.

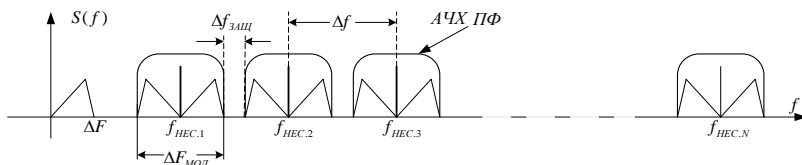


Рис.3.3–2. Спектральная диаграмма работы многоканальной системы с АМ.

Такая помеха, свойственная многоканальным системам, называется **переходной или взаимной помехой**. Переходная помеха в системах с частотным разделением может также возникать из-за нелинейностей в различных звеньях системы связи. Это объясняется тем, что в результате нелинейности в сигнале образуются новые гармоники, которые могут оказаться в полосе другого канала.

4. Цифровое представление аналоговых сигналов

4.1. Амплитудно-Импульсная Модуляция (АИМ)

Одним из методов передачи аналоговых сигналов является также амплитудно-импульсная модуляция (АИМ) [1-§2.4, 2-§5.1, 3-§3.1]. При амплитудно-импульсной модуляции несущим сигналом является импульсная последовательность с периодом T , и в процессе модуляции изменяется высота (амплитуда) импульсов в соответствии с передаваемым (управляющим, информационным) сигналом $E(t)$ (Рис.4.1–1).

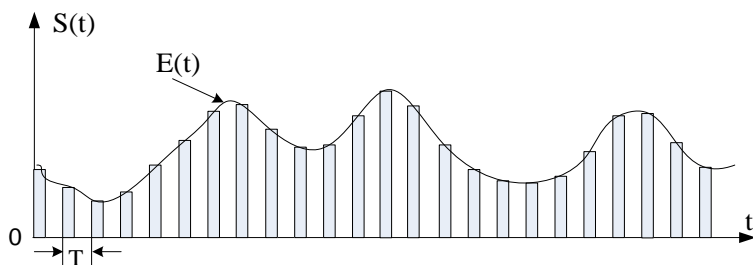


Рис.4.1–1. Временная диаграмма АИМ сигнала.

Примем, что спектр $E(f)$ управляющего сигнала ограничен частотой F (Рис.4.1–2). Спектр периодической последовательности импульсов (без модуляции амплитуды) был рассмотрен в разделе 2.3 и изображен на Рис.2.3–1. В более упрощенном виде этот спектр, то есть спектр немодулированного несущего колебания, представлен ниже на Рис.4.1–2.

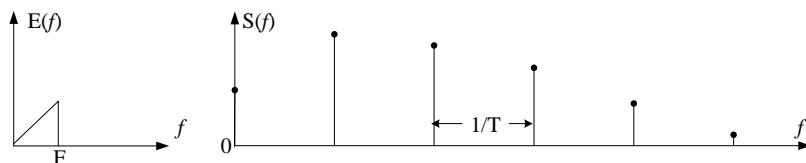


Рис.4.1–2. Спектры модулирующего $E(f)$ и несущего $S(f)$ сигналов.

При модуляции импульсной последовательности по амплитуде каждая гармоника этой последовательности, в том числе и постоянная составляющая, подвергается амплитудной модуляции, в результате чего около каждой гармоники образуются две боковые полосы (Рис.4.1–3). Постоянная составляющая также модулирована, поэтому в спектре АИМ сигнала оказывается также спектр модулирующего сигнала, как показано на Рис.4.1–3.

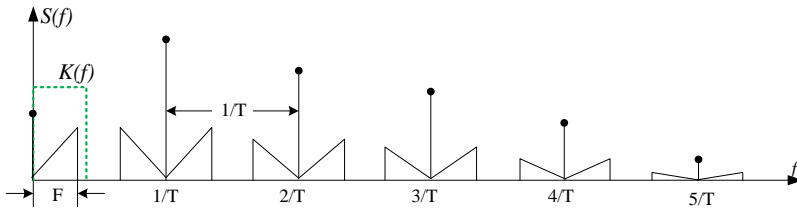


Рис.4.1–3. Спектр несущего сигнала, модулированного управляющим сигналом $E(t)$ в случае, когда удовлетворяется условие $T > 1/2F$.

Из спектральной диаграммы видно, что выделение управляющего (информационного) сигнала сводится к НЧ фильтрации. Там же изображена амплитудно-частотная характеристика (АЧХ) $K(f)$ идеального фильтра нижних частот (ФНЧ), предназначенного для выделения управляющего сигнала на приемном конце (демодуляции). Из рисунка видно, что выделение управляющего (информационного) сигнала путём НЧ фильтрации возможно, если удовлетворяется условие $T \leq 1/2F$.

Если условие $T \leq 1/2F$ не выполняется, то спектр левой боковой полосы первой гармоники накладывается на спектр информационного сигнала (Рис.4.1–4). Чистая фильтрация (демодуляция) в этом случае неосуществима. Мы или потеряем часть полезного сигнала (Рис.4.1–4а), или пропустим помеху (Рис.4.1–4б).

Передача сигнала методом АИМ означает, что мы **передаём** не весь сигнал, а его значения в дискретные моменты времени $0, T, 2T, \dots, nT, \dots$. Значения сигнала в эти дискретные моменты

времени $S(0), S(T), S(2T), \dots, S(nT), \dots$ называются **отсчетами сигнала**, а их совокупность – **выборкой**.

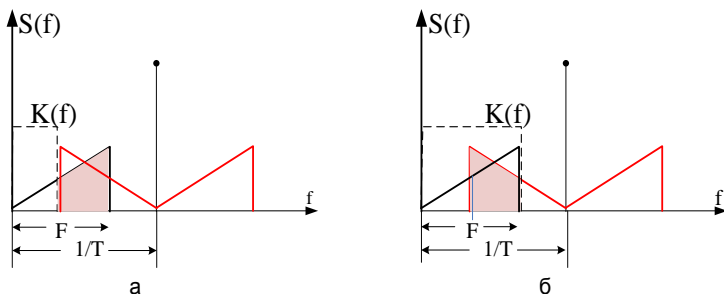


Рис.4.1–4. Спектр несущего сигнала, модулированного управляющим сигналом $E(t)$ в случае, когда $T > 1/2F$.

Представление сигнала в виде последовательности отсчётов называется дискретизацией сигнала. Временной интервал Δt называется **шагом дискретизации**, который в случае АИМ равен периоду T несущей последовательности. Соответственно, обратная величина называется частотой дискретизации $F_d = 1/T = 1/\Delta t$.

Рассматривая АИМ, мы пришли к выводу, что дискретизированный сигнал может быть восстановлен, если отсчеты брать достаточно часто – с шагом $T \leq 1/2F$. Другими словами, частота дискретизации должна быть $F_d \geq 2F$. При выполнении этого условия дискретизация является обратимым преобразованием, а значит, не приводит к потерям информации.

Так как дискретизация предполагает ограниченность спектра передаваемого сигнала, то спектр реальных сигналов перед дискретизацией необходимо ограничить величиной $F = 1/2\Delta t$. Фильтр, выполняющий данную функцию, называется **антиалийзинговым фильтром** (фильтром против наложения). В частности, в телефонии частота дискретизации выбрана равной 8000 отсчетов

в секунду. Это означает, что полоса сигнала, передаваемого по телефонному каналу, должна быть ограничена частотой 4 кГц.

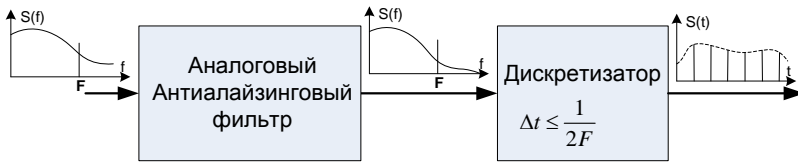


Рис. 4.1–5. Схема дискретизации аналогового сигнала

Таким образом, схема дискретизации аналогового сигнала имеет вид, данный на Рис.4.1–5. Частота дискретизации в данной схеме выбирается близкой к оптимальной и требуется хороший аналоговый фильтр, который **трудно реализуем**. Не идеальность фильтра приводит к тому, что после фильтрации в спектре остаётся «хвост», который приводит к наложению спектров при дискретизации, что проявляется как помеха (шум). Есть возможность увеличить частоту дискретизации, то есть осуществить так называемую «выборку с запасом». Это уменьшит влияние «хвостов» и, тем самым, облегчит задачу аналоговой фильтрации. Однако это означает расширение полосы передачи, что снижает эффективность использования канала связи. Ниже, при рассмотрении цифровой передачи, будет рассмотрен принцип радикального решения данной проблемы.

4.2. Теорема Котельникова

Возможность дискретизации сигнала, допускающей последующее его восстановление, является содержанием одного из важнейших результатов теории сигналов, сформулированного в известной теореме Котельникова (Найквиста) [1-§2.4, 2-§3.6, 3-§3.1]. Наиболее общая формулировка данной теоремы следующая:

Теорема Котельникова: Произвольный сигнал $S(t)$, ширина спектра которого ограничена и равна F , может быть полностью

восстановлен, если известны его отсчеты, взятые через одинаковые промежутки времени, равные $\Delta t \leq 1/2F$.

Обсуждение теоремы Котельникова в данной формулировке и тем более ее практического применения в таком виде достаточно сложная задача. Мы рассмотрим ее для случая «низкочастотного» сигнала, для которого спектр начинается почти с нуля. Таким является большинство первичных сигналов, которые подвергаются дискретизации в системах связи, например, звуковой сигнал (в том числе, телефонный). Теорема Котельникова в применении к таким сигналам имеет следующую формулировку:

Теорема Котельникова: Произвольный сигнал $S(t)$, спектр которого ограничен сверху частотой $F_{\text{ВЕРХ}} = F$, может быть полностью восстановлен, если известны его отсчеты, взятые через одинаковые промежутки времени, равные $\Delta t \leq 1/2F$. Сигнал выражается через свои отсчёты $S(0), S(\Delta t), S(2\Delta t), S(3\Delta t), \dots, S(n\Delta t)$ рядом Котельникова:

$$S(t) = \sum_{n=-\infty}^{+\infty} S(n\Delta t) \frac{\sin 2\pi F(t - n\Delta t)}{2\pi F(t - n\Delta t)}. \quad (4.2-1)$$

АИМ является одним из способов передачи отсчетов сигнала, при котором каждый отсчёт представляется амплитудой соответствующего импульса. В теореме Котельникова предполагается, что отсчеты передаются посредством δ -импульсов $S(n\Delta t) \cdot \delta(t - n\Delta t)$ с соответствующими временными сдвигами и площадью, пропорциональной величине отсчета. При этом на выходе восстанавливающего ФНЧ мы получаем последовательность отсчетных функций (Рис.4.2–1) с соответствующими временными сдвигами и коэффициентами. Таким образом, в соответствии с выражением (4.2–1), выходной сигнал будет равен первоначальному сигналу $S(t)$.

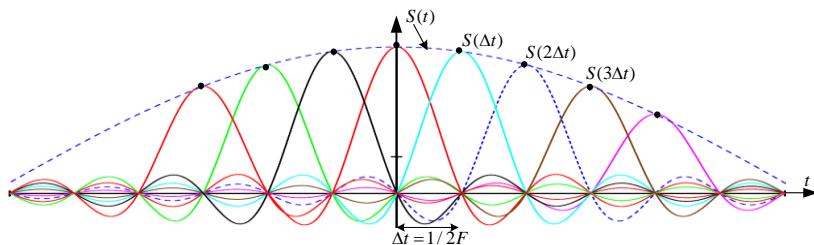


Рис. 4.2–1. Суммирование отсчётных функций на выходе восстанавливающего ФНЧ.

Из формулы (4.2–1) следует, что для восстановления сигнала в каждый момент времени t необходимо знать все отсчеты этого сигнала, определённые на всей временной оси. Поскольку теорема Котельникова относится к сигналам с конечным спектром, то, согласно основному свойству преобразования Фурье (см. раздел 2.4), длительность сигнала бесконечна. Отсюда естественно ожидать, что число отсчетов, требуемых для восстановления сигнала, бесконечно, что делает всю процедуру непригодной с точки зрения техники связи. Обойти эту сложность можно, используя понятие эффективной длительности сигнала (см. раздел 2.4) и, тем самым, использовать ограниченное число отсчетов. При этом возникает ошибка восстановления, связанная с тем, что часть отсчётов не учитывается.

Есть класс сигналов, для которых возможно полное восстановление. Это так называемые «сигналы с конечной базой». **Базой сигнала** N называется число ненулевых отсчётов сигнала. База сигнала равна

$$N = 2F\Delta T, \quad (4.2-2)$$

где F – ширина спектра сигнала, а ΔT – длительность интервала, вне которого все **отсчеты сигнала равны нулю**.

Минимальной базой $M=1$ обладает сигнал с равномерным ограниченным спектром, так называемый идеальный низкочастотный сигнал (Рис.2.4–2)

$$S(t) = 2F \cdot \frac{\sin 2\pi Ft}{2\pi Ft}. \quad (4.2-3)$$

Этот сигнал можно восстановить по одному отсчету. Это становится ясным, если вспомнить, что данный сигнал получается как импульсная характеристика идеального ФНЧ (Рис.2.5-3). Таким образом, если передаваемый сигнал имеет вид (4.2-3), то можно передать его единственный отсчёт. На приёмном конце, подавая этот отсчёт в виде δ -импульса идеальный ФНЧ с полосой равной F , мы получим на выходе восстановленный сигнал (4.2-3). Это же следует из ряда Котельникова (4.2-1).

Базу произвольного сигнала можно **приближенно** определить, используя понятие эффективной длительности и эффективной ширины спектра:

$$N = 2F_{\text{эфф}} \Delta T_{\text{эфф}}. \quad (4.2-4)$$

Передача непрерывного сигнала методом дискретизации предполагает передачу отсчетов данного сигнала каким-либо способом и дальнейшего его восстановления в соответствии с выражением (4.2-1). Передача с помощью коротких импульсов, как показано на Рис.4.2-1, является неэффективной, так как, очевидно, требует частотной полосы, существенно превышающей полосу передаваемого сигнала. Если передавать отсчеты посредством последовательности отсчетных функций (Рис. 4.2-1), то полоса передачи совпадет с полосой сигнала. Однако такая передача тождественна непосредственной передаче первоначального сигнала. Практически пригодным способом передачи отсчетов аналогового сигнала является импульсно-кодовая модуляция, рассматриваемая ниже, то есть цифровая передача.

4.3. Квантование

В разделе 4.1 была рассмотрена дискретизация аналогового сигнала, то есть его преобразования в последовательность отсчетов. Было показано, что такое преобразование возможно осу-

ществить без потери информации, то есть существует способ **его восстановления** на приемном конце. Такое преобразование является дискретизацией по времени, то есть непрерывная временная ось заменяется последовательностью моментов времени. В данном разделе рассматривается **дискретизация по значениям, которая называется квантованием сигнала**[3-§3.2]. Такое преобразование означает, что непрерывная область значений сигнала заменяется дискретным множеством «разрешенных» значений, которые называются «уровнями квантования». Процедура квантования в принципе применима к первоначальному сигналу, однако мы будем рассматривать только квантование **отсчетов дискретизированного сигнала**.

На Рис.4.3–1 изображен процесс квантования. Он заключается в том, что значение каждого отсчета заменяется ближайшим к нему уровнем квантования. На рисунке изображено равномерное квантование, когда шаг квантования Δ один и тот же между всеми соседними уровнями квантования. В технике связи, однако, чаще применяется неравномерное квантование.

Как видно из рисунка, процесс квантования сопряжен с искажением отсчетов сигнала. Если дискретизация сигнала в принципе возможна без потери информации, то квантование всегда необратимо искажает сигнал, то есть приводит к потерям информации.

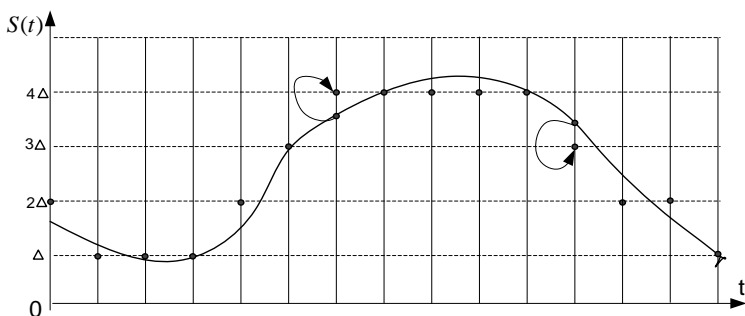


Рис 4.3–1. Процесс квантования.

В результате квантования отсчётов аналогового сигнала мы получаем последовательность уровней квантования (Рис. 4.3–2) – дискретную последовательность, соответствующую последовательности отсчётов первоначального сигнала.

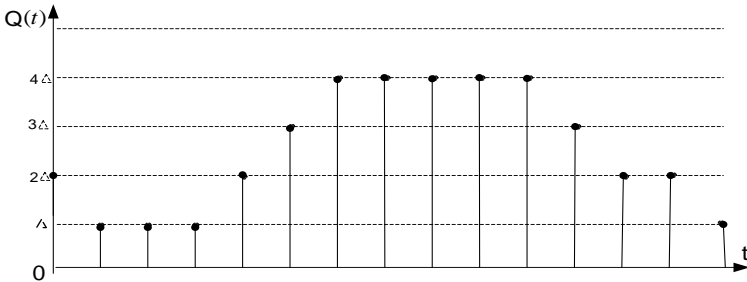


Рис 4.3–2. Квантованный сигнал

Разница между квантованным значением отсчёта и его истинным значением является «шумом» квантования. Его роль такая же, как у аддитивного шума, действующего в канале. На Рис 4.3–3 изображен шум квантования для сигнала, приведённого на Рис 4.3-1.

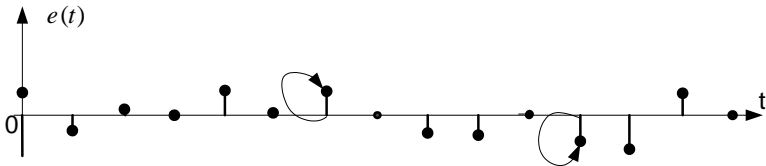


Рис 4.3–3. Шум квантования.

Мощность шума квантования, рассчитанная в предположении, что он имеет равномерное распределение на интервале $[-\Delta / 2, \Delta / 2]$, равна

$$P_{ш.кв} = \frac{\Delta^2}{12}. \quad (4.3-1)$$

Здесь следует уточнить, что в реальных аналого-цифровых преобразователях (АЦП) значение отсчета заменяется не на

ближайший уровень квантования, а на ближайший **снизу** уровень. В этом случае шум распределен на интервале $[0, \Delta]$. В этом случае квантование в среднем приводит к сдвигу сигнала на величину $\Delta/2$. Мощность же шума квантования, как и в предыдущем случае, описывается выражением(4.3–1).

Выбор шага квантования, а, следовательно, и числа уровней квантования зависит от точности, с которой требуется воспроизвести сигнал на приемном конце. Например, для стандартного телефонного сигнала число уровней квантования принято $N=256$. Такой выбор числа уровней квантования связан с тем, что для стандартного телефонного канала динамический диапазон должен быть не хуже 40–50 дБ. Действительно, отношение максимального значения к минимальному значению такого квантованного сигнала получается $P_{MAX} / P_{MIN} = 256$. С этой точки зрения динамический диапазон лежит в допустимых пределах $D = 20 \log 256 \approx 48$ дБ. То же значение получается, если рассчитать отношение средней мощности сигнала к мощности шума квантования:

$$D = 10 \log \frac{A^2 / 12}{\Delta^2 / 12} \approx 48 \text{ дБ}. \quad (4.3-2)$$

Формула (4.3–2) получена в предположении, что мощность сигнала распределена равномерно во всем диапазоне A его изменения. Для реального сигнала это не так, однако, точный расчет дает результат, также лежащий в допустимых пределах.

Как было отмечено, квантование представляет собой преобразование, которое безвозвратно искажает передаваемый аналоговый сигнал. Оправдать применение квантования можно тем, что аналоговый сигнал всегда в какой-то мере искажается в системе связи, на него всегда накладывается шум, и качество аналоговой связи определяется степенью искажения сигнала, обусловленного этими факторами (см. Введение, (1.11)). Допустимая степень этого искажения задается техническими условиями на систему связи.

Поэтому, если шаг квантования выбрать достаточно малым, так чтобы искажения, сопряженные с квантованием, не превышали допустимую норму, то вреда от квантования не будет. Однако «безвредность» не дает основания для использования лишнего, достаточно сложного преобразования, если оно бесполезно. Остается вопрос: какие преимущества дает применение квантования при передаче отсчетов аналогового сигнала. Ответы на эти вопросы будут даны в двух следующих параграфах.

4.4. Импульсно-Кодовая Модуляция. Цифровой сигнал. Дискретизация с запасом

Квантование открывает возможность цифровой передачи отсчетов дискретизированного сигнала. А именно: уровни квантования можно пронумеровать и передавать номер уровня квантования, соответствующий данному отсчету. Номер обычно выражается в двоичной системе счисления, и таким образом мы приходим к двоичной цифровой передаче сигналов[3-§3.2].

В предыдущем примере (Рис 4.2–2) последовательность уровней квантования, соответствующая передаваемому сигналу, имеет вид: 11123444443221. Поскольку в нашем случае имеем 6 уровней квантования, то для того, чтобы представить их номера в виде двоичных кодов, длину кода достаточно взять $m=3$. Коды, соответствующие номерам уровней квантования, будут $1=1002=0103=1104=0015=1016=011$. В результате получаем последовательность кодовых групп для нашего сигнала, которая изображена на Рис.4.4–1. Представление аналогового сигнала в виде последовательности кодовых групп называется **импульсно-кодовой модуляцией (ИКМ)**, а такой сигнал называется **цифровым сигналом**. Применение ИКМ сводит передачу аналогового сигнала к передаче двоичной последовательности.

Двоичная последовательность в цифровом канале неоднородна (как показано на Рис.4.4–1 разными цветами) и представляет собой последовательность кодовых групп. Поэтому на приёмном

конец возникает задача правильной группировки принятых битов. Это так называемая задача «синхронизации по кодовым группам».

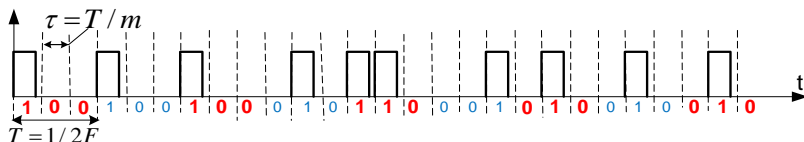


Рис. 4.4–1. Цифровая последовательность, соответствующая последовательности квантованных отсчетов, изображённой на Рис.4.3–2

Синхронизация по кодовым группам осуществляется путем ввода в структуру цифровой последовательности специального сигнала – синхросигнала (СС). Для этого применяются различные методы, аналогичные тому, как в обычном тексте между словами вставляется «пробел». Совокупность устройств, обеспечивающих передачу и прием СС, называется каналом синхронизации (КС).

Если дополнить схему Рис.4.1–5 аналого-цифровым преобразователем, то получим схему преобразования аналогового сигнала в цифровую форму (оцифровки) как показано на Рис.4.4–2 [3-§3.2]. АЦП преобразует отсчёты дискретизированного сигнала в двоичный код определённой длины m , равной разрядности АЦП. Число уровней квантования при этом, очевидно, равно $N = 2^m$. Тем самым, АЦП осуществляет одновременно и квантование и оцифровку. Например, для **стандартного телефонного канала** число уровней квантования принято $N=256$, и передача сигнала ведется двоичными блоками (циклами) из $m=8$ символов – **байтами**.

Частотная полоса телефонного **сигнала** лежит в пределах от 0,3кГц до 3,4 кГц. Полоса стандартного телефонного **канала** выбрана равной 4кГц, так что частота дискретизации в соответствии с теоремой Котельникова, равна 8кГц. Учитывая, что каж-

дый отсчёт передается восьмью битами, получим, что скорость передачи цифрового телефонного канала равна 64 кбит/с.

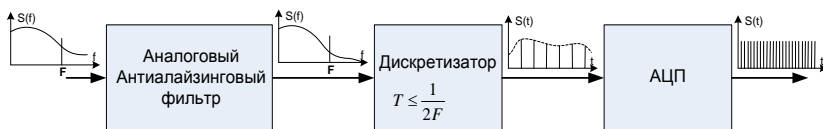


Рис. 4.4–2. *Схема преобразования аналогового сигнала в цифровую форму (оцифровки).*

Как уже было отмечено, несмотря на применение антиалийзингового фильтра, некоторое наложение все-таки остается, так как идеальный ФНЧ нереализуем, и в спектре сигнала, подаваемого на оцифровку, остаются частоты выше разрешенной $F = 1/2T$. Реализация достаточно хороших антиалийзинговых фильтров является серьезной технической трудностью в цифровой связи. Радикальным решением данной проблемы является **выборка с запасом** (дискретизация с запасом) и последующая цифровая обработка полученного цифрового сигнала (Рис.4.4–3).

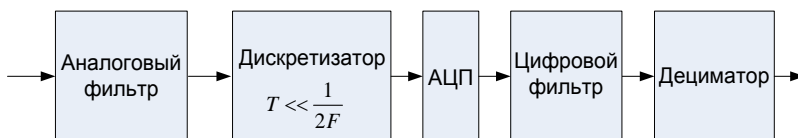


Рис. 4.4–3. *Схема оцифровки аналогового сигнала методом «выборки с запасом».*

При выборке с запасом дискретизация проводится с частотой в несколько раз выше оптимальной (для определенности, в 3 раза, то есть $F_d = 3 \cdot 2F$). При этом, аналоговый фильтр легко реализуем, так как требования к нему ниже. Затем сигнал оцифровывается, и мы получаем цифровой сигнал с полосой, равной $3F$. Далее полоса цифрового сигнала ограничивается значением F с помощью цифрового ФНЧ. Цифровые фильтры легче реализуются и дают лучшее качество фильтрации и избавления от наложения. После

фильтрации мы получаем цифровой сигнал, спектр которого ограничен нужным пределом, однако частота дискретизации полученного сигнала остается в три раза выше частоты, оптимальной для данного сигнала. Это означает, что большая часть отсчетов (две-трети) лишняя. Лишние отсчеты удаляются путем «децимации». Например, если мы взяли частоту дискретизации в 3 раза больше, то при децимации сохраняем каждую третью кодовую группу, а остальные удаляем.

4.5. Временное уплотнение линии (Временное разделение каналов)

4.5.1. Временное уплотнение сигналами АИМ

Дискретизация открывает новую возможность уплотнения линии [1-§9.2, 3-§4.6], а именно: методом **временного** разделения каналов (ВРК). Временное уплотнение использует тот факт, что при передаче дискретизированного сигнала линия связи используется в течение длительности импульсов, передающих отсчеты (при АИМ). В промежутках между импульсами линия свободна и может быть использована для передачи импульсов последовательности отсчетов другого сигнала (другого источника). Если при частотном разделении мы распределяли передаваемые сигналы по частотной оси, то при **временном** разделении мы распределяем сигналы по временной оси, сдвигая их друг относительно друга во времени.

На Рис.4.5–1 принцип временного уплотнения проиллюстрирован на примере двухканальной системы. Дискретизируя два информационных сигнала $E_1(t)$ и $E_2(t)$, мы получаем два дискретных сигнала, то есть две последовательности $S_1(t)$ и $S_2(t)$. Эти последовательности сдвигаются друг относительно друга на половину шага дискретизации T . Длительность импульсов выбирается меньше половины шага дискретизации так, что, объединяя эти две последовательности, мы получаем последовательность непере-

секающихся импульсов. Эта последовательность содержит оба сигнала и называется групповым (дискретным) сигналом $S_{ГР}(t)$.

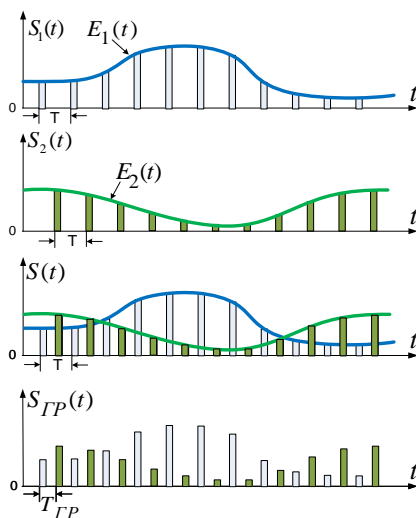


Рис.4.5–1. Принцип разделения двух дискретных потоков во времени

Объединение канальных сигналов в групповой сигнал, то есть **мультиплексирование** в данном случае означает установление очередности передачи импульсов отдельных каналов в течение каждого тактового интервала T . Так мы обеспечиваем передачу сигналов отдельных каналов во взаимно-непересекающихся временных интервалах. Такое мультиплексирование, как и в случае частотного разделения, обеспечивает взаимную ортогональность сигналов, что дает возможность их разделения на приёмном конце. На приемном конце соответственно производится **демультиплексирование**, то есть направление каждого принятого импульса соответствующему получателю.

Принцип временного объединения каналов удобно пояснить с помощью синхронно вращающихся распределителей на передающей и приемной стороне (Рис.4.5–2).

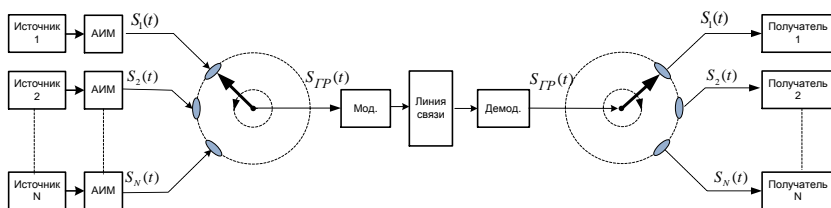


Рис.4.5–2. Упрощенная схема многоканальной системы передачи аналоговых сигналов с временным разделением.

Демультимплексирование в данном случае предполагает синхронизацию момента прихода импульса данного канала с моментом переключения к соответствующему получателю. Обычно, синхронизация устанавливается в начале передачи и поддерживается и перепроверяется в течение всей передачи. Для этого существуют специальные алгоритмы в зависимости от принятого способа импульсной модуляции, способа временного объединения каналов, структуры сигналов синхронизации и т.д.

В примере, рассмотренном на **Рис.4.5–1** для передачи каждого импульса группового сигнала отводится интервал, равный половине шага дискретизации исходного сигнала T . Соответственно, при объединении N каналов максимально возможная длительность импульса группового сигнала, будет $T_{gp} = T/N$. Исходя из этого, учитывая свойства преобразования Фурье, можно сделать вывод, что частотная полоса, требуемая для передачи группового сигнала, пропорциональна числу каналов.

Используя результаты раздела 4.2, можно показать, что теоретически возможная минимальная полоса, требуемая для передачи группового сигнала равна $I_{gp} = N \cdot F_c / 2$. Такая полоса соответствует идеальному случаю использования в качестве импульсов отсчётных функций, что не реализуемо. Использование прямоугольных (или иных) импульсов, ограничивая их спектр, приведет к их расползанию и, соответственно, к **межсимвольной интерференции** [3-§4.6] в групповом канале, то есть проникновению

расплывшегося импульса во временной интервал «соседнего» канала (Рис.4.5–2).

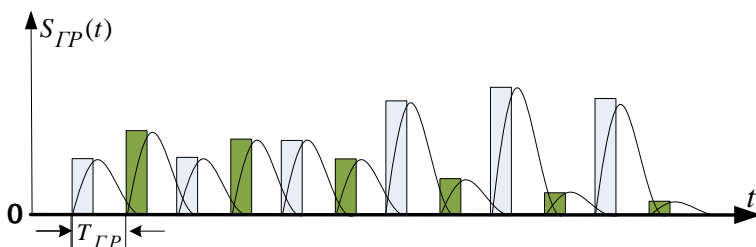


Рис.4.5–2. МСП вследствие расплывания импульсов.

Такое проникновение означает **переходную (взаимную) помеху**, то есть мешающее влияние каналов друг на друга. В случае АИМ борьба с подобными проблемами не приводит к существенным результатам, то есть низкое качество разделения каналов не позволяет строить эффективные многоканальные системы. Поэтому временное уплотнение каналов сигналами АИМ не нашло применения.

4.5.2. Временное разделение в системах с ИКМ

В современной технике связи **временное** разделение применяется в цифровых системах передачи. При цифровой передаче значения сигнала передаются кодовыми группами (ИКМ). Как и в случае рассмотренной выше АИМ, сокращая длительность битов и, соответственно, кодовых групп можно освободить место для кодовых групп других каналов. Объединенные таким образом двоичные потоки образуют групповой сигнал. **В данном случае в одном интервале дискретизации содержатся кодовые группы всех каналов, которые вместе образуют «цикл».** В методе «чередования кодовых групп» в одном интервале дискретизации последовательно размещаются кодовые группы отдельных каналов. Более удобным способом является «чередование битов», при котором в каждом **тактовом интервале** последовательно

размещаются биты отдельных каналов. В обоих случаях мы обеспечиваем передачу сигналов отдельных каналов во взаимно-непересекающихся временных интервалах. На Рис.4.5–3 приведена временная диаграмма объединения двух цифровых потоков (мультиплексирования) методом чередования битов.

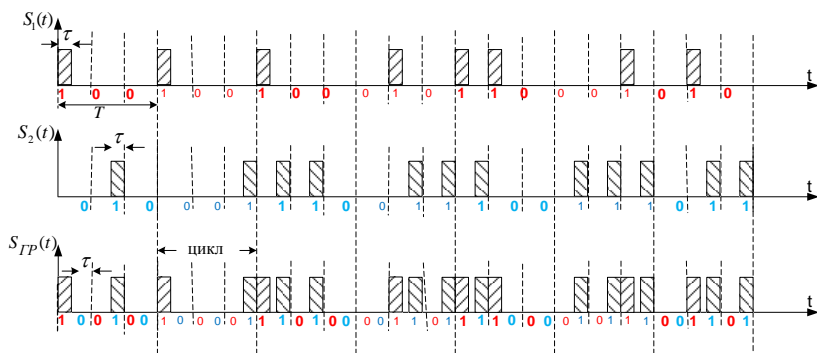


Рис.4.5–3. Принцип объединения двух цифровых потоков методом чередования битов.

При объединении цифровых потоков возникают те же вопросы, связанные с межсимвольной интерференцией и, соответственно, с взаимным влиянием каналов, что и при АИМ. Однако импульсы цифрового канала являются дискретными сигналами и не несут информации об амплитуде. Они менее чувствительны к действию помех, так как в данном случае отпадает необходимость определять высоту импульса, требуется лишь определить его наличие или отсутствие в данном тактовом интервале. Теория и практика показали, что в такой ситуации оказывается возможным построение эффективной системы многоканальной системы с временным разделением.

Групповой сигнал представляет собой поток двоичных символов. Длительность тактового интервала τ связана с шагом дискретизации T , числом каналов N и числом двоичных символов в кодовой группе m соотношением

$$\tau = \frac{T}{mN}. \quad (4.5-1)$$

Соответственно, для передачи группового сигнала требуется полоса в N раз бóльшая, чем при передаче цифрового потока от одного источника. Для приёма отдельных символов группового сигнала необходима тактовая синхронизация, которая **будет** рассмотрена в разделе 5.2.

После приема групповой последовательности требуется цикловая синхронизация, задачей которой является определение **границ циклов** группового сигнала. Далее производится демультимплексирование, то есть перенаправление двоичных символов каждого цикла к соответствующим получателям (в соответствии с цветами на Рис.4.5–3). Для этого применяются специальные алгоритмы в зависимости от принятого способа импульсной модуляции, способа временного объединения каналов, структуры сигналов синхронизации и т.д.

4.6. Регенерация дискретных сигналов

Если передаются квантованные отсчеты сигнала, то на приемном конце они будут также квантованы, но шаг квантования на приемном конце **будет** отличаться коэффициентом передачи линии связи: $\Delta_{ПРМ} = K \cdot \Delta_{ПРД}$. При прохождении сигнала по линии, на него, как известно, накладывается шум, величина которого случайна. Таким образом, на приемном конце мы получим значения, отличающиеся от уровней квантования (Рис.4.6–1).

Если на приемном конце мы повторим процедуру квантования, но с шагом $\Delta_{ПРМ}$, то получим некоторую последовательность квантованных отсчетов. При этом если величина шума не превосходит половины $\Delta_{ПРМ}$, то мы получаем исправленное значение данного квантованного отсчета. Такое восстановление квантованного сигнала называется **регенерацией**.

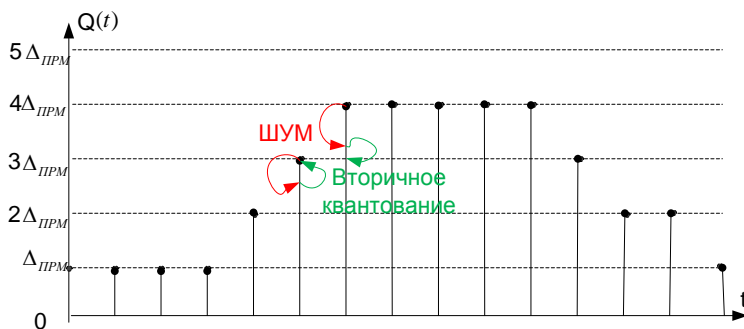


Рис. 4.6–1. Регенерация квантованного сигнала

Регенерация становится возможной благодаря тому, что при передаче квантованного сигнала передаются не любые значения, а вполне определенный набор сигналов – уровней квантования. На приемном конце путем повторного квантования выбирается в качестве переданного ближайший уровень (то есть наиболее вероятный). Если шум превышает значение $\Delta_{ПРМ} / 2$, то регенерация невозможна. Более того, процесс регенерации в данной ситуации приносит дополнительную ошибку. Качество регенерированного сигнала зависит от того, какая доля отсчетов была неверно регенерирована, то есть, была искажена на величину более $\Delta_{ПРМ} / 2$. Разрешенная доля неверно восстановленных отсчетов определяется техническими условиями на систему связи.

Если шум настолько велик, что доля искаженных отсчетов недопустимо велика, то можно воспользоваться тем обстоятельством, что шум растет с длиной линии, а сигнал затухает. Это значит, что линию можно разбить на участки с такой длиной, что соотношение величины сигнала и шума обеспечивает допустимую долю ошибок. В конце каждого участка (регенерационного участка) ставится регенератор с усилителем, и «очищенный» от шумов сигнал передается к следующему пункту регенерации. Конечно, такая процедура не может проводиться бесконечное число раз, поскольку происходит некоторое накопление ошибок. Однако, расстояние, на которое удается передать сигнал с удов-

летворительным качеством, оказывается значительно больше, чем при непосредственной передаче аналогового сигнала с использованием промежуточных усилительных пунктов.

Таким образом, применяя квантование, мы заранее привносим допустимое искажение в отсчеты аналогового сигнала, однако получаем возможность «очистки» этих отсчетов от шума, действующего в линии. Правильно организовав процесс передачи по линии, мы с большой вероятностью обеспечиваем допустимый уровень искажения сигнала на приемном конце.

Импульсы, с помощью которых передаются кодовые группы цифрового сигнала, по сути, являются сигналами, квантованными на два уровня. Поэтому становится очевидным, что процедура регенерации применима и к цифровым сигналам. Более того, поскольку в данном случае всего два уровня квантования, то различие между ними велико и очевидно, что такой сигнал допускает больший уровень помех. Это преимущество в некоторой степени компенсируется уменьшением длительности импульсов и, тем самым, расширением частотной полосы. Математический расчет показывает, что сигнал ИКМ в результате оказывается более помехоустойчивым, то есть обеспечивает надежную передачу при более высоком уровне шума, чем АИМ. Это является примером эффективного обмена полосы частот на динамический диапазон.

5. Передача дискретных сигналов в основной (первичной) полосе

5.1. Передача импульсов

Дискретные системы связи предназначены для передачи дискретных сигналов (последовательностей), поступающих с источника. Мы будем рассматривать двоичные последовательности, учитывая, что таковыми являются почти все последовательности, поступающие в системы передачи данных и в цифровые системы передачи. Итак, задачей дискретной системы связи является передача последовательностей двоичных сигналов, которые обычно обозначаются «1» и «0». Двоичными являются сигналы, поступающие от источника и передаваемые на приемном конце потребителю. Однако передача двоичных сигналов не означает, что во всех звеньях системы связи сигналы двоичные [3-§4.1].

Двоичный источник выдает элементы последовательности с некоторой скоростью R , означающей число двоичных символов (битов) в секунду:

$$R = n / t \text{ (бит/с)}. \quad (5.1-1)$$

Не следует путать «бит», используемый здесь как «двоичный символ» с битом как единицей количества информации. Соотношение (5.1-1) определяет «техническую» скорость источника в отличие от производительности источника, определяемой в рамках теории информации как количество информации, выдаваемое источником в секунду.

Временной интервал, который отведён для передачи каждого бита, называется **тактовым** интервалом. Из соотношения (5.1-1) получаем, что длительность тактового интервала равна

$$T = \frac{1}{R} \text{ (с)}. \quad (5.1-2)$$

Для построения системы связи необходимо выбрать два сигнала $S_1(t)$, $S_0(t)$, $0 < t < T$ для передачи соответственно «1» и

«0»). Представление двоичного символа соответствующим ему сигналом называется **дискретной модуляцией** или манипуляцией. Если в качестве сигналов выбираются **видео импульсы**, определенные на интервале T , то такая передача называется передачей в **основной (первичной) полосе или низкочастотной передачей**. Спектр такого сигнала широкополосный и начинается практически с нулевой частоты. Временная диаграмма такого сигнала имеет вид одного-двух колебаний, в отличие от радиоимпульса (см. Рис. 2.4–3).

Поскольку частотная полоса любой системы связи конечна, тот факт, что сигнал должен иметь конечную длительность, определяет главное теоретическое противоречие и практическую трудность при реализации дискретных (и цифровых) систем связи. Одно из свойств преобразования Фурье гласит, что длительность и ширина спектра сигнала не могут одновременно быть ограниченными. Поэтому, имея конечную частотную полосу для передачи, мы должны смириться с тем, что сигнал полностью не укладывается в выделенный временной интервал. Это вызывает **межсимвольную интерференцию (МСИ)** – явление, когда сигналы проникают во временные интервалы соседних символов и ухудшают их распознавание.

В качестве простейшей абстракции сигнала с ограниченным спектром применяется сигнал с равномерным ограниченным спектром (Рис.5.1–1). Это так называемая отсчетная функция типа $\frac{\sin x}{x}$, для нее имеется особое обозначение $\text{sinc}(x)$. Такой сигнал имеет базу, равную единице, то есть может быть восстановлен по одному отсчету и может быть передан системой связи, имеющей полосу пропускания F , однако у этого сигнала тот естественный недостаток, что его длительность неограниченна.

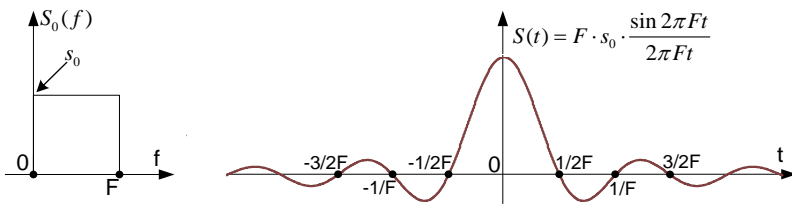


Рис.5.1-1. Сигнал с равномерным ограниченным спектром.

Теоретически, этот сигнал может быть использован для передачи со скоростью $R = 2F$, если принять, что импульс отсчитывается в середине тактового интервала. Согласно свойству отсчётной функции, ее отсчеты в точках, отстоящих на промежутки, кратные $\Delta t = 1/2F$ равны нулю и поэтому не влияют на отсчеты остальных импульсов, как показано на Рис.5.1–2.

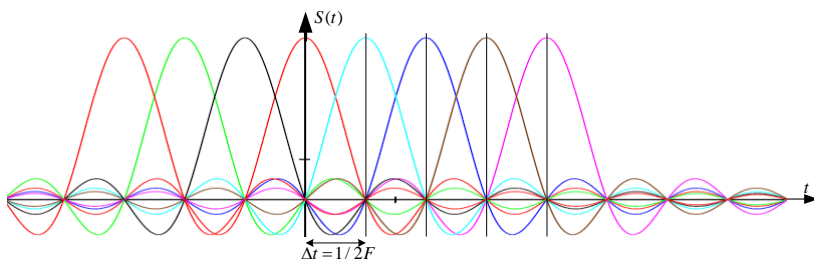


Рис.5.1–2. Последовательность импульсов в виде отсчётных функций при передаче со скоростью $R = 2F$

Таким образом, мы приходим к одному из основных результатов теории связи, называемому условием (барьером) Найквиста [3-§4.1].(Nyquist): максимальная скорость, с которой возможна передача импульсной последовательности, без МСИ, равна удвоенной ширине полосы канала:

$$R = 2F. \quad (5.1-3)$$

В технике связи используется понятие **плотности передачи**, определяемое как

$$\delta = R / F, \quad (5.1-4)$$

где R – скорость передачи, а F – частотная полоса, используемая для передачи.

Плотность передачи имеет размерность бит/с/Гц, то есть показывает, какая скорость в бит/с приходится на 1Гц полосы. Из (5.1–3) следует, что при передаче двоичных последовательностей посредством импульсов **максимальная теоретически возможная плотность передачи равна 2 бит/с/Гц.** Скорость передачи в реальных системах связи несколько ниже теоретической величины (5.1–3), так как использование импульсов в виде отсчетных функций невозможно.

В качестве простейшей абстракции сигнала, ограниченного по времени, обычно рассматривается идеальный прямоугольный импульс длительности T . На Рис.5.1–3 изображен такой импульс и его односторонний (физический) спектр. Спектр такого импульса бесконечен, так что никакая реальная система связи передать такой импульс не может.

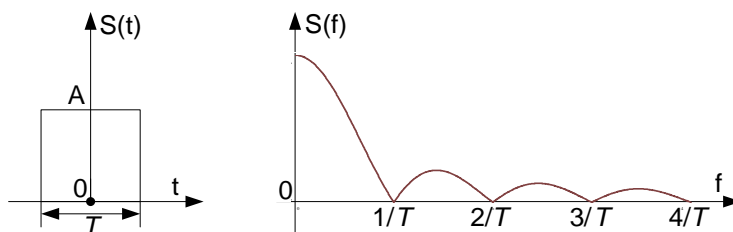


Рис.5.1–3. Прямоугольный импульс и его односторонний (физический) спектр.

На практике спектр импульсов ограничивается до величины, при которой импульс остаётся пригодным для передачи с нужной скоростью. На Рис.5.1–4 дана иллюстрация того, как ограничение спектра прямоугольного импульса величиной F , влияет на его форму.

Импульс со спектром, ограниченным величиной $1/2\tau$, «похож» на отсчетную функцию, но отсчёты в точках, отстоящих на промежутки, кратные $\Delta t = 1/2F$, не равны нулю. Ширина этого импульса почти вдвое больше ширины τ исходного импульса. Это

означает, что при использовании таких импульсов будет наблюдаться сильная МСИ. При более слабом ограничении спектра ($1/\tau$ или $3/2\tau$) ширина импульса меньше, однако применение импульсов с широким спектром нецелесообразно. На практике оптимальным считается ограничение спектра импульса величиной $0.86/\tau$ [3-§4.2].

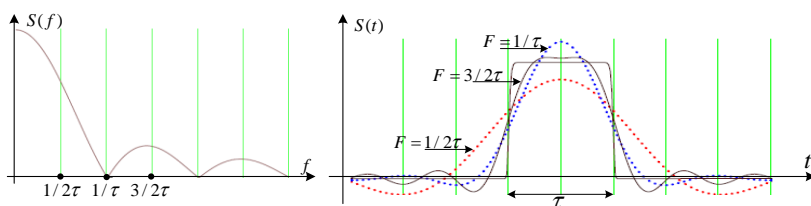


Рис.5.1–4. Изменение формы импульса при ограничении его спектра.

Мы рассмотрели простейший способ формирования импульса передачи путем идеального ограничения спектра идеального прямоугольного импульса. В реальных системах связи применяются различные практически реализуемые методы для формирования импульса передачи.

Увеличение плотности передачи достигается применением **многоуровневой модуляции**, когда для передачи выбирается большее число различных сигналов длительностью T , и в течение тактового интервала передается один из них. Обычно число различных сигналов выбирается равным степени двойки. В этом случае передача одного сигнала соответствует одновременной передаче нескольких битов.

Для выделения символов переданной последовательности на приемном конце необходимо знать моменты прихода импульсов (границы тактового интервала каждого принимаемого импульса). Это так называемая задача тактовой (символьной) синхронизации. Легко видеть, что при ограничении спектра передаваемых импульсов, при котором имеет место МСИ и «округление» импульсов, значение импульса в тактовом интервале наиболее четко

различно в некоторой области около середины тактового интервала. Поэтому отсчет принятого сигнала надо брать именно около середины тактового интервала. Это тем более важно, если учесть, что в линии на сигнал накладывается шум, так что вероятность ошибки будет минимальной именно при отсчете в точке (области), где сигналы «0» и «1» наиболее различимы. Все сказанное означает, что точность тактовой синхронизации играет важную роль в обеспечении надежного восстановления переданной последовательности на приемном конце.

5.2. Тактовая (символьная) синхронизация

На передающем конце дискретной системы связи формируется сигнал в виде последовательности двоичных символов (битов) в соответствии передаваемым сообщением. Дискретная электроника, формирующая данный сигнал, управляется **тактовым генератором** – генератором прямоугольного сигнала «CLOCK» с частотой, равной скорости передачи, то есть периодом, равным длительности тактового интервала.

На приемном конце для выполнения обратной задачи также необходим тактовый генератор. Он задает периодический импульсный сигнал, необходимый для выделения из принимаемого сигнала отдельных битов и формирования двоичной последовательности для передачи ее получателю. Очевидно, что этот генератор должен быть согласован (синхронизирован) с тактовым генератором передающего конца. Это означает, во-первых, что периоды (частоты) обоих генераторов должны быть равны. Тем самым, тактовая синхронизация обеспечивает равенство скоростей обработки цифровых сигналов в линейных и стационарных регенераторах, кодеках и других устройствах цифровой системы передачи (ЦСП), осуществляющих обработку сигнала с заранее заданной тактовой частотой. Кроме того, на приемном конце необходимо установить фазу тактового генератора, в соответствии с моментами прихода импульсов[3-§4.2].

Существует два основных режима цифровой передачи, основанных на принципиально различных методах согласования тактовых генераторов передающего и приемного конца. Первый метод – это **асинхронная передача**, которая подразумевает отдельную передачу определенных групп битов (например, байтов). При этом моменты передачи отдельных групп никак не связаны, что и означает асинхронность. При втором методе, называемом **синхронной передачей**, цифровые сигналы посылаются непрерывным потоком с постоянной скоростью (тактовой частотой). В этом случае на приемном конце необходимо обеспечить синхронность тактового генератора в течение неопределенного периода времени.

5.2.1. Асинхронная передача

При асинхронной передаче информация передается небольшими группами битов, обычно байтами [3–§4.2]. Моменты передачи отдельных групп никак не связаны, что и означает асинхронность. Период (частота) тактового генератора заранее задан равным периоду тактового генератора передатчика, и в процессе приема никак не синхронизируется.

Асинхронная линия в периодах между передачей находится в неактивном состоянии, которое обычно означает высокий уровень. Начало передачи очередной группы отмечается стартовым битом, то есть переходом линии в активное состояние (низкий уровень) на длительность, равную тактовому интервалу (Рис.5.2–1).

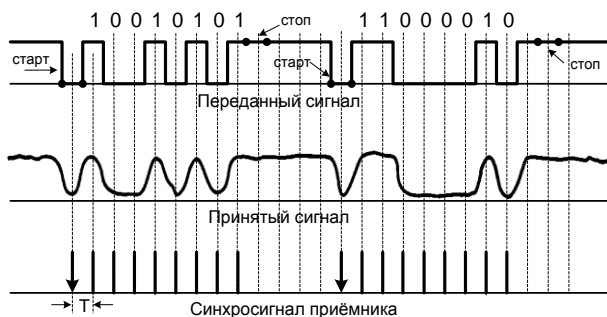


Рис.5.2–1. Старт-стопный принцип передачи.

При получении стартового бита в приемном устройстве запускается тактовый генератор. Таким образом, происходит синхронизация фазы тактового генератора приемной стороны, начальная фаза которого устанавливается в середине стартового бита, так что все последующие синхроимпульсы оказываются в середине тактового интервала. Тем самым тактовый генератор задает моменты отсчетов значений принимаемых информационных битов. Небольшое рассогласование частот генераторов передающей и приемной сторон может привести к постепенному сдвигу моментов отсчетов от середины тактового интервала. Однако длина принимаемой группы невелика, поэтому в течение приема одной группы существенный сдвиг произойти не может, то есть синхронизация сохраняется. После информационных битов передаётся один или несколько стоповых битов, после которых линия остается в неактивном состоянии до прихода очередного стартового бита.

То, насколько точно синхрогенератор приемника будет задавать отсчет каждого бита в середине каждого тактового интервала, зависит от точности определения середины стартового бита. Из-за наличия шумов стартовый бит искажается, и моменты отсчетов определяются с погрешностью тем большей, чем больше шум. Поэтому асинхронная передача плохо работает в системах с высоким уровнем шума. Применение асинхронного режима передачи ограничено низкоскоростной передачей на небольшие расстояния.

5.2.2. Синхронная передача

При синхронной передаче цифровые сигналы посылаются непрерывным потоком с постоянной скоростью (тактовой частотой). В этом случае на приемном конце необходимо обеспечить синхронность тактового генератора в течение неопределенного периода времени [3-§4.2]. В современных системах связи главным образом применяется метод «замкнутых синхронизаторов». В таких синхронизаторах осуществляется автоподстройка частоты тактового генератора приемника по принципу фазовой автопод-

стройки частоты (ФАПЧ). Синхронизирующий сигнал для работы ФАПЧ выделяется из принимаемой информационной последовательности в соответствии с поступающим двоичным потоком.

Принцип любой ФАПЧ заключается в том, что фазовый детектор (ФД) сравнивает фазы пришедшего сигнала и сигнала тактового генератора приемника (локального генератора) и вырабатывает сигнал, пропорциональный этой разности. Частота генератора управляется напряжением, поступающим на управляющий вход. Это так называемый «генератор, управляемый напряжением» (ГУН). Выходное напряжение ФД используется для управления ГУН (Рис.5.2–2). Если, например, фаза генератора начинает опережать фазу принимаемого сигнала, то это означает, что его частота выше частоты сигнала и выходной сигнал ФД должен уменьшить частоту генератора. Роль ФНЧ в данной схеме заключается в том, чтобы процесс регулирования частоты протекал плавно, без скачков.

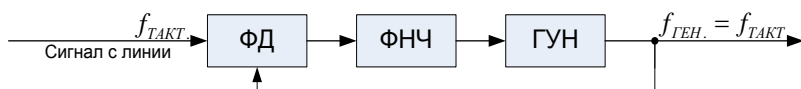


Рис.5.2–2. Простейшая схема ФАПЧ.

Поскольку входной сигнал, с которым должен синхронизироваться ГУН, в нашем случае — это двоичный поток, то в соответствии с этим строится схема ФД. Одной из схем автоподстройки частоты, применяемых для синхронизации в системах цифровой передачи, является схема с опережающим и запаздывающим стробированием (Рис.5.2–3).

Временная диаграмма, поясняющая его работу, приведена на Рис.5.2–4. Его работа заключается в выполнении двух отдельных интегрирований входного сигнала по двум различным интервалам одинаковой длительности $T-d$, где T — тактовый интервал. Первое интегрирование начинается с начала тактового интервала, определенного приемником, которое мы условно обозначаем «0», и

заканчивается через $T-d$ секунд. Второе интегрирование начинается с задержкой d секунд и соответственно заканчивается в конце тактового интервала. Интервалы интегрирования задаются соответственно сигналами «задний строб» и «передний строб». Если приемник правильно определил границы тактового интервала (Рис.5.2–4а), то оба интегрирования дадут один и тот же результат (штриховое заполнение), управляющее напряжение генератора не меняется и его частота сохраняется. Если же приемник неверно определил границы тактового интервала (Рис.5.2–4б), то интегрирование даст разные результаты и разность сигналов интеграторов является мерой ошибки синхронизации и используется для коррекции частоты генератора. В зависимости от знака этой разности частота генератора уменьшается или увеличивается.

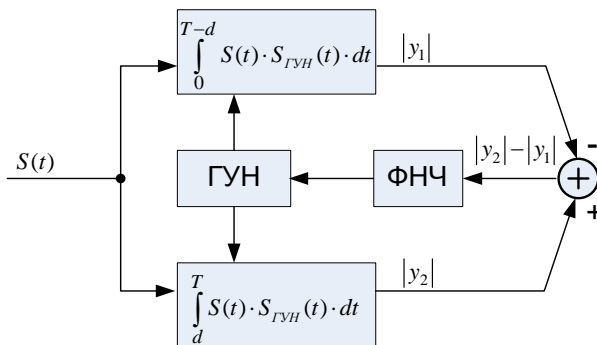


Рис.5.2–3. Схема автоподстройки частоты «с опережающим и запаздывающим стробированием».

В примере, проиллюстрированном на Рис.5.2–4, рассматривался отдельный импульс, до и после которого произошло изменение, то есть кодовая комбинация «010». В такой идеализированной ситуации, соответствующей передаче сигнала «...1010101010...», синхронизация будет хорошо сохраняться. В реальной системе передачи сигнал представляет собой случайную двоичную последовательность. В таком сигнале могут попадаться длинные последовательности одинаковых символов. В такой ситуации сигнал ошибки

синхронизации всегда равен нулю, подстройка не происходит, и генератор работает в самостоятельном режиме. Это требует высокой стабильности ГУН.

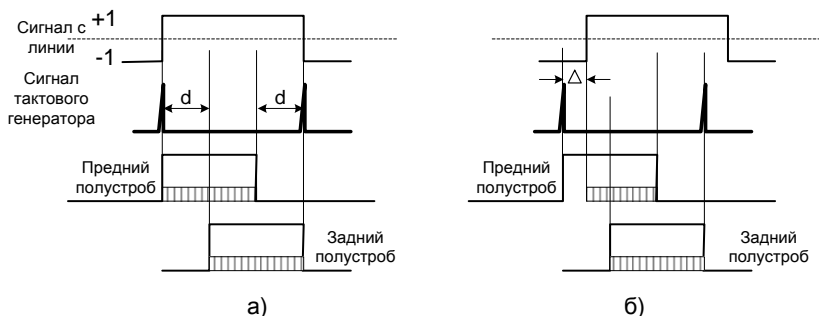


Рис. 5.2–4. Временная диаграмма работы синхронизатора.

С другой стороны, шумы могут влиять на результат интегрирования, в результате чего будет вырабатываться ошибочный сигнал ошибки синхронизации. Мы, конечно, ожидаем, что среднее значение шума равно нулю и поэтому вводим в схему фильтр ФНЧ, который усредняет эту ошибку, тем самым, делая ее меньше. Чем **уже** полоса ФНЧ, тем больше интервал усреднения и тем меньше ошибка (закон больших чисел). Однако при слишком узкой полосе фильтра система становится слишком медлительной. Нахождение компромиссного решения – сложная теоретическая и техническая задача.

Для обеспечения устойчивой работы систем синхронизации применяются различные методы перекодировки передаваемой последовательности с тем, чтобы обеспечить достаточно большое число переходов ($0 \rightarrow 1$ и $1 \rightarrow 0$) и минимизировать длины блоков из одинаковых символов. К таким методам относятся некоторые разновидности **кодирования для линии и скремблирования**.

5.3. Передача последовательностей

На Рис. 5.3–1 изображена некоторая импульсная последовательность, посредством которой передаётся двоичная (бинарная) информация[3-§4.3]. Здесь изображён простейший способ передачи, при котором «1» передаётся как высокий уровень в течение всего тактового интервала, а «0» –как нулевой уровень. Такой способ передачи означает, что для передачи «1» используется положительный импульс, а для передачи «0» – отсутствие сигнала (Рис. 5.3–2). Это так называемый униполярный код «без возвращения к нулю» (NRZ).

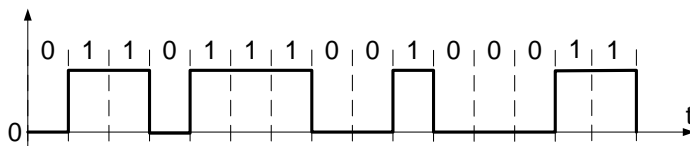


Рис.5.3–1. Представление двоичной информации в виде последовательности импульсов.

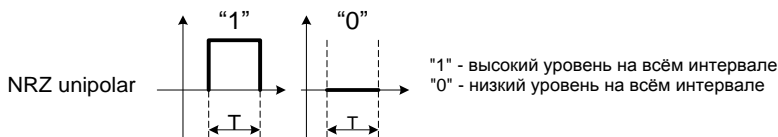


Рис.5.3–2. Выбор сигналов для «1» и «0» при кодировании «без возвращения к нулю».

Рассмотрим проблемы, которые могут возникнуть при передаче такого кода по линии связи. Во-первых, последовательность таких импульсов содержит постоянную составляющую, а большинство проводных линий связи содержат трансформаторы и конденсаторы, которые не пропускают постоянную составляющую. Поэтому применение униполярного кода «без возвращения к нулю» ограничивается соединениями в логических цепях с гальванической внутрисхемной и межкаскадной связью.

Для устранения этого недостатка в проводных линиях связи применяется биполярный вариант кода NRZ (Рис.5.3–3). Такой код, например, применяется в интерфейсе RS232.

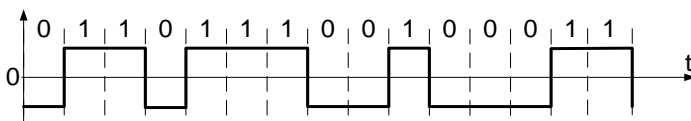


Рис.5.3–3. Передача биполярным вариантом кода NRZ.

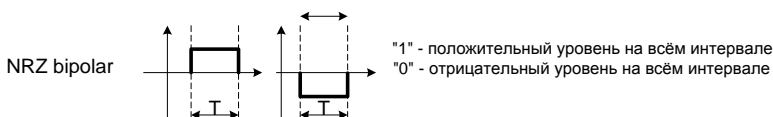


Рис.5.3–4. Выбор сигналов для «1» и «0» при биполярном коде NRZ.

Еще одно **преимущество биполярного кода** заключается в том, что при одинаковой разнице между уровнями средняя мощность сигнала вдвое меньше. Разница между уровнями определяет помехоустойчивость сигнала, то есть возможность различения уровней при наложении шума (помехи). Получается, что мы выигрываем по мощности вдвое, не ухудшая помехоустойчивости.

Однако переход от униполярного кодирования к биполярному коду не избавляет от проблемы постоянной составляющей. Появление в сигнале больших серий нулей или единиц эквивалентно появлению в его спектре очень низких частот, то есть «почти постоянной» составляющей, которая воспринимается индуктивной и емкостной связью между цепями «почти» как постоянная составляющая. Это выражается так называемым дрейфом нуля (Рис.5.3–5). Следовательно, кроме адекватного выбора сигналов, представляющих «1» и «0» (способа модуляции), требуется перекодировка импульсной последовательности перед её подачей в линию (кодирование для линии).

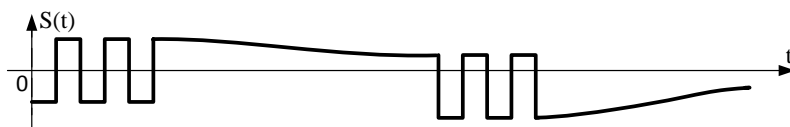


Рис. 5.3–5. Дрейф нуля при появлении серий одинаковых символов в последовательности.

Соображения, приведенные выше, относятся к физической возможности передачи импульсных последовательностей, однако есть целый ряд соображений, делающих кодирование для линии необходимым звеном в системах передачи дискретной информации.

5.4. Кодирование для линии

При формировании импульсных последовательностей, предназначенных для передачи информационной последовательности по линии, необходима оптимизация как отдельных сигналов (импульсов), так и структуры последовательностей. Первое называют модуляцией, а второе – кодированием для линии. Оптимизация означает формирование таких последовательностей, чтобы обеспечить возможность передачи сигнала по линии, а на приемном конце создать условия для правильного выделения переданной последовательности. Таким образом, выбором модуляции и линейного кодирования решается ряд практических задач с целью минимизации числа ошибок при приеме.

С точки зрения передачи, спектр сигнала должен быть по возможности узким. Это уменьшает влияние шумов, а также позволяет увеличить число каналов при частотном уплотнении линии. Формирование сигналов с узкой полосой является задачей как модуляции, так и линейного кодирования.

Для устойчивого приема необходимо на приемном конце иметь сигнал синхронизации, который указывает на моменты, когда надо отсчитывать значения принятых импульсов. Этот сигнал формируется на приемном конце на основе переходов в принятой

последовательности (от «1» к «0» или наоборот). Длинные отрезки из нулей или единиц могут нарушить синхронизацию и, тем самым, привести к ошибкам. Поэтому от линейного кодирования требуется обеспечить достаточную частоту переходов. Со спектральной точки зрения это означает отсутствие слишком низких частот в спектре сигнала. В литературе для этого свойства принят неудачный термин DC-free («без постоянной составляющей»).

Реальная постоянная составляющая зависит от того, какие импульсы используются. Например, при передаче в униполярном формате NRZ постоянная составляющая всегда имеется и ее существование не зависит от кодирования. Тем не менее, если при NRZ достаточная частота переходов обеспечивается, то сигнал будет «DC-free». Постоянная составляющая недопустима с точки зрения электроники, если связь между устройствами осуществляется с емкостной или трансформаторной развязкой. Для избавления от «неизбежной» постоянной составляющей применяются биполярные сигналы.

Одной из проблем при приеме последовательностей является межсимвольная интерференция (МСИ), означающая влияние предыдущих символов на приём каждого конкретного символа. Она возникает из-за расширения импульсов вследствие линейных искажений (ограничения спектра) и усугубляется плохой синхронизацией. Расширение импульса приводит к тому, что он проникает в соседний временной интервал (такты́вый интервал) и изменяет значение отсчета соседнего импульса, что повышает вероятность ошибочного решения. Уменьшение МСИ обеспечивается правильным выбором как типа модуляции, так и линейного кодирования [3-§4.3].

Кодирование для линии может быть простым и обусловленным (conditional).

При простом кодировании каждый символ «0» или «1» имеет свой «собственный» сигнал (импульс), с помощью которого этот

символ передается. Таким образом, простое кодирование означает выбор способа модуляции и никакого дополнительного кодирования не происходит. Наиболее распространенными простыми кодами являются:

1. код без возвращения к нулю (NRZ);
2. код с возвращением к нулю (RZ);
3. типа L (манчестерский, абсолютный бимпульсный) код, который называется также Phase Encoding (PE).

На Рис.5.4–1. показан выбор сигналов для указанных кодов, а на Рис.5.4–2. приведен пример кодирования некоторой последовательности.

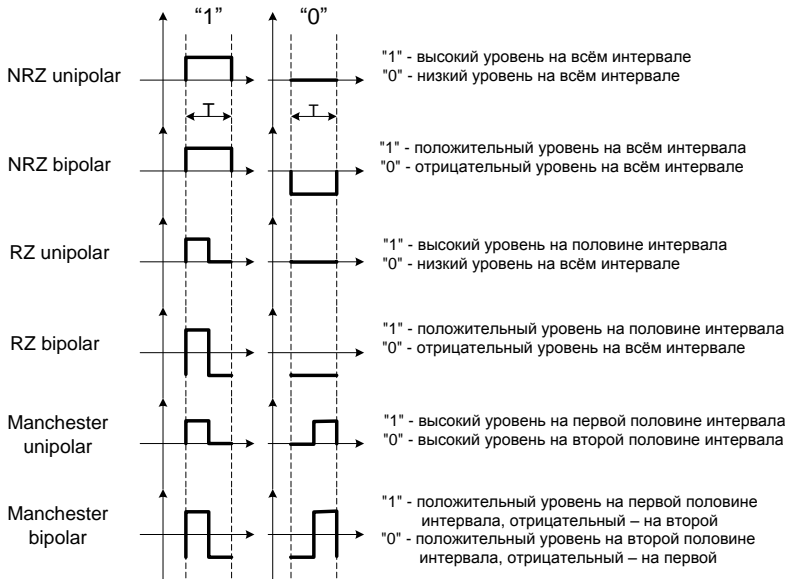


Рис.5.4–1. Выбор сигналов для некоторых простых кодов.

Изображенный на Рис.5.4–2 первый сигнал (CLOCK) представляет собой **синдром-последовательность**, которая управляет дискретной электроникой, формирующей импульсы. На передающем конце синхро-последовательность вырабатывается генератором синхронизирующих импульсов, который обеспечивает задан-

ную скорость передачи, задает длительность T и моменты начала тактового интервала. На приемном конце синхро-последовательность формируется также с помощью синхро-генератора, однако частота и фаза данного генератора подстраивается равной частоте и фазе передающего синхрогенератора. Подстройка осуществляется с помощью системы фазовой автоподстройки частоты (ФАПЧ). Назначение синхросигнала на приемном конце заключается в том, чтобы указывать на моменты времени, когда необходимо брать отсчет каждого поступающего импульса. Как уже было отмечено, для большей надежности отсчёт надо брать в середине тактового интервала. Это соответствует отрицательному фронту синхросигнала (CLOCK), показанного на Рис.5.4–2.

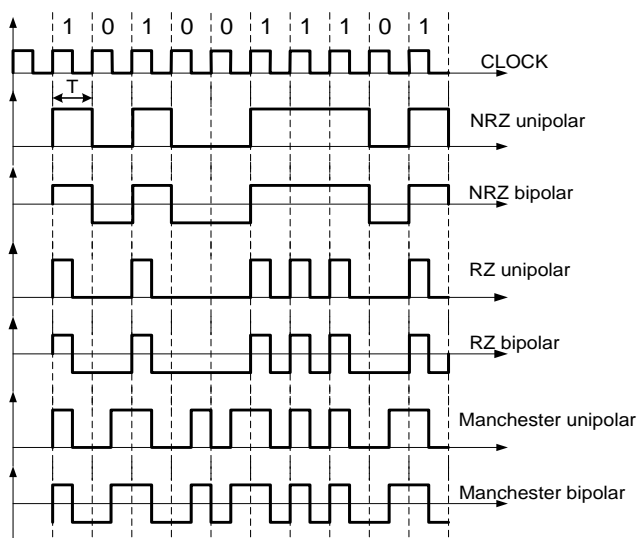


Рис.5.4–2. Пример кодирования двоичной последовательности некоторыми «простыми» кодами.

Ещё одно свойство кода, **полярность**, означает, что код не допускает перестановки проводников линии \pm . Соответственно, если код **неполярный**, то можно не обращать внимания на

полярность подключения проводников в каждой паре. Все описанные выше коды полярные.

При обусловленном (conditional) кодировании каждый символ передаваемой последовательности кодируется с учетом состояния определенных предыдущих битов. На Рис.5.4–3 изображены некоторые коды и описаны правила кодирования.

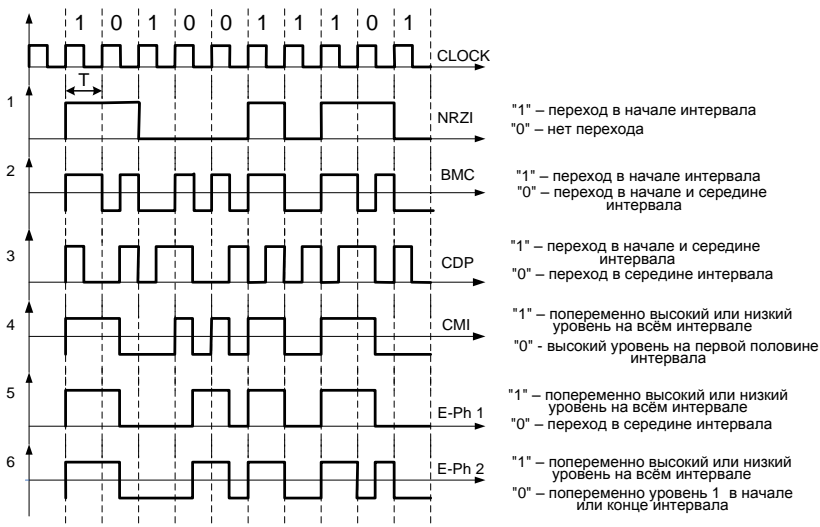


Рис.5.4–3. Пример кодирования двоичной последовательности некоторыми «обусловленными» (conditional) кодами.

1. Код без возвращения к нулю с инверсией единиц NRZI. Название связано с тем, что правило кодирования фактически означает, что единицы попеременно представляются то высоким, то низким уровнем, то есть в зависимости от того, каким уровнем была представлена предыдущая единица.

2. Двухфазный код типа S (BMC– Biphase Manchester Code). Здесь единицы представляются попеременно то высоким, то низким уровнем, а нули по ситуации высоким уровнем на первой или второй половине интервала.

3. Дифференциальный Манчестерский (CDP – conditional-diphase).

4. Код обращением «CMI coded mark inversion». В данном коде единица представляется попеременно: то высоким уровнем, то низким уровнем. Ноль представляется фиксированным сигналом – высоким уровнем на первой половине и низким – на второй половине интервала.

5. Электронно-фотонный код 1 (E-Ph 1)

6. Электронно-фотонный код 2 (E-Ph 2)

Все коды кроме NRZ удобно представлять в виде так называемых 1B2B- кодов. Это означает, что каждый бит представляется двумя битами.

Дальнейшее развитие методов кодирования связано с применением **троичных кодов**. Эти коды используют три уровня передачи, обычно -1, 0, 1 (биполярный троичный сигнал) и могут работать как в режиме RZ, так и NRZ.

Простейший троичный код – это «код с чередованием полярности» (AMI). Такой код использует для передачи «1» попеременно положительные и отрицательные уровни. Поэтому реальная постоянная составляющая при таком кодировании равна нулю. Однако при появлении длинной серии нулей из-за отсутствия переходов ухудшается синхронизация.

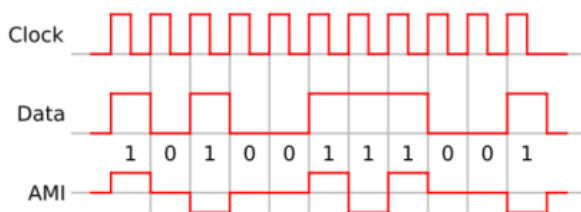


Рис.5.4–4. Временная диаграмма кода AMI в формате NRZ.

Этот недостаток устранен в кодах HDB3 (European E-carrier). Кодирование производится аналогично AMI, но длинные серии нулей заполняются двухполярными парами битов по определен-

ному правилу. Спектр таких кодов аналогичен коду «с обращением» (СМІ), то есть имеет ширину соответственно как RZ или NRZ, но низкие частоты малы (DC – free), а постоянная составляющая отсутствует.

Простейший метод NRZ используется в протоколах на базе интерфейса RS232. В сетях Ethernet применяется кодирование PE (phase encoding). В телефонии используется алгоритм HDB3. Этот метод служит для кодирования сигналов в потоках E1, E2 и E3 со скоростями, соответственно 2.048 Мбит/с, 8.448 Мбит/с и 34.368 Мбит/с. В потоке E4 со скоростью 139.264 Мбит/с используется код СМІ. Двухфазный код типа S используется в одной из самых известных сетей Token-Ring компании IBM. Код NRZI применяется в оптоволоконных сетях. Выбор метода передачи зависит от полосы канала связи, используемой кабельной системы, скорости передачи данных и других параметров.

5.5. Частотная полоса

Как мы видели, теоретическая минимальная полоса пропускания, необходимая для передачи со скоростью R , равна $\Delta F=0,5R$ при передаче отсчетной функцией. Таким образом, теоретически максимальная плотность передачи $\text{равна } \delta=2b\text{um/c/Гц}$.

Полоса пропускания, необходимая при передаче простейшим кодом NRZ с заданной скоростью передачи R , оценивается величиной $\Delta F=0,86R$. При использовании кода RZ требуемая полоса, очевидно, удваивается, $\Delta F=1,72R$, так как длительность импульсов вдвое меньше.

Все рассмотренные методы кодирования для линии (кроме NRZ) содержат импульсы половинной длительности и, естественно, их использование требует полосы, близкой к $1,72R$. На Рис.5.5–1 представлены спектральные плотности для некоторых рассмотренных кодов.

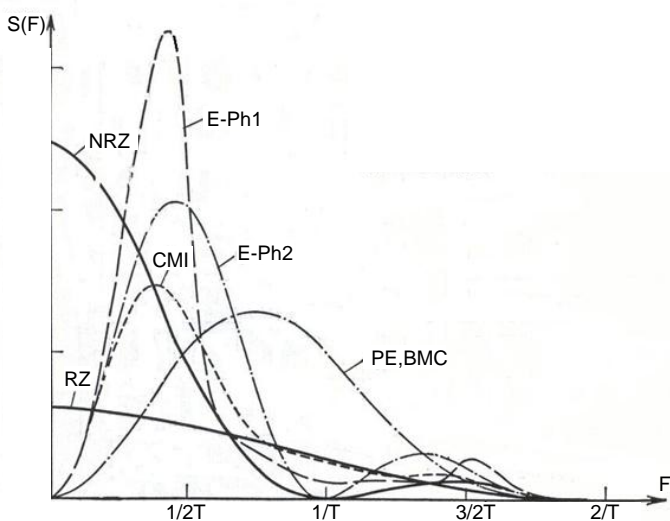


Рис. 5.5–1. Спектральная плотность некоторых кодов для линии.

Из графика видно, что коды RZ и NRZ содержат постоянную составляющую, так что их использование нежелательно с точки зрения обеспечения синхронизации. Остальные коды не имеют постоянной составляющей, однако коды PE и BMC невыгодны с точки зрения частотной полосы. Судя по графикам, наиболее удачными являются коды CMI, E-Ph2 и E-Ph1, причём полоса последнего самая узкая и оценивается величиной $1,52R$.

6. Дискретная модуляция гармонического несущего сигнала (полосовая передача)

Полосовая передача осуществляется путем модуляции гармонического несущего сигнала посредством передаваемой дискретной последовательности. В данном случае имеет место дискретная модуляция [1-§3.5,3-§6.1]. Соответственно, модуляция осуществляется путем изменения одного или нескольких параметров гармонического сигнала. Гармонический сигнал $S(t) = A \cdot \cos(\omega t + \varphi)$ имеет три параметра, а именно: амплитуда A , частота ω и фаза φ . Соответственно, выбирая один или несколько из них в качестве параметра модуляции, можно получить различные способы дискретной модуляции. В отличие от аналоговой модуляции, при которой параметр модуляции изменяется непрерывным образом, здесь при переходе от одного тактового интервала к другому параметр модуляции совершает скачок, соответствующий переходу в передаваемой последовательности. В простейшем случае параметры модуляции (A , ω и φ) принимают фиксированное значение в течение всего тактового интервала для каждого символа последовательности.

6.1. Двоичные методы модуляции

Двоичная модуляция ($m=2$) называется также **манипуляцией** [1-§3.5]. Если в качестве параметра модуляции выбрать один из параметров синусоиды, то можно получить три вида манипуляции – АМ, ФМ и ЧМ. Временные диаграммы сигналов этих видов двоичной модуляции на Рис.6.1–1.

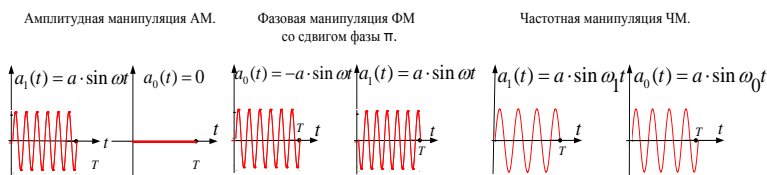


Рис.6.1-1. Временные диаграммы сигналов различных видов двоичной модуляции

Спектр амплитудно-манипулированного сигнала получается сдвигом спектра модулирующей последовательности на величину частоты несущего сигнала и образованием левой боковой полосы. Приблизительно этот спектр можно описать как спектр радиопульса (см. Рис. 2.4–3). Аналогичный спектр имеет сигнал фазовой манипуляции, но в нем отсутствует сигнал несущей частоты, как в сигнале БАМ.

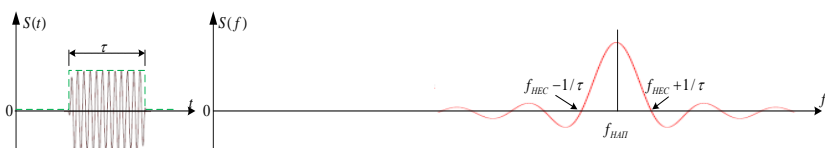


Рис.6.1–2. Временная и спектральная диаграммы сигнала АМ.

При частотной манипуляции возникает вопрос, чему должен быть равен сдвиг частот $\Delta f = f_1 - f_0$. Оптимальным является сдвиг частот, равный $\Delta f = 1/2T$, где T – длительность тактового интервала. Такая модуляция называется модуляцией с минимальным частотным сдвигом (MSK – minimum shift keying). Такое название связано с тем, что частотный сдвиг при MSK – это минимальный сдвиг, при котором выбранные сигналы являются ортогональными. При таком выборе частот полоса, занимаемая сигналом, достаточно узкая, и вместе с тем сигналы достаточно хорошо различимы на частотной оси (Рис.6.1-3).

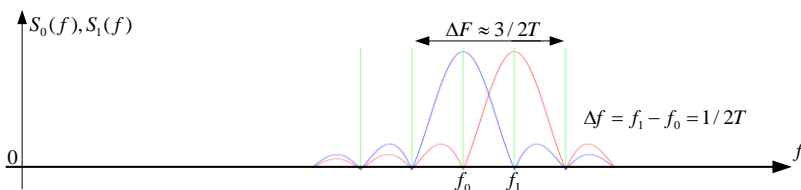


Рис.6.1–3. Спектральная диаграммы сигнала MSK.

Спектральная диаграмма сигналов MSK, изображенная на Рис.6.1–3 предполагает, что передача ведется прямоугольными

радиоимпульсами (Рис.2.4–3). На практике спектр передаваемых импульсов ограничивается, и импульс имеет форму, отличающуюся от прямоугольной (Рис.2.4–4). Если полосу сигнала MSK ограничить величиной ΔF , как показано на Рис.6.1–3, то импульсы будут иметь вид, показанный на (Рис.2.4–4). Как и при передаче в основной полосе (видеоимпульсами), это неизбежно приводит к межсимвольной интерференции (МСИ).

При передаче с модуляцией частотная полоса соответственно удваивается и плотность передачи вдвое меньше. Однако если применить ОБП (Рис.3.1–5), то плотность передачи совпадает с плотностью передачи в основной полосе.

6.2. Недвоичная модуляция

Для увеличения плотности передачи необходимо перейти от двоичной к многоуровневой передаче. В системах связи нашли применение многоуровневые методы ЧМ и ФМ, а также комбинированная «амплитудно-фазовая модуляция», которая называется **квадратурной АМ (КАМ)**. Многоуровневая АМ в чистом виде практически не применяется из-за плохих характеристик. Мы рассмотрим многоуровневую ФМ и КАМ [1-§3.5, 3-§6.1].

Для практической реализации многоуровневой ФМ удобно квадратурное представление гармонического сигнала с произвольной фазой (формула (2.2–2)):

$$\begin{aligned} S(t) &= \cos(\omega t + \varphi) = \cos \varphi \cos \omega t - \sin \varphi \sin \omega t = \\ &= A_{C\varphi} \cos \omega t + A_{KB} \sin \omega t. \end{aligned} \quad (6.2-1)$$

Так, мы представляем гармонический сигнал с произвольной фазой в виде суммы двух ортогональных составляющих — «синфазной» (к синусоидальный, с нулевой начальной фазой) и «квадратурной» (ортогональной, с фазой 90°) с амплитудами, соответственно, $A_{C\varphi} = \cos \varphi$ и $A_{KB} = -\sin \varphi$. Эти амплитуды называются квадратурными коэффициентами. Такое представление дает способ практического построения сигнала с произвольной фазой путем

сложения сигналов $\cos \omega t$ и $\sin \omega t$ с соответствующими коэффициентами. В сущности, сигналы $\cos \omega t$ и $\sin \omega t$ представляют собой базисные векторы в двумерном пространстве фазовой диаграммы (Рис.6.2–2). Для синфазной и квадратурной составляющих приняты обозначения соответственно **I** и **Q**.

Для получения 4-х уровневой ФМ («двукратной» фазовой манипуляции) выбираются четыре равномерно распределенные фазы, обычно, $\pi / 4$, $3\pi / 4$, $5\pi / 4$, $7\pi / 4$. Очевидно, для квадратурных коэффициентов получается следующая таблица.

Таблица 6.2–1

Сигнал	Квадратурные коэффициенты		Двоичная комбинация
	$A_{C\phi} = \cos \varphi$	$A_{KB} = -\sin \varphi$	
$\cos(\omega t + \pi / 4)$	0,707	-0,707	00
$\cos(\omega t + 3\pi / 2)$	-0,707	-0,707	01
$\cos(\omega t + 5\pi / 4)$	-0,707	0,707	10
$\cos(\omega t + 7\pi / 4)$	0,707	0,707	11

На практике в модуляторе берется один генератор несущего колебания, формируется ортогональный ему сигнал и далее в течение каждой пары тактовых интервалов суммируются эти ортогональные сигналы с соответствующими коэффициентами. Схематически, описанная процедура приведена на Рис.6.2–1. Аналогично можно описать квадратурную реализацию трехкратной ФМ (8-ФМ).

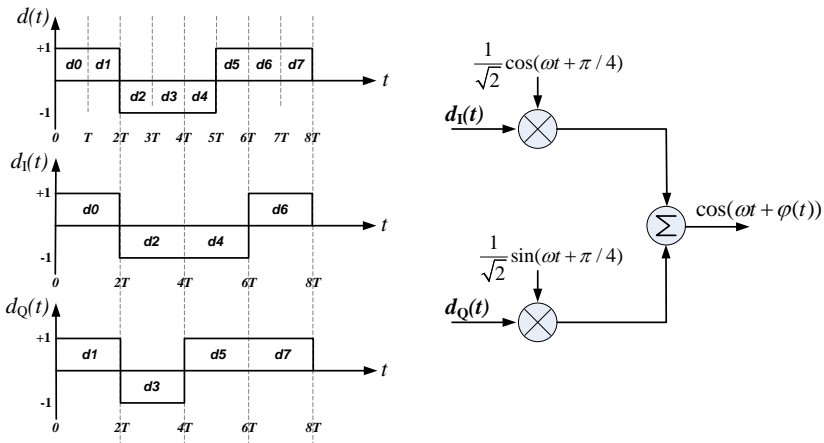


Рис. 6.2–1. Принцип кодирования для передачи двукратной ФМ.

Мы знаем, что гармонические сигналы можно представлять векторной диаграммой, которая показывает амплитуду и фазу сигнала, то есть полностью его описывает. Множество ФМ сигналов можно изобразить множеством векторов. Чтобы не загромождать диаграмму, вместо векторов можно показать только их концы. Учитывая, что у нас сигналы единичной амплитуды, мы получим множество точек, равномерно расположенных на единичной окружности (Рис. 6.2–2).

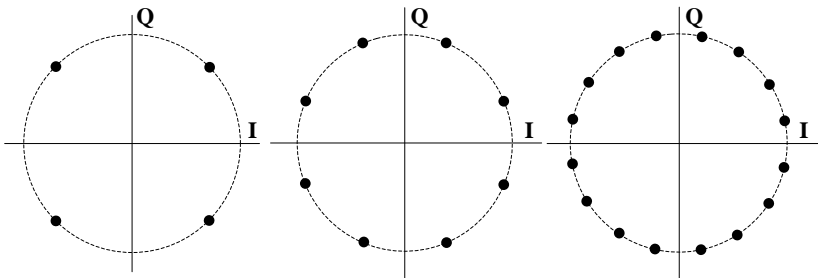


Рис. 6.2–2. Векторные (фазовые) диаграммы сигналов 4-ФМ, 8-ФМ и 16-ФМ.

Мы видим, что с увеличением числа ФМ сигналов уменьшается расстояние между сигналами, что означает снижение помехоустойчивости. Поэтому ФМ с числом сигналов **больше 8** практически не применяется.

КАМ отличается от ФМ тем, что сигналы отличаются не только фазой, но и амплитудой. Поэтому в случае КАМ квадратурное представление сигналов следует записать в более общем виде

$$\begin{aligned} S_i(t) &= A_i \cos \varphi_i \cos \omega t - A_i \sin \varphi_i \sin \omega t = \\ &= A_{C\Phi} \cos \omega t + A_{KB} \sin \omega t. \end{aligned} \quad (6.2-1)$$

Здесь $S_i(t)$ – это сигнал, предназначенный для передачи i -го символа. Соответственно, амплитуды «синфазной» и «квадратурной» компонент данного сигнала принимают вид $A_{C\Phi} = A_i \cos \varphi_i$ и $A_{KB} = -A_i \sin \varphi_i$. Если для ФМ задается набор фаз, то для КАМ задается набор пар (A_i, φ_i) . Таким способом единичный круг можно заполнить сигнальными точками более выгодным образом. На Рис.6.2–3 для сравнения приведены диаграммы 16-ФМ и 16-КАМ.

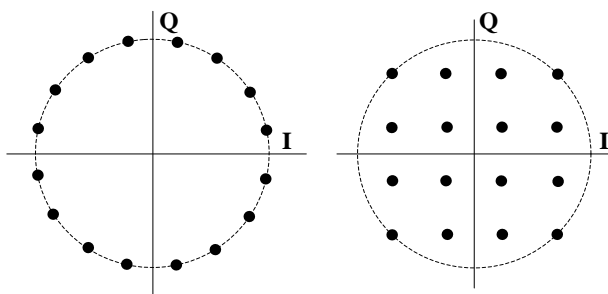


Рис.6.2–3. Фазовые диаграммы 16-ФМ и 16-КАМ.

Из рисунка видно, что при той же максимальной мощности расстояние между сигналами у 16-КАМ больше, чем у 16-ФМ, что означает лучшую помехоустойчивость. Очевидно также, что средняя мощность у 16-КАМ меньше, что также является преимуществом. При увеличении числа сигналов преимущество КАМ растет. По фазовой диаграмме Рис.6.2–3 легко рассчитать квадратурные коэффициенты для КАМ аналогично тому, как это делалось в случае ФМ.

7. Основы помехоустойчивого кодирования

7.1. Постановка задачи помехоустойчивого кодирования

При передаче двоичной последовательности по каналу связи некоторые символы на приемном конце могут быть приняты неверно. Задача помехоустойчивого кодирования состоит в том, чтобы так закодировать передаваемую информацию, чтобы на приёмном конце можно было бы **обнаружить** или **исправить** произошедшие ошибки [1-§7.3]. «Обнаружить» означает, что на приемном конце станет известным, что в принятой последовательности есть неверно принятые символы, однако не будет возможности определить их позиции. Если имеется возможность определить также позиции ошибочных символов, то их можно исправить. В таком случае говорят, что код обеспечивает исправление ошибок. В зависимости от практической задачи могут применяться как коды, исправляющие ошибки, так и коды, обнаруживающие ошибки. В последнем случае обычно предоставляется обратный канал для возможности переспроса.

Мы будем рассматривать так называемые **блоковые коды**. При блоковом кодировании информационная последовательность разбивается на **блоки (информационные слова)** определенной длины k . Кодирование заключается в том, что каждое информационное слово $A^{(k)} = (a_1, a_2, \dots, a_k)$ длины k заменяется некоторым **кодовым словом** $C^{(n)} = (c_1, c_2, \dots, c_n)$ длины $n > k$. Это означает, что с каждым словом будут передаваться на $m = n - k$ символов больше, чем их содержится в информационном слове, то есть оно содержит избыточные символы. Таким образом, для передачи будут использоваться не все последовательности длины n , а некоторое подмножество таких последовательностей. Число всех последовательностей длины n равно 2^n , а число кодовых слов равно 2^k .

Кодом называется подмножество $\{C^{(n)}\}$ кодовых слов длины n , предназначенных для передачи информационных слов $A^{(k)}$.

Величина

$$R = k / n \quad (7.1-1)$$

называется относительной скоростью кода. Величина же

$$R = m / n \quad (7.1-2)$$

называется избыточностью кода. Именно наличие избыточных символов определяет способность кода исправлять или обнаруживать ошибки. Это обусловлено тем, что не все последовательности длины n являются кодовыми словами, то есть не все будут передаваться по каналу.

Корректирующая (исправляющая) **способность кода** обычно задается максимальным числом ошибок t в принятом слове, при котором код исправляет или обнаруживает эти ошибки. Соответственно, код называется «кодом, исправляющим t ошибок» или «кодом, обнаруживающим t ошибок».

Основной задачей теории помехоустойчивого кодирования является разработка корректирующих кодов, то есть выделения из множества всех последовательностей длины n подмножества (кода), состоящего из 2^k элементов, применение которых в качестве кодовых слов обеспечит заданную корректирующую способность. Дальнейшей задачей является создание практически реализуемого алгоритма, по которому каждому из 2^k информационных слов ставится в соответствие кодовое слово (алгоритм кодирования). И, наконец, наиболее сложной задачей является разработка практически реализуемого алгоритма декодирования.

Сложность построения кодов, то есть выбора подмножества кодовых слов, заключается в том, что нужно выбрать кодовые слова в необходимом количестве, причем выбрать так, чтобы они отличались друг от друга в максимальном числе позиций. Если учесть, что эффективность кода тем выше, **чем больше его длина**

n , то понятно, что разработка кода является сложной математической задачей, которую перебором не решить [3-§4.5].

Если код с нужными параметрами уже построен, то далее возникает практическая **проблема кодирования**. Для кодирования можно использовать таблицу, показывающую, какое кодовое слово длины n соответствует данному информационному слову длины k , то есть дающая соответствие $(a_1, a_2, \dots, a_k) \rightarrow (c_1, c_2, \dots, c_n)$.

Такая таблица содержит 2^k строк, и перебор при кодировании отнимает много времени. Если учесть, что для эффективного кодирования приходится применять достаточно длинные блоки, то в такой постановке проблема кодирования в реальном времени оказывается безнадежной.

Еще хуже дело обстоит на приемном конце при декодировании. В случае кода с **обнаружением ошибок** алгоритм декодирования достаточно прост, он проверяет, является ли принятое слово кодовым. Если принятое слово не совпадает ни с одним кодовым словом, значит, оно содержит ошибочные символы. Здесь также требуется перебор в том же объеме, что и при кодировании, так что табличное декодирование в реальном времени практически не применимо.

В случае кодов с **исправлением ошибок** алгоритм декодирования должен ставить в соответствие каждому **принятому** слову то **кодовое** слово, которое получается из принятого путем искажения наименьшего числа позиций. Это так называемый алгоритм декодирования по минимальному «расстоянию». Данный алгоритм предполагает, что меньшее число искажений при передаче по каналу более вероятно. Несмотря на столь простой принцип, его табличная реализация требует еще большего числа операций, чем при декодировании с обнаружением ошибок.

Для создания практически реализуемых корректирующих кодов и алгоритмов кодирования и декодирования необходимо множество кодовых слов наделять определенной математической структурой и затем, используя соответствующий математический аппарат, строить коды и алгоритмы[4-§1.2].

7.2. Математическое описание двоичных кодов

В данном разделе рассматривается наиболее распространенный математический аппарат, позволяющий адекватно описывать кодовые слова, коды и процедуры кодирования и декодирования [4-Гл.2].

Двоичные символы рассматриваются как элементы поля GF(2). **Поле** – это алгебраическая структура с двумя операциями, которые называются «сложением» и «умножением». Поле GF(2) – это конечное поле, состоящее из двух элементов, которые обозначаются «0» и «1». Операции сложения \oplus и умножения \otimes , так называемые арифметические операции «по модулю 2», задаются таблицами 7.2–1, причем вычитание совпадает со сложением.

Таблица 7.2-1

\oplus	0	1	\otimes	0	1
0	0	1	0	0	0
1	1	0	1	0	1

Соответственно, слова S^n длины n рассматриваются как n -мерные векторы. В этом случае множество S^n последовательностей S^n длины n образует линейное (векторное) пространство размерности n над GF(2). Сложение векторов определяется обычным образом как покомпонентное сложение. При этом, компоненты (координаты) векторов складываются по модулю 2:

$$S_1^n = (s_1^1, s_2^1, \dots, s_n^1), \quad S_2^n = (s_1^2, s_2^2, \dots, s_n^2),$$

$$S_1^n \oplus S_2^n = (s_1^1 \oplus s_1^2, s_2^1 \oplus s_2^2, \dots, s_n^1 \oplus s_n^2). \quad (7.2-1)$$

На пространстве $S^{(n)}$ задаётся метрика, то есть функция, определяющая расстояние между любыми двумя элементами этого пространства, которая называется **метрикой Хэмминга**. Пусть $S_1^n = (s_1^1, s_2^1, \dots, s_n^1)$ и $S_2^n = (s_1^2, s_2^2, \dots, s_n^2)$ – двоичные векторы. Расстояние в метрике Хэмминга определяется как:

$$d_H(S_1^n, S_2^n) = \sum_{i=1}^n (s_i^1 \oplus s_i^2). \quad (7.2-2)$$

В формуле (7.2-2) суммирование слагаемых проводится как обычно, а сложение в каждом слагаемом (в скобках) – по модулю 2 [4-§3.1]. Поэтому расстояние Хэмминга фактически равно числу отличающихся **символов в S_1^n и S_2^n** [4-§3.1].

Норма (длина, вес) вектора определяется как

$$\|S^n\| = d_H(S^n, 0) = \sum_{k=1}^n s_n, \quad (7.2-3)$$

и равна числу ненулевых символов данного вектора. Из определений (7.2-1, 2, 3) следует, что расстояние Хэмминга равно весу суммарного (разностного) вектора:

$$d_H(S_1^n, S_2^n) = \|S_1^n \oplus S_2^n\| = \sum_{i=1}^n (s_i^1 \oplus s_i^2). \quad (7.2-4)$$

Теперь мы можем описать процесс появления ошибок в принятом слове как сложение его с «вектором ошибки» E^n , то есть принятое слово

$$Z^n = S^n + E^n. \quad (7.2-5)$$

При таком описании единица «1» в векторе ошибки указывает на ошибочно принятые символы. Очевидно также, что кратность ошибки (число ошибок) равна весу Хэмминга вектора ошибки.

В свете введенных понятий можно дать геометрическую интерпретацию принципов помехоустойчивого кодирования. На Рис.7.2-1 **изображено в пространстве $S^{(n)}$** двоичных слов длины n , в котором выбрано 2^k кодовых слов C_1, C_2, \dots , множество которых представляют собой код $C^{(n)}$.

Кодовым расстоянием кода d называется минимальное расстояние между кодовыми словами:

$$d = \min_{k \neq l} \{d_{k,l}\}, \text{ где } d_{k,l} = d_H(C_k, C_l) = \sum_{i=1}^n (c_i^k - c_i^l). \quad (7.2-6)$$

Очевидно, что код с кодовым расстоянием d позволяет обнаруживать любые ошибки кратности $r \leq d - 1$. Это понятно из того, что ошибка веса $d - 1$ и меньше не может превратить ни одно кодовое слово в некоторое другое кодовое слово. Поэтому принятое слово будет не кодовым и приемник опознает его как ошибочно [3-§4.5, 4-§1.6].

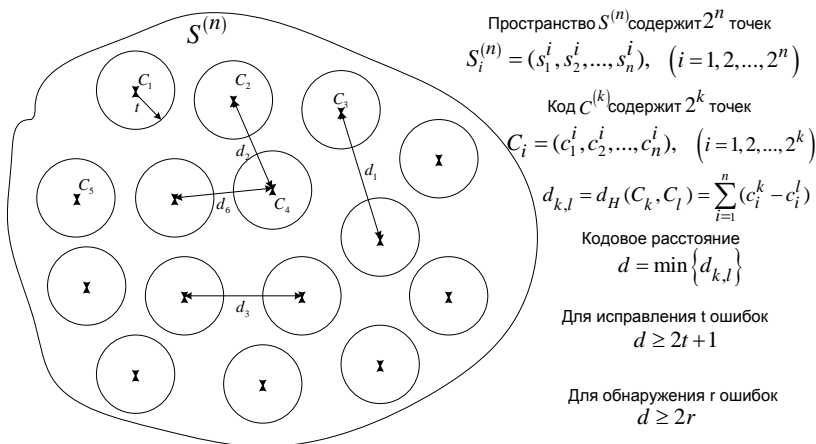


Рис.7.2–1. Геометрическая интерпретация кодирования и коррекции ошибок.

Если говорить об исправлении ошибок, то надо вспомнить, что декодер ищет кодовое слово, ближайшее к принятому слову (считая его наиболее вероятным). Для того чтобы код гарантированно исправлял все ошибки веса до t , необходимо, чтобы ни один вектор ошибки веса t , прибавленный к любому кодовому слову не сделал его ближе к другому кодовому слову. Тогда принятое слово останется ближайшим к переданному слову и декодер его правильно найдет. Если произойдет ошибка кратности больше t , то возможны ситуации, когда принятое (искаженное) слово ближе к соседнему кодовому слову, и декодер произведет неверное декодирование.

Ситуация с исправлением ошибок аналогична регенерации квантованного сигнала, рассмотренной в разделе 4.5.

Код, гарантированно исправляющий все ошибки веса до t , должен иметь кодовое расстояние $d \geq 2t+1$. В последнем выражении единица добавлена для того чтобы исключить ситуацию, когда полученное слово окажется на расстоянии t сразу от двух кодовых слов.

Корректирующие коды чаще всего строятся в систематическом виде. В случае **систематических кодов** кодирование осуществляется путем добавления к информационному слову $A = (a_1, a_2, \dots, a_k)$ проверочных символов (g_1, g_2, \dots, g_m) , в результате чего получается кодовое слово $C = (a_1, a_2, \dots, a_k, g_1, g_2, \dots, g_m) = (c_1, c_2, \dots, c_n)$. Проверочные символы g_1, g_2, \dots, g_m вычисляются по информационным символам a_1, a_2, \dots, a_k согласно заданному алгоритму, который и определяет код и, соответственно, **процедуру кодирования**.

7.3. Линейные коды

Код мы определили как подмножество $\{C^{(n)}\}$ слов длины n , предназначенных для передачи информационных слов $A^{(k)}$. Линейные коды [1-§7.3, 4-§3.1, 3.2] описываются как линейное подпространство $C^{(k)}$ размерности k линейного пространства $S^{(n)}$ размерности n . Здесь k означает именно размерность кодового подпространства, а не длину кодовых слов, которая равна n .

Из данного определения следует, что все кодовые слова должны удовлетворять некоторой однородной системе линейных уравнений над полем $GF(2)$.

$$C \cdot H^\dagger = 0, \quad (7.3-1)$$

где H – матрица коэффициентов этой системы, которая называется проверочной матрицей данного кода (здесь † означает транспо-

нирование матрицы). Матрица H состоит из n столбцов и $n - k = m$ строк:

$$H = \begin{pmatrix} h_{11} \dots h_{1n} \\ h_{21} \dots h_{2n} \\ \dots \dots \dots \\ h_{m1} \dots h_{mn} \end{pmatrix}. \quad (7.3-2)$$

Строки матрицы представляют собой базисные векторы подпространства, ортогонального к $C^{(k)}$.

Кодовые слова линейного кода образуют группу по сложению, то есть сумма или разность любой пары кодовых слов также является некоторым кодовым словом. Отсюда следует (см.7.2-4), что **кодвое расстояние линейного кода равно минимальному ненулевому весу Хемминга кодовых слов.**

Как известно из линейной алгебры, при решении системы из m уравнений с n неизвестными, значения $n - m$ неизвестных могут быть выбраны произвольно. Остальные m неизвестных находятся как решения оставшейся системы размерности $m \times m$. Таким образом, $n - m = k$ неизвестных можно выбрать в качестве информационных символов, то есть задавать в соответствии с передаваемым (информационным) словом. Проверочные символы вычисляются как решение системы линейных уравнений (7.3-1), или в раскрытом виде

$$\begin{aligned} h_{11}a_1 \oplus h_{12}a_2 \oplus \dots \oplus h_{1k}a_k \oplus h_{1(k+1)}g_1 \dots \oplus h_{1n}g_m &= 0 \\ h_{21}a_1 \oplus h_{22}a_2 \oplus \dots \oplus h_{2k}a_k \oplus h_{2(k+1)}g_1 \dots \oplus h_{2n}g_m &= 0 \\ \dots \dots \dots & \\ h_{m1}a_1 \oplus h_{m2}a_2 \oplus \dots \oplus h_{mk}a_k \oplus h_{m(k+1)}g_1 \dots \oplus h_{mn}g_m &= 0 \end{aligned} \quad (7.3-3)$$

Проверочную матрицу можно преобразовать так, что справа образуется единичная матрица размером $m \times m$, то есть она примет вид:

Синдром является m -компонентным вектором. Раскрывая выражение (7.3–6), получаем соотношения для вычисления компонентов синдрома:

$$\begin{aligned} S_1 &= h_{11}b_1 \oplus h_{12}b_2 \oplus \dots \oplus h_{1k}b_n = h_{11}e_1 \oplus h_{12}e_2 \oplus \dots \oplus e_{1k}b_n \\ S_2 &= h_{21}b_1 \oplus h_{22}b_2 \oplus \dots \oplus h_{2k}b_n = h_{21}e_1 \oplus h_{22}e_2 \oplus \dots \oplus h_{2k}e_n \end{aligned} \quad (7.3-7)$$

$$S_m = h_{m1}b_1 \oplus h_{m2}b_2 \oplus \dots \oplus h_{mk}b_n = h_{m1}e_1 \oplus h_{m2}e_2 \oplus \dots \oplus h_{mk}e_n.$$

Декодирование с исправлением ошибок для линейного кода также начинается с подстановки символов принятого слова $B = b_1, b_2, \dots, b_n$ в систему проверочных уравнений (7.3–7), то есть с получения синдрома. Если синдром равен нулю, то ошибки нет. Если же синдром не равен нулю, то необходимо найти вектор ошибки E . Однако система уравнений $S = E \cdot H^\dagger$ имеет множество решений. В технике связи нас интересует решение с минимальным весом Хемминга E_{MIN} , поскольку оно соответствует наиболее вероятному вектору ошибки. Таким образом, решение с минимальным весом даст нам наиболее вероятное значение переданного слова:

$$C = B - E_{MIN}. \quad (7.3-8)$$

Соотношение (7.3–8) есть алгоритм исправления ошибки. Если вес Хемминга ошибки меньше половины кодового расстояния, то мы гарантированно правильно декодируем переданное слово. Нахождение решения с минимальным весом Хемминга является наиболее сложной частью алгоритма декодирования линейного кода.

7.4. Простейшие линейные коды

7.4.1. Код с одной проверкой на четность

Простейший линейный код «с одной проверкой на чётность» строится по проверочной матрице

$$H = (1, 1, \dots, 1). \quad (7.4-1)$$

Проверочное уравнение для кодирования при этом получается

$$\bigoplus \sum_{i=1}^n c_i = 0, \quad (7.4-2)$$

где знак \bigoplus означает сложение в поле $GF(2)$, то есть по модулю 2. Здесь $n-1$ символов можно выбирать произвольно, а последний, проверочный символ вычисляется как

$$c_n = \bigoplus \sum_{i=1}^{n-1} c_i. \quad (7.4-3)$$

Таким образом, для простейшего линейного кода с проверкой на четность $R = (n-1) / n$.

В результате получается, что любое кодовое слово кода с проверкой на чётность содержит четное число единиц. Легко увидеть, что минимальный ненулевой вес кодового слова, а следовательно, и **кодовое расстояние простейшего линейного кода с проверкой на четность равно 2**. Это значит, что данный код не даёт возможности исправлять ошибки, но позволяет обнаруживать одну ошибку. **Более точно, код с проверкой на четность позволяет обнаруживать любую ошибку нечетной кратности.**

Синдромом ошибки кода с проверкой на четность для принятого слова $B = b_1, b_2, \dots, b_n$ вычисляется как:

$$S = \bigoplus \sum_{i=1}^{n-1} b_i. \quad (7.4-4)$$

По значению синдрома устанавливается наличие ошибки в принятом слове. А именно, если $S = 0$, то принимаем решение, что ошибки нет, и принятое слово предоставляем получателю. Если же $S \neq 0$, то получателю сообщается о том, что слово принято неверно.

7.4.2. Код Хемминга

Код Хемминга [3-§4.5, 4-§5.1] является линейным кодом, исправляющим одну ошибку, то есть имеет кодовое расстояние равное $d = 3$. Длина кодовых слов равна $n = 2^m - 1$, число проверочных символов равно m и число информационных символов $k = 2^m - 1 - m$. Таким образом, относительная скорость кода Хемминга

$$R = \frac{(2^m - 1 - m)}{2^m - 1}. \quad (7.4-5)$$

Проверочная матрица кода состоит из всех ненулевых двоичных последовательностей длины m . Например, при $m=3$ имеем код Хемминга (7,4) с проверочной матрицей

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left(\begin{array}{ccccccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right). \end{matrix} \quad (7.4-6)$$

Над матрицей написаны числа, показывающие, что столбцы упорядочены так, что столбец является двоичной записью номера позиции, на которой он расположен. При любом расположении столбцов проверочной матрицы позиции проверочных символов следует выбирать так, чтобы они совпадали со столбцами, в которых только одна единица. Это возможно, так как таких столбцов точно m штук. При этом решение системы уравнений относительно проверочных символов упрощается, так как в каждое уравнение входит только один проверочный символ. Поэтому при расположении столбцов как в (7.4-6) выбираем в качестве прове-

рочных символов первый, второй и четвертый символы. Тогда, в соответствии с (7.3–5), соотношения для нахождения проверочных символов будут:

$$\begin{aligned} g_1 &= a_1 \oplus a_2 \oplus a_4 \\ g_2 &= a_1 \oplus a_3 \oplus a_4 \\ g_3 &= a_2 \oplus a_3 \oplus a_4 \end{aligned} \quad (7.4-7)$$

Код в этом случае имеет структуру:

$$= (g_1, g_2, a_1, g_3, a_2, a_3, a_4) = (c_1, c_2, \dots, c_7).$$

Аналогично можно построить матрицу кода Хемминга любого порядка и с любым расположением столбцов. Чтобы получить код Хемминга (7,4) в систематической форме, следует расположить столбцы следующим образом:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}. \quad (7.4-8)$$

При таком расположении столбцов проверочными будут первые три символа, то есть код имеет следующую структуру:

$$C = (g_1, g_2, g_3, a_1, a_2, a_3, a_4) = (c_1, c_2, \dots, c_7).$$

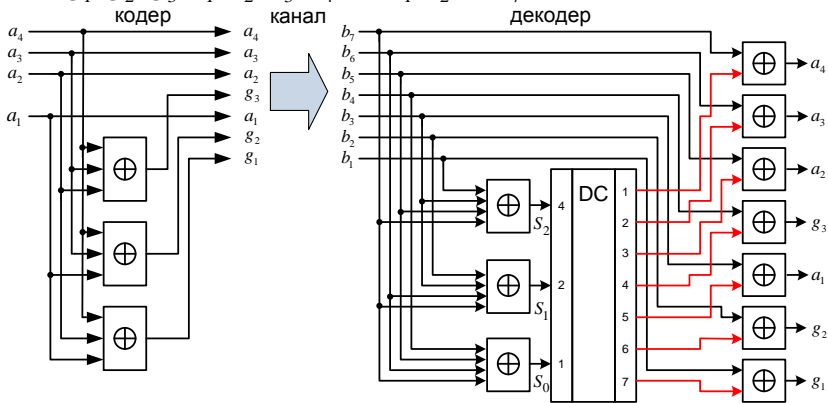


Рис. 7.4–1 Схема кодирования и декодирования кода Хемминга (7,4)

На Рис.7.4–1 приведена схема кодирования и декодирования кода Хемминга (7,4) с проверочной матрицей вида (7.4–6). Здесь элементы \oplus – сумматоры по модулю 2.

Кодер реализует алгоритм (7.4–8) , формируя код со структурой $g_1g_2a_1g_3a_2a_3a_4$.

Декодер вычисляет синдром в соответствии с выражением (7.3–7). Если раскрыть данное выражение для нашего случая, **получим** следующие соотношения для трех компонентов синдрома:

$$\begin{aligned} S_1 &= b_1 \oplus b_3 \oplus b_5 \oplus b_7 \\ S_2 &= b_2 \oplus b_3 \oplus b_6 \oplus b_7 . \\ S_3 &= b_4 \oplus b_5 \oplus b_6 \oplus b_7 \end{aligned} \quad (7.4-9)$$

Если при передаче в канале произошла одна ошибка, то вектор ошибки содержит одну единицу в позиции, где эта ошибка произошла. Синдром в этом случае равен столбцу проверочной матрицы, соответствующему позиции ошибки. И, поскольку мы расположили столбцы матрицы как номера столбцов в двоичном выражении, то синдром совпадает с двоичной записью номера позиции ошибки. Поэтому на вход дешифратора ДС с сумматоров декодера поступает двоичный номер позиции ошибки. Дешифратор – это электронная схема, у которой «1» появляется на том выходе, номер которого совпадает с числом, соответствующим двоичной записи, поданной на вход $i = S_0 \cdot 2^0 + S_1 \cdot 2^1 + S_2 \cdot 2^2$. Таким образом, дешифратор прибавляет единицу в позиции, соответствующей ошибке и, тем самым, исправляет ошибочно принятый символ.

Если в канале произойдут 2 ошибки, то синдром окажется равным двоичной сумме двух столбцов, соответствующих позициям ошибок. При этом дешифратор произведет «исправление» ошибки в позиции, соответствующей этому суммарному столбцу, то есть внесет третью ошибку.

7.5. Прочие методы построения корректирующих кодов

В предыдущих параграфах были описаны принципы построения линейных корректирующих кодов и рассмотрены простейшие коды. В настоящее время разработано и широко применяется большое число линейных кодов, исправляющих многократные ошибки. Наиболее общим классом линейных кодов, позволяющим строить коды с заданной корректирующей способностью, являются коды Боуза-Чоудхури-Хоквингема (БЧХ)[1-§7.3].

Обсудим, в чем преимущество линейных кодов. Благодаря тому, что мы наделили код определенной математической структурой, его задание ограничилось заданием проверочной матрицы размера $n \times k$, в отличие от задания таблицы кодирования размера 2×2^k (см. параграф 7.1). Проверочная матрица требует значительно меньшего объема памяти, что весьма существенно при больших k . Одновременно, процедура кодирования также требует меньшего числа операций. Развитый математический аппарат позволяет строить реализуемые в реальном времени алгоритмы кодирования и декодирования.

Дальнейшее упрощение задания линейного кода, его кодирования и декодирования связано с использованием алгебраического аппарата полей Галуа. Его использование позволяет задавать код с помощью одного многочлена степени k , то есть проще, чем в рассмотренном случае задания кода посредством проверочной матрицы. Упрощается также реализация кодирования и декодирования. Эти коды являются так называемыми циклическими кодами, то есть кодами, в которых циклический сдвиг любого кодового слова также является кодовым словом. Коды БЧХ также могут быть реализованы как циклические коды.

Помимо блоковых кодов существует множество методов построения помехоустойчивых последовательностей, в которых путем введения избыточных символов обеспечивается возможность обнаружения или исправления ошибок [4-Гл.1]. В этих методах также используется разнообразный и сложный математический аппарат, вследствие чего достаточно простое и, вместе с

тем, содержательное описание этих методов представляется затруднительным. Поэтому ознакомление с этими кодами возможно только в рамках специального курса по теории кодирования.

Заключение

Первая часть нашего курса завершена. Знания, полученные в процессе изучения материала данного пособия, являются фундаментом для дальнейшего изучения современной теории связи. Введены основные понятия, без знания и понимания которых невозможна профессиональная деятельность в области телекоммуникаций. Рассмотрена общая схема системы связи, обсуждены ее работа и роль ее звеньев. Отмечена центральная задача теории связи – это борьба с помехами, влияние которых приводит к потере части переданной информации.

Помехи принципиально невозможно полностью исключить из системы связи, поэтому после всех мер по уменьшению мощности помех, действующих в системе, осуществляются меры по уменьшению их влияния на верность передачи информации. Главным результатом теории является вывод о том, что мерой **поумень-**шению влияния помех является введение избыточности в передаваемый сигнал. Это означает, что в передаваемый сигнал добавляется дополнительная информация. Ее назначение – помочь полностью восстановить на приёмном конце переданную информацию, частично потерянную из-за воздействия шумов. Задача теории связи состоит в том, чтобы показать, как именно преобразовать передаваемый сигнал, чтобы достичь требуемого результата. В этом контексте мы рассмотрели понятие обмена параметров сигнала. Было отмечено, что формирование помехоустойчивого сигнала принципиально связано с обменом динамического диапазона на ширину спектра. Исторически первым примером такого обмена является частотная модуляция. Современные методы известны под названием «методы расширенного спектра». Принцип исправления сигнала, пораженного шумом, наиболее прозрачно виден на примере корректирующих кодов, рассмотренных в последней главе настоящего пособия.

Мы рассмотрели главные этапы обработки сигналов при их передаче от источника к потребителю. Это формирование первичного электрического сигнала, формирование сигнала для линии, передача по линии, восстановление переданного сигнала на приемном конце. Предполагается, что преобразование сигнала на каждом этапе является обратимым, что является гарантией того, что информация не теряется. Если речь идет о цифровой передаче, то дополнительно производится преобразование аналогового сигнала в цифровую форму. Данное преобразование не является обратимым, восстановление сигнала всегда имеет погрешность, обусловленную параметрами оцифрования. Однако погрешность **всегда можно** установить в допустимых пределах выбором этих параметров. Преимущество цифровой связи в том, что цифровой сигнал значительно меньше подвержен действию помех.

В настоящем пособии мы ограничились рассмотрением детерминированных сигналов – случайные сигналы будут предметом рассмотрения второй части курса. Случайные сигналы являются **сутью теорией связи**. Именно случайным характером обусловлено информационное содержание сигналов и деструктивное влияние помех. Вероятностные характеристики позволяют определить количество информации в сигнале, оценить оптимальность восстановления сигнала из помех.

Автор надеется, что основа, данная в рамках первой части, позволит лучше усвоить вторую часть, посвященную основам вероятностных методов теории связи.

Список рекомендуемой литературы по разделам

К разделу 1 (введению).

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 1. СС. 10–27.

К разделу 2.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 2. СС. 29–44

Баскаков С.И. «Радиотехнические цепи и сигналы», «ВШ». М.: 1988. 448 с.

Главы 1–3. СС. 11–90.

Глава 8, §8.2. СС. 193–200.

К разделу 3, §3.1, 3.2.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 3, §3.2, 3.3. СС. 88–102.

Баскаков С.И. «Радиотехнические цепи и сигналы», «ВШ». М.: 1988. 448 с.

Глава 4, §4.1, 4.2. СС. 92–108.

К разделу 3, §3.3.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 9, §9.2. СС. 353–355.

К разделу 4. §4.1, 4.2.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 2, §2.4. СС. 44–49.

Глава 3, §3.6. СС. 112–115.

Баскаков С.И. «Радиотехнические цепи и сигналы», «ВШ». М.: 1988. 448 с.

Глава 5, §5.1, 5.2. СС. 119–127.

Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.

Глава 3, §3.1. СС. 111–115.

К разделу 4. §4.3, 4.4.

Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.

Глава 3, §3.2.1–3.2.4. СС. 115–125.

К разделу 4. §4.5.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 9, §9.2. СС. 353–355.

Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.

Глава 4, §4.6. СС. 218–228.

К разделу 5.

Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.

Глава 4, §4.1–4.3. СС. 173–200.

К разделу 6.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.

Глава 3, §3.5. СС. 103–112.

Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.

Глава 6, §6.1.1–6.1.5. СС. 283–312.

К разделу 7.

Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи: Учебник для вузов / М.: Радио и связь, 1999. 432с.

Гл.7, §7.3.1-7.3.5. СС. 263–289.

У. Питерсон, Э. Уэлдон. Коды, исправляющие ошибки. Пер. с англ. / Под ред. Р.Л. Добрушина и С.И. Самойленко / «МИР», М., 1976. 593с.

Гл.1, §1.2. СС. 11–13, §2.4. СС. 15–19. Гл.2, Гл.3, §3.1, 3.2. СС. 52–59, §5.1. СС. 134–138, Гл.9.

Основная литература

1. *Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В.* Теория электрической связи: Учебник для вузов/ М.: Радио и связь, 1999. 432с.
2. *Баскаков С.И.* «Радиотехнические цепи сигналы», «ВШ». М.: 1988. 448 с.
3. *Беллами Дж.* Цифровая телефония: Пер. с англ. / Под ред. Ф.Н. Берлина, Ю.Н. Чернышева. М.: Эко-Трендз, 2004. 640с.
4. *Питерсон У., Уэлдон Э.* Коды, исправляющие ошибки. Пер. с англ. Под ред. Р.Л. Добрушина и С.И. Самойленко / «МИР», М., 1976. 593с.

Дополнительная литература

1. *Зуев Н.* История развития средств связи до открытия электричества. <http://www.istorya.ru/referat/6198/1.php>
2. *Иванов М.Т., Сергиенко А.Б., Ушаков В.Н.* Теоретические основы радиотехники: Учеб. пособие / Под ред. В.Н. Ушакова. М.: Высш. шк., 2002. 306 с.: ил.
3. *Юкио Сато.* «Обработка сигналов. Первое знакомство», 1999.
4. *Simon Haykin, Michael Moher:* “Introduction to Analog and Digital Communications” JOHN WILEY & SONS, INC. 2007.
5. *Возенкрафт Дж., Джекобс И.* Теоретические основы техники связи. М., Мир, 1969. 630 с.
6. *Прокис Джон.* Цифровая связь. Пер. с англ. / Под ред. Д.Д. Кловского. М.: Радио и связь. 2000. 800 с.

АРА ВАГРАМОВИЧ ДАРЬЯН

ОБЩАЯ ТЕОРИЯ СВЯЗИ

Курс лекций

Часть первая

Главный редактор – М.Э. Авакян

Редактор – Э.А. Рухкян

Компьютерная верстка – А.С. Бжсикян

Адрес Редакции научных изданий
Российско-Армянского
университета:

*0051, г. Ереван, ул. Овсена Эмина, 123
тел/факс: (+374 10) 27-70-52 (внутр. 42-02)
e-mail: redaction.rau@gmail.ru*

Заказ № 6

Подписано к печати 22.04.2017г.

Формат 60x84¹/₁₆. Бумага офсетная № 1.

Объем усл. 8. 07 п.л. Тираж 200 экз.