

Л.Л.Клюев

Теория электрической связи

*Допущено Министерством образования
Республики Беларусь в качестве учебного
пособия для студентов специальности
"Телекоммуникационные системы" высших
учебных заведений*

Минск
Издательство "Дизайн ПРО"
1998

УДК 621.391(075)

ББК 31.264я7

К52

Рецензенты: кафедра "Системы передачи информации
"Белорусского государственного университета
транспорта, зав. кафедрой М.С. Костенок; д.т.н.,
профессор Я.В.Алишев

Клюев Л.Л.

**К52 Теория электрической связи. — Мн.: Дизайн ПРО, 1998, — 336 с.: ил.
ISBN 985-6182-71-9**

Изложены важнейшие вопросы теории электрической связи: освещены принципы построения различных систем электросвязи; дан анализ сообщений, сигналов и помех и их математических моделей; рассмотрены методы формирования и преобразования сигналов в каналах систем электросвязи. Значительное внимание в книге уделено: теории передачи и кодированию сообщений; теории помехоустойчивости систем электросвязи; оценке эффективности систем. Подробно рассмотрены цифровые способы передачи и методы разделения сигналов.

Для студентов высших учебных заведений, слушателей курсов повышения квалификации, специалистов народного хозяйства.

Гос.заказ

УДК 621.391(075)

ББК 31.264.я7

ISBN 985-6182-71-9

© Клюев Л.Л., 1998

ПРЕДИСЛОВИЕ

Настоящая книга является учебным пособием по курсу "Теория электрической связи". Она предназначена для студентов высших учебных заведений связи, обучающихся по специальности "Телекоммуникационные системы".

Значительное внимание в книге уделено физическому толкованию понятий и математическим выкладкам. Объем сведений из высшей математики, читаемого на первом и втором курсах ВУЗов, достаточен для понимания излагаемого материала. Усвоению материала способствует включение в некоторые главы ряда задач. Математический анализ доведен до соотношений, позволяющих производить инженерные расчеты. Для закрепления материала и контроля за его усвоением каждая глава содержит контрольные вопросы.

Автор выражает благодарность ректору Высшего колледжа связи, профессору Баркуну М.А. и сотрудникам колледжа: Прановичу В.Г., Мартюченко Н.И. за внимание и оказанную помощь автору при подготовке книги к изданию. Автор благодарит коллектив кафедры систем передачи информации Белорусского государственного университета транспорта, заведующего кафедрой к.т.н. доцента Костенка М.С., профессора Семенюту Н.Ф., д.т.н., профессора Белорусского государственного университета информатики и радиоэлектроники Алишева Я.В. за обсуждение рукописи и высказанные замечания, которые способствовали ее улучшению.

Автор надеется, что данное пособие облегчит работу студентам при изучении курса "Теория электрической связи", а также будет полезно при изучении специальных дисциплин.

ВВЕДЕНИЕ. Краткая история развития теории и техники электрической связи

Истоки развития теории и техники электросвязи (далее в тексте — связи) связаны с открытиями, которые были сделаны в начале XIX века.

В 1800 г. А.Вольта создал химический источник тока. Открытие, сделанное в 1820 г. Г. Х. Эрстедом, касалось способности электрического тока отклонять магнитную стрелку. В том же году А.М.Ампер предложил использовать магнитное действие электрического тока для передачи по проводам буквенного текста. В 1831 г. М.Фарадей открыл закон электромагнитной индукции, устанавливающий связь между электрическими и магнитными возмущениями. С.Б.Якоби в 1839 г. предложил пишущий телеграфный аппарат с записью сигнала в виде ломаной линии. С.Морзе создал в 1844 г. конструкцию более совершенного телеграфного аппарата и предложил телеграфную азбуку, в которой буквы обозначались комбинацией точек и тире. Эта азбука используется и до сих пор.

В России первая линия связи с пишущим телеграфным аппаратом была введена в 1841 г. между Зимним дворцом и Главным штабом в Петербурге, а в США — в 1844 г. между Вашингтоном и Балтимором.

Развитие телеграфа послужило толчком к появлению телефона. Первый опыт использования электричества для передачи речи на расстояние относится к 1837 г. Спустя 40 лет А.Г.Белл запатентовал телефон. Для передачи речи вначале использовались телеграфные линии, но качество принимаемых сообщений было низким. Поэтому началось строительство специальных двухпроводных телефонных линий. Одна из таких линий была построена в 1898 г. между Петербургом и Москвой.

Начало внедрения телефонной связи в нашей республике относится к 1891 году, когда в г. Минске был установлен первый телефон. Строительство телефонной сети общего пользования началось в 1896 г.

Теоретические основы радиосвязи были разработаны Д.К.Максвеллом. В его работе "Трактат об электричестве и магнетизме" опубликованной в 1873 г., впервые сформулирован важный для практического использования вывод: любой проводник с переменным током излучает в пространство электромагнитные волны, которые распространяются со скоростью света. Экспериментально это было подтверждено Г.Герцом в 1880 г.

Впервые передача сообщения на расстояние с помощью электромагнитных волн была продемонстрирована в 1895 г. А.С.Поповым при помощи разработанного для этой цели приемника. Позднее, в 1899 г. им была организована радиосвязь на линии протяженностью в 52 км.

Увеличение дальности связи и повышение ее качества стало возможным благодаря изобретению диода А.Флемингом (1904 г.) и триода Л.Д.Форестом (1906 г.). Эти изобретения положили начало развитию электроники.

Первая в мире радиовещательная станция мощностью 12 кВт была построена в 1922 г. в Москве. Позднее была построена радиостанция РВ-10 в г.Минске. Радиопередачи на этой радиостанции начались в 1925 г.

Развитие радио стимулировало рождение нового средства для передачи информации — телевидения. Первые прототипы телевизионной техники, в которых были реализованы два конкурентных принципа —

механический и электронный — появились в двадцатых годах XX века. Механический не выдержал испытание временем, а первые экспериментальные телевизионные передачи о применении электрического способа начались в 1929 г. Через семь лет на Би-би-си в Англии организуется постоянная телевизионная служба. Телевидение в нашей республике развивается с 1955 г.

Рождение ЭВМ, благодаря работам Дж.Сибица и Д.Мовгли в середине 40-х годов, оказало существенное влияние на развитие систем передачи информации.

ЭВМ и микропроцессоры становятся неотъемлемой частью различных сетей и систем, обеспечивая их требуемые качественные показатели и удобство обслуживания.

Запуск первого искусственного спутника Земли 4 октября 1957 г. открывает новую страницу в развитии теории и техники связи. На околоземные орбиты начали выводить специальные спутники с аппаратурой, обеспечивающей: передачу информации на большие расстояния (телевизионных программ, многоканальных телефонных сигналов, сигналов передачи данных и т.д.); охват значительных территорий вещательными и телепрограммами; связь между абонентами, находящимися в любой точке Земного шара.

1958 год ознаменовался изобретением лазера. С появлением первых промышленных образцов лазера (1961 г.) делаются попытки его внедрение в системы связи. Успех сопутствовал К.С.Као и сотрудникам из лаборатории телекоммуникационных стандартов в Харлоу (Англия). Они предложили для передачи светового сигнала использовать длинные оптические волокна, что послужило началом развития теории и техники волоконно-оптической связи. Применение лазера и направляющей среды — оптического волокна позволило существенно увеличить скорость и достоверность передачи информации, а связь сделать более надежной.

Теория электрической связи — интенсивно развивающаяся область науки. Фундаментом этой науки являются работы выдающихся ученых. Имена некоторых из них названы выше.

Значительный вклад в современную теорию связи внесли также А.Н.Колмогоров и А.Я.Хинчин (разработали основы математического аппарата); В.А.Котельников (создал теоретические основы цифровых систем передачи информации, разработал теорию потенциальной помехоустойчивости); К.Э.Шеннон (создал теорию кодирования для повышения качества передачи сообщений).

Дальнейшее развитие теория связи получила в работах А.А.Харкевича, В.И.Сифорова, Х.Найквиста, Р.Хартли, Н.Винера, Ф.М.Вудворда, Я.З.Ципкина и др. ученых.

Глава 1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ СВЯЗИ

1.1. Общие сведения о системах связи.

Информация, сообщение, сигнал

В данной книге под словом "*информация*" будем понимать сведения, которые описывают то или иное событие или свойство какого-либо объекта. Когда эти сведения становятся объектом передачи, распределения, преобразования, хранения или непосредственного использования, то они могут быть представлены в виде символов (знаков) или в каком либо другом виде, например, в виде фотографии, рисунка, киноплетки, произнесенных слов, выражающих мысль говорящего и т.д.

Сообщение является формой представления информации. Одно и то же сообщение может быть представлено в различной форме. Например, сведения о Вашем приезде можно сообщить по телефону или телеграммой. Сообщение, которое передается телеграммой, кроме информационной части, из которой получателю становится известным день и время приезда, номер поезда и вагона, содержит также сведения, позволяющие доставить это сообщение в нужное время и по указанному адресу.

Телефонное сообщение с той же информацией содержит, кроме речи, сведения для установления и разъединения тракта передачи.

Сообщение, передаваемое по сети связи, таким образом содержит информационную часть и дополнительную (вспомогательную) часть, позволяющую доставить сообщение.

Различают *аналоговые* и *дискретные* сообщения. Аналоговое сообщение например, разговорная речь описывается функцией времени и непрерывным множеством возможных значений. Источник, на выходе которого вырабатывается непрерывное сообщение, называется *непрерывным*. Дискретное сообщение, например, телеграмма или данные электронно-вычислительной машины (ЭВМ) является функцией дискретного или непрерывного времени и принимает конечное число возможных значений. Источник на выходе которого вырабатывается дискретное сообщение, называется *дискретным*.

Передача сообщений на расстояние осуществляется с помощью какого-либо материального носителя (бумаги, перфоленты, магнитной ленты и т.д.) или физического явления (электрического тока, звукового, электромагнитного поля и т.д.). Физический процесс (колебания электрического тока или напряжения, звуковые и электромагнитные волны), один или несколько параметров

которого изменяются случайно в соответствии с передаваемым случайным сообщением или детерминированно (по заранее заданному закону), называется *сигналом*. Сигнал, предназначенный для передачи информации, является формой представления этой информации. В канале связи эта форма, как правило, не остается постоянной, так как сигнал подвергается различным преобразованиям. Преобразование одного (преобразуемого) сигнала в другой (преобразованный) путем изменения параметров третьего сигнала переносчика называется *модуляцией*. Особенностью сигнала переносчика является постоянство его параметров до передачи и их изменение в соответствии с сообщением при передаче. Сигнал переносчик в режиме передачи является преобразованным (*модулированным*) сигналом. Преобразуемый сигнал называется *модулирующим*, а сигнал переносчик — *модулируемым*. Поскольку сообщение является случайным по своей природе, т.е. непредсказуемым для получателя, то как модулирующий, так и модулированный сигналы являются *случайными*.

Модулирующий сигнал, полученный непосредственным преобразованием сообщения, называется *первичным*. Примером первичного сигнала является электрический ток на выходе микрофона. Преобразованный сигнал на выходе модулятора по отношению к преобразованному первичному сигналу называется *вторичным*.

В системах связи, кроме случайных, находят применение также неслучайные (*детерминированные*) сигналы, используемые для настройки, управления или контроля. Эти сигналы являются физическими процессами с постоянными или с изменяющимися по известному закону параметрами. По этой причине они никакой информации не несут. Однако, при прохождении по каналам связи принимаемые сигналы могут стать носителями информации. Эту информацию можно извлечь, измеряя отклонения параметров принимаемого сигнала от параметров передаваемого. Они возникают из-за каких-то нарушений в работе аппаратуры или изменений показателей качества и параметров системы связи.

Параметр сигнала переносчика, отображающий изменения передаваемого сообщения, называется *представляющим* (информационным). Для гармонического колебания это амплитуда, фаза или частота.

Сигнал, у которого представляющий параметр описывается функцией непрерывного или дискретного времени и непрерывным множеством возможных значений, называется *аналоговым*.

Сигнал, у которого представляющий параметр описывается функцией непрерывного времени и конечным множеством возможных значений, называется *дискретным*.

Дискретный сигнал, у которого представляющий параметр описывается функцией дискретного времени и конечным множеством возможных значений, называется *цифровым*. Если возможные значения представлены совокупностью двоичных чисел, то цифровой сигнал называется *двоичным*. Двоичные цифровые сигналы являются самыми простыми и по этой причине используются чаще других сигналов. Кроме двоичных, в системах связи применяются троичные, четверичные, восьмеричные и т.д. n -ичные цифровые сигналы.

Сигналы в системах связи являются средством для транспортировки сообщений, а техника связи по существу является техникой транспортирования (передачи) сообщений.

Основными параметрами сигнала с точки зрения транспортировки являются: длительность T_c , ширина спектра F_c , динамический диапазон D_c .

Длительность сигнала определяет интервал времени, в котором этот сигнал существует. *Ширина спектра* сигнала — это интервал (диапазон) частот, в пределах которого сосредоточена основная часть энергии сигнала. Для сигнала, определенного функцией времени $u(t)$, его *энергия* равна

$$E = \int_0^{T_c} u^2(t) dt. \quad (1.1)$$

Динамический диапазон — это отношение наибольшей мгновенной мощности сигнала к наименьшей. Динамический диапазон обычно выражают в децибелах:

$$D_c = 10 \lg \frac{P_{\max}}{P_{\min}}, \quad (1.2)$$

где $P_{\max}$, $P_{\min}$ — максимальное и минимальное значения мгновенной мощности.

Обобщающей характеристикой является объем сигнала, равный произведению его длительности, ширины спектра и динамического диапазона

$$V_c = T_c F_c D_c. \quad (1.3)$$

1.2. Связь, сеть связи, система связи

Всякое средство, позволяющее корреспонденту доставлять одному или нескольким определенным (в телефонии и телеграфии) или возможным (в радиовещании) корреспондентам информацию любого рода в любой используемой форме (письменный или печатный документ, неподвижное или подвижное изображение, речь, музыка, видимые или слышимые сигналы, сигналы управления механизмами и т.д.) с использованием любой систе-

мы передачи (проводной, радио, оптической и т.д.) называется *электросвязью* (связью). Связь классифицируется по видам в зависимости от характера передаваемых сообщений или среды распространения сигналов (рис.1.1).



Рис.1.1. Виды электросвязи

Совокупность узлов и трактов, позволяющая соединять два и более определенных пункта для обеспечения связи между ними, называется *сетью связи*. Здесь под пунктами следует понимать источники и получатели (потребители) информации, а под трактом — непрерывный путь связи между двумя пунктами. Узел — это промежуточный пункт на сети, в котором возможно соединение нужных входов и выходов.

Система, посредством которой источник может достаточно эффективно и надежно передавать информацию в пункт назначения, называется *системой связи*. Система связи, таким образом, является подсистемой сети связи. Обобщенная структурная схема системы электросвязи приведена на рис.1.2.

В состав системы связи входят передатчик 1, *среда распространения* 2 и приемник 3. Источник сообщения 4 и получатель сообщения 5 в состав системы не входят. Передатчик является устройством, которое преобразует сообщение в сигнал. Вначале сообщение преобразуется в первичный сигнал. Если этот сигнал возможно передать через линию связи, включающую соответствующее оборудование и среду распространения, то он без каких-

либо дополнительных преобразований поступает к приемнику 3. Часто такая передача оказывается невозможной, и первичный сигнал посредством модуляции преобразуется во вторичный, параметры которого оказываются согласованными с линией связи.

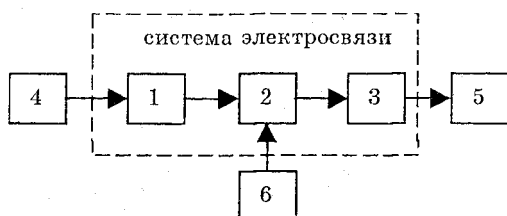


Рис.1.2. Обобщенная структурная схема системы связи: 1 — передатчик; 2 — линия связи; 3 — приемник; 4 — источник сообщения; 5 — получатель сообщения; 6 — источник помех

Среда распространения 2 — это физическая цепь, которая содержит одну или две пары токопроводящих металлических проводов или светопроводящих стеклянных волокон. В качестве среды распространения используется также открытое пространство. Соответственно системы связи называются *проводными*, *радио* или *оптическими*.

Приемник является устройством, с помощью которого осуществляется восстановление сообщения по принятому сигналу. Преобразования, выполняемые в приемнике, обратны преобразованиям в передатчике. Это позволяет сохранить информацию, содержащуюся в сообщении при ее передаче. Сигналы на входе приемника отличаются от переданных из-за *искажений* и *помех*, что является причиной воспроизведения сообщения с некоторой погрешностью (ошибкой). Искажения возникают из-за несовершенства аппаратуры связи и отклонений характеристик каналов связи от требуемых. Эти искажения в принципе могут быть устранены соответствующим конструированием аппаратуры связи и коррекцией характеристик каналов. Помехи заранее неизвестны, поэтому их влияние полностью устранить невозможно.

Помехи, действующие на выходе источника помех 6, подразделяются на *внешние* и *внутренние*. К внешним относятся космические, атмосферные, промышленные, помехи от посторонних радиостанций и каналов. Внутренние помехи возникают в аппаратуре систем связи. Они обусловлены хаотическим движением носителей заряда в резисторах, электронных и других приборах, взаимным влиянием друг на друга каналов в многоканальных системах связи.

По типу передаваемых сообщений системы связи делятся на *аналоговые* и *дискретные (цифровые)*. Аналоговые системы предназначены для передачи непрерывных, а цифровые — для передачи цифровых сигналов. Используются также и *смешанные*

системы, в которых непрерывные сообщения преобразуются в дискретные с целью использования цифровых методов передачи.

По назначению различают следующие типы систем связи: *телефонные, телеграфные, фототелеграфные, телевизионные телеметрические, телеуправления, системы передачи данных* и т.д.

По режиму использования системы связи делятся на *симплексные* и *дуплексные*. Система связи называется *многоканальной*, если она обеспечивает передачу нескольких сообщений по одной общей среде распространения.

Система связи называется *автоматизированной*, если она обеспечивает связь в основном автоматическими устройствами.

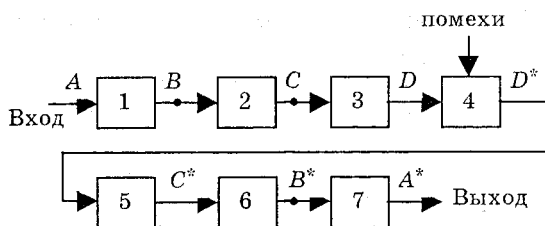


Рис.1.3. Структурная схема цифровой системы передачи: 1 — кодер источника (блок эффективного кодирования);

2 — кодер канала (блок помехоустойчивого кодирования); 3 — модулятор; 4 — линия связи; 5 — демодулятор; 6 — декодер канала (блок помехоустойчивого декодирования);

7 — декодер источника (блок эффективного декодирования)

Понятие *канал связи* поясним с использованием структурной схемы цифровой системы связи (рис.1.3). В данной системе кроме пути прохождения сигнала от входа (точка A) до выхода (точка A*), могут быть выделены также другие отрезки пути: от точки A до точки D*, от точки B до точки B* и т.д.

Соответствующие этим отрезкам технические средства и среда распространения

являются *каналами связи*. В зависимости от характера сигналов на входе и выходе каналы делятся на *дискретные, дискретно-непрерывные* и *непрерывные*. Дискретные каналы на рис.1.3 обеспечивают прохождение сигналов по путям A, A* и B, B*. Сигналы на входах и выходах этих каналов дискретные. Дискретно-непрерывным является канал, соответствующий отрезку пути между точками A, D*. Сигнал на входе этого канала дискретный, а на выходе — непрерывный. Непрерывным каналом является линия связи (между точками D, D*). На его входе и выходе сигналы непрерывные.

По характеру применения каналы связи делятся на: *телефонные, телевизионные, телеграфные, звукового вещания, передачи данных* и другие. В зависимости от передачи сигнала по

направляющим линиям или в свободном пространстве, каналы называются *каналами связи* или *радиоканалами*. К каналам связи относятся проводные (воздушные и кабельные) каналы, работающие на частотах до тысяч килогерц, а также волоконно-оптические, работающие в оптическом диапазоне волн (10^{13} - 10^{15} Гц).

К радиоканалам относятся каналы связи, работающие в диапазоне волн от сверхдлинных (3 кГц) до децимиллиметровых (3000 ГГц).^{*}

Каналы связи, также как и сигнал, характеризуются тремя параметрами: *временем передачи* T_k , *полосой пропускания* F_k и *динамическим диапазоном* D_k .

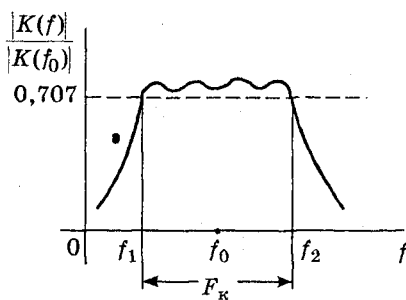


Рис.1.4. Нормированная амплитудно-частотная характеристика канала

Время передачи T_k — это время, в течении которого ведется передача информации. *Полоса пропускания* канала F_k — это диапазон частот, в пределах которого значение нормированной амплитудно-частотной характеристики (АЧХ) канала больше уровня, равного 0,707 (Рис.1.4). *Динамический диапазон* D_k определяется отношением допустимой максимальной мощности передаваемого сигнала к мощности помехи, присутствующей в канале

$$D_k = 10 \lg \frac{P_{\max}}{P_{\Pi}}. \quad (1.4)$$

Обобщающей характеристикой канала связи является его объем

$$V_k = T_k F_k D_k. \quad (1.5)$$

Для неискаженной передачи сигнала по каналу связи необходимо, чтобы объем сигнала был меньше или по крайней мере равен объему канала. Это условие называют *условием согласования сигнала с каналом* и записывают в виде

$$V_c \leq V_k. \quad (1.6)$$

1.3. Кодирование и модуляция

Кодированием и модуляцией достигается согласование источника сообщений с каналом связи.

Рассмотрим источник дискретного сообщения. Пусть сообщение выражено при помощи некоторых символов, например,

букв печатного текста. Если далее эти буквы преобразуются в другие символы или последовательности из них по заранее установленному и известному при передаче и приеме правилу, причем преобразование является взаимнооднозначным, то говорят, что сообщение дискретного источника кодируется.

Набор символов (букв), над которым определена последовательность символов на выходе дискретного источника, называется *алфавитом*. Число символов в алфавите называется *объемом алфавита*. Рассмотрим пример.

Пример 1.1. Для набора (алфавита) букв a_1, a_2, a_3, a_4 объем равен 4. Определим правило, по которому каждой букве a_i , $i = 1, 2, 3, 4$ будет соответствовать последовательность b_i над алфавитом объемом 2. Символы этого алфавита 0 и 1. Такой алфавит называется двоичным. Пусть это правило следующее:

a_i	b_i
a_1	00
a_2	01
a_3	10
a_4	11

В этом случае, например, дискретной последовательности $a_1 a_2 a_3$ соответствует кодовая последовательность 001011.

Отрезок кодовой последовательности (комбинации) называется *блоком*. Число символов (элементов) в блоке называется *значностью* или *длиной* блока.

Блок, выделенный по определенному признаку и рассматриваемый как одно целое, называется *словом*. Словом, например, является кодовая последовательность (комбинация). Совокупность (множество) всех кодовых комбинаций, возможных при заданном правиле кодирования, образует *код*.

Кодирование источника непрерывных сообщений не отличается от кодирования дискретного, если непрерывное сообщение преобразуется в дискретное. Это преобразование включает дискретизацию, при которой непрерывное во времени сообщение преобразуется в последовательность отсчетов. Далее каждый отсчет сравнивается с заранее заданными уровнями квантования. Ближайший к отсчету уровень переводится в число. Набору квантованных уровней соответствует набор символов (букв), а объем алфавита равен числу квантованных уровней.

Кодовая последовательность $b_{n-1}b_{n-2}...b_2b_1b_0$, где $b_{n-1}, b_{n-2}, ..., b_2, b_1, b_0$ разряды числа, начиная со старшего, определяет число в системе счисления с основанием m

$$M = b_{n-1}m^{n-1} + b_{n-2}m^{n-2} + ... + b_1m^1 + b_0m^0. \quad (1.7)$$

Коэффициенты b_i , $i = 0, 1, 2, \dots, n-1$, принимают значения от 0 до $m-1$ и образуют кодовый алфавит. Число кодовых комбинаций равно m^n .

В современных системах связи нашла широкое применение *двоичная система счисления* с основанием $m = 2$. Коэффициенты b_i при $m = 2$ принимают значения 0 и 1 и могут передаваться отрезками колебаний различных частот, отрезками колебаний с отличающимися начальными фазами или постоянным током с разной амплитудой или полярностью. Это обеспечивает простоту аппаратуры.

Блоки системы связи, в которых осуществляется кодирование, называются *кодерами*.

Коды, у которых все кодовые комбинации имеют одинаковую длину, называются *равномерными*. *Неравномерные* коды содержат кодовые комбинации неравной длины.

Одной из задач кодирования является задача согласования алфавита, из которого построены дискретные сообщения с кодовым алфавитом выходных комбинаций. Примером решения этой задачи является международный телеграфный код №2 (МТК-2). В этом коде передача 32 букв русского алфавита осуществляется двоичными комбинациями длиной в 5 символов. Код МТК-2 является равномерным.

Кодирование позволяет решить также задачу устранения избыточности источника сообщения и тем самым повысить скорость передаваемой информации. Примерами кодов, решающими эту задачу (задачу эффективного кодирования), являются коды Шеннона-Фано и Хаффмена. Эти коды являются неравномерными. Для повышения достоверности передачи дискретной информации применяют *помехоустойчивые коды*. В отличие от простых, у которых все возможные кодовые комбинации используются для передачи информации, помехоустойчивые коды строятся так, что для передачи используется лишь часть от общего числа возможных комбинаций. Благодаря этому появляется возможность обнаруживать и исправлять ошибки. Способность кода обнаруживать и исправлять ошибки достигается также и тем, что, наряду с информационными, по каналу связи передаются *проверочные символы*.

В системах связи применяется дискретная и непрерывная модуляция. Под *дискретной модуляцией* понимают такой процесс формирования дискретного сигнала, при котором каждому символу кодовой комбинации ставится в однозначное соответствие форма единичного элемента. Примерами форм единичных элементов сигнала являются импульсы постоянного тока заданной длительности с амплитудой, неравной или равной нулю, или им-

пульсы переменного тока заданной длительности с отличающейся частотой.

Под *непрерывной модуляцией* понимают процесс преобразования непрерывного сигнала в другой непрерывный сигнал путем изменения параметров сигнала носителя сообщения (переносчика) в соответствии с преобразуемым сигналом.

Устройства, с помощью которых осуществляется модуляция, называются *модуляторами*. Они служат для согласования параметров сигнала с параметрами линии связи. Используются следующие носители сообщения: постоянный ток, переменный ток (гармоническое колебание), периодическая последовательность импульсов, шумоподобный сигнал (ШС).

В системах связи применяют от одной до нескольких ступеней модуляции. На 1-й ступени сообщение преобразуется в первичный сигнал. На 2-й и последующих ступенях первичный или другой сигнал модулирует высокочастотную несущую (сигнал переносчик). Принципиальной разницы между первичной и последующей модуляциями нет. Отличие заключается только в значениях частот модулируемых колебаний.

Модуляцией первичным непрерывным сигналом одного из параметров гармонического сигнала получают амплитудную модуляцию (АМ), частотную модуляцию (ЧМ), фазовую модуляцию (ФМ). Воздействуя первичным непрерывным сигналом на один из параметров периодической последовательности видеоимпульсов, получают амплитудно-импульсную модуляцию (АИМ), широтно-импульсную модуляцию (ШИМ), фазо-импульсную (ФИМ) и частотно-импульсную (ЧИМ).

Модулируя один из параметров гармонического сигнала импульсами постоянного тока (двоичной кодовой комбинацией), можно получить амплитудно-модулированный (АМ), частотно-модулированный (ЧМ) или фазо-модулированный (ФМ) дискретный сигнал. Разновидностью дискретного ФМ сигнала является сигнал с относительной фазовой модуляцией (ОФМ). В цифровых системах связи в качестве носителя сообщения также используют и постоянный ток. Изменяя его параметры (величину или направление тока) в соответствии с двоичной кодовой комбинацией, получают сигналы передачи данных или телеграфные сигналы.

1.4. Демодуляция и декодирование

Демодуляция в системах связи — это процесс преобразования модулированного сигнала (обычно прошедшего канал и искаженного помехами) в модулирующий. Восстановление дискретного сообщения по выходному сигналу демодулятора, осуществляемое с учетом правил кодирования, называется

декодированием. Устройства, с помощью которых производится демодуляция, называются *демодуляторами*. Устройства, с помощью которых ведется декодирование, называются *декодерами*.

Прием сигналов ведется на фоне помех. Они ухудшают качество восстановленного приемником сообщения. Чтобы свести к минимуму влияние помех, смесь сигнала и помех, поступающая на вход приемника, подвергается определенным линейным и нелинейным преобразованиям, которые называют *обработкой сигнала*. Структура приемника, обеспечивающая наилучшее преобразование (с точки зрения заранее сформулированного критерия оптимальности) называется *оптимальной*, а приемник — *оптимальным*.

Характер оптимальной обработки зависит от свойств сигнала и помех, свойств канала, методов модуляции и кодирования, а также от критерия оптимальности.

При оптимальной обработке приемник обеспечивает наибольшую помехоустойчивость. Под *помехоустойчивостью* понимают способность системы связи противостоять вредному влиянию помех на передаваемое сообщение.

Демодуляция относится к числу основных операций, совершаемых приемником. На выходе аналогового демодулятора присутствует колебание, с некоторой погрешностью повторяющее первичный сигнал. Это колебание при последующей обработке преобразуется в сообщение.

В цифровых системах связи восстановление сообщения осуществляется чаще всего в два этапа. На первом этапе сигнал демодулируется и на выходе демодулятора выделяется кодовая комбинация. Если решение о каждом элементе кодовой комбинации принимается в результате анализа смеси сигнала и шума только в течение длительности этого элемента и без учета результатов анализа предыдущих элементов, то такой способ обработки называется *позлементным (посимвольным) приемом*. На втором этапе кодовая комбинация декодируется.

Возможен также и другой способ приема — *прием в целом*. При этом способе демодуляция и декодирование совмещаются и выполняются одним устройством, в котором отрезок смеси сигнала и помехи, соответствующий передаваемой кодовой комбинации, анализируется целиком, в отличие отazoleментного приема. Это позволяет реализовать оптимальную обработку сигнала, следовательно, и наибольшую помехоустойчивость. К сожалению, приемник, в котором используется прием в целом, оказывается во много раз сложнее приемника сazoleментным приемом. По этой причине прием в целом в системах связи, как правило, не применяется.

1.5. Основные характеристики систем связи

Система связи характеризуется совокупностью параметров. Те из них, которые связаны с качеством системы монотонной зависимостью, называются *показателями качества системы*. Чем больше (меньше) значение показателя качества, тем лучше (хуже) при прочих равных условиях система.

При проектировании системы принимают во внимание большое количество показателей качества и параметров в соответствии с заранее обоснованным критерием оптимальности. Наилучшей (оптимальной) считается такая система, которой соответствует наибольшее (наименьшее) значение некоторой *целевой функции* от показателей качества. Показатели качества и параметры систем связи условно делят на *информационные*, *техничко-экономические* и *техничко-эксплуатационные*. К информационным относятся *верность*, *скорость* и *задержка передачи информации*. К технико-экономическим — *стоимость*, *габаритные размеры*, *масса*. К технико-эксплуатационным — *среднее время безотказной работы*, *температурный диапазон работы* и т.д.

Выделим показатели, характеризующие систему связи с точки зрения передачи информации.

Помехоустойчивость является одним из основных показателей качества системы связи. Помехоустойчивость при заданной помехе характеризуется *верностью передачи* — степенью соответствия принятого сообщения переданному. При передаче непрерывных сообщений мерой верности служит среднеквадратичное отклонение $\bar{\varepsilon}^2$ между принятым $a^*(t)$ и переданным $a(t)$ сообщениями:

$$\bar{\varepsilon}^2 = \frac{1}{T} \int_0^T [a^*(t) - a(t)]^2 dt,$$

где T — время, в течении которого принимается сообщение.

Первичный сигнал $b(t)$ связан с сообщением $a(t)$ линейной зависимостью, т.е.

$$b(t) = ka(t),$$

где k — коэффициент преобразования.

Для простоты будем считать, что $k = 1$. С учетом этого формулу для среднеквадратичного отклонения запишем в виде:

$$\bar{\varepsilon}^2 = \frac{1}{T} \int_0^T [b^*(t) - b(t)]^2 dt. \quad (1.8)$$

Звездочкой обозначают оценку, которая отличается от передаваемого сигнала на величину погрешности.

Мерой верности может быть также вероятность того, что отклонение ε не превзойдет заранее заданную величину ε_0 :

$$P(|\varepsilon| \leq \varepsilon_0). \quad (1.9)$$

Чем меньше $\bar{\varepsilon}^2$ или больше вероятность (1.9), тем выше помехоустойчивость.

Другой важный показатель качества это *скорость передачи* R , определяемая количеством передаваемых двоичных символов в единицу времени. Для дискретного канала связи этот показатель

$$R = \frac{1}{T} \log_2 m, \quad (1.10)$$

где T — длительность посылки, m — основание кода.

Скорость передачи дискретной информации измеряется в битах в секунду (бит/с). Максимальная скорость передачи информации, которая обеспечивается соответствующим распределением вероятности выдачи символа на выходе дискретного источника, называется *пропускной способностью дискретного канала связи*. Аналогично определяются скорость передачи информации и пропускная способность непрерывного канала связи. Определение скорости и пропускной способности для непрерывного канала связи приведено в гл.6. Задержка передачи определяет время, прошедшее от момента начала передачи сообщения в передатчике до момента выдачи восстановленного сообщения на выходе приемника. Она зависит от протяженности канала связи и длительности обработок сигнала в передатчике и приемнике.

Контрольные вопросы

1. Что понимают под сообщением и сигналом?
2. Какие основные блоки входят в систему связи?
3. В чем различие между системой и каналом связи?
4. В чем различие и единство процессов кодирования и модуляции? Какова роль этих процессов в системах связи?
5. Какие процессы при декодировании и демодуляции Вы знаете?
6. Что Вы понимаете под оптимальным приемником?
7. Чем отличается поэлементный прием от приема в целом?
8. Какие характеристики систем связи Вам известны?
9. Что Вы понимаете под сетью связи?

Глава 2. ДЕТЕРМИНИРОВАННЫЕ СИГНАЛЫ И ИХ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

2.1. Разложение детерминированных функций в ортогональные ряды

Детерминированными называются такие функции, которые точно определены в любой момент времени. Как показано в гл.1, при передаче информации сигналы являются случайными процессами, поэтому детерминированные функции не могут быть их математическими моделями. Однако, изучение детерминированных функций полезно по ряду причин. Назовем некоторые из них.

1. Результаты анализа (функционального, спектрального и т.д.) детерминированных функций являются базовыми при изучении случайных функций.

2. Случайные функции в ряде случаев могут быть представлены такими математическими моделями, которые строятся с использованием множества случайных чисел и детерминированных функций (канонические разложения случайных функций). Канонические разложения позволяют существенно упростить решение некоторых задач (вычисление корреляционных функций, определение интегралов от случайных функций), так как операции над случайными функциями заменяются операциями над случайными числами и детерминированными функциями.

3. Применение детерминированных сигналов оказывается полезным также при решении задач анализа линейных и нелинейных узлов, блоков и систем, оценке искажений, возникающих при передаче сигналов по каналам связи и т.д.

4. Детерминированные сигналы имеют и самостоятельное значение. Они специально создаются для целей измерения, наладки и регулирования систем связи, выполняя роль эталонов.

Детерминированный сигнал выражается математически определенной функцией времени $u(t)$. Примером математического выражения детерминированного сигнала является косинусоида

$$u(t) = u_m \cos(\omega t + \varphi), \quad (2.1)$$

которая описывает колебание с постоянной амплитудой u_m , частотой ω и начальной фазой φ . При известной амплитуде, частоте и начальной фазе функция (2.1) однозначно определена для любого момента времени. Такая функция относится к *простым сигналам*.

В системах связи, кроме простых, используются также сложные детерминированные сигналы, которые удобно описывать совокупностью простых (элементарных) функций. Такие функции

называются *базисными*. Например, некоторый сигнал $u(t)$ можно представить в виде суммы произведений некоторых коэффициентов a_k на базисные функции φ_k :

$$u(t) = \sum_k a_k \varphi_k(t), \quad (2.2)$$

где a_k — либо простые, либо комплексные числа, а число " k " лежит в пределах от $-\infty$ до ∞ .

Разложение (2.2) справедливо для заданного конечного интервала времени. Условно этот интервал будем обозначать квадратными скобками с указанием начального и конечного значений времени: $[t_1, t_2]$.

При выбранном наборе базисных функций сигнал (2.2) полностью определяется совокупностью коэффициентов a_k . Такие совокупности называются *дискретными спектрами сигналов*. Базисные функции выбираются так, чтобы они имели простое аналитическое выражение, обеспечивали быструю сходимость ряда, позволяли легко вычислять коэффициенты a_k . Вычисление коэффициентов существенно облегчается, если в качестве базисных выбрать ортогональные функции. Систему комплексных функций $\varphi_k(t)$, $\varphi_l(t)$ называют *ортогональной* на отрезке времени $[t_1, t_2]$, если для всех значений $k = 0, \pm 1, \pm 2, \dots$ и $l = 0, \pm 1, \pm 2, \dots$ (за исключением случая $k = l$) удовлетворяется условие

$$\int_{t_1}^{t_2} \varphi_k(t) \varphi_l^*(t) dt = 0, \quad k \neq l, \quad (2.3)$$

где $\varphi_l^*(t)$ — комплексно сопряженная функция.

Система комплексных функций называется *нормированной*, если выполняется условие

$$\int_{t_1}^{t_2} \varphi_l^2(t) dt = 1, \quad \text{для } l = 0, \pm 1, \pm 2, \dots \quad (2.4)$$

Базисные функции, удовлетворяющие условиям (2.3) и (2.4), называются *ортонормированными*. Умножим левую, и правую части выражения (2.2) на $\varphi_l^*(t)$ и результат проинтегрируем

$$\int_{t_1}^{t_2} u(t) \varphi_l^*(t) dt = \sum_k a_k \int_{t_1}^{t_2} \varphi_l^*(t) \varphi_k(t) dt.$$

Учитывая выражения (2.3) и (2.4), получим

$$a_l = \int_{t_1}^{t_2} u(t) \varphi_l^*(t) dt, \quad l = k.$$

Коэффициенты a_k называют обобщенными коэффициентами Фурье, а ряд (2.2) — обобщенным рядом Фурье.

Если в качестве базисных выбраны периодические экспоненциальные функции

$$\varphi_k(t) = \frac{1}{\sqrt{T}} e^{jk\omega_1 t}, \quad k = 0, \pm 1, \pm 2, \dots,$$

то выражение (2.2) можно записать в виде

$$u(t) = \frac{1}{\sqrt{T}} \sum_{k=-\infty}^{\infty} a_k(jk\omega_1) e^{jk\omega_1 t}, \quad (2.5)$$

где $T = t_2 - t_1$ — интервал разложения сигнала, причем: $t_1 = -T/2$, $t_2 = T/2$, $\omega_1 = 2\pi/T$,

$$a_k(jk\omega_1) = \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} u(t) e^{-jk\omega_1 t} dt.$$

Выражение (2.5) преобразуем к виду

$$u(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \frac{2a_k(jk\omega_1)}{\sqrt{T}} e^{jk\omega_1 t} = \frac{1}{2} \sum_{k=-\infty}^{\infty} a(jk\omega_1) e^{jk\omega_1 t}, \quad (2.6)$$

где

$$a(jk\omega_1) = \frac{2}{T} \int_{-T/2}^{T/2} u(t) e^{-jk\omega_1 t} dt. \quad (2.7)$$

Выражение (2.6) известно как ряд Фурье, записанный в комплексной форме, а выражение (2.7) — как спектр сигнала $u(t)$.

От двустороннего спектрального представления можно перейти к одностороннему (не имеющему отрицательных частот). В этом случае получаем ряд Фурье в тригонометрической форме.

Преобразуем выражение (2.7), применяя формулу

$$e^{-jk\omega_1 t} = \cos k\omega_1 t - j \sin k\omega_1 t.$$

Получим:

$$a(jk\omega_1) = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cos k\omega_1 t dt - \frac{2}{T} j \int_{-T/2}^{T/2} u(t) \sin k\omega_1 t dt = a_k - jb_k, \quad (2.8)$$

$$\text{где } a_k = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cos k\omega_1 t dt, \quad b_k = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \sin k\omega_1 t dt.$$

Для чисел k , равных по величине, но разных по знаку, справедливы равенства

$$a_k = a_{-k}, \quad b_k = -b_{-k},$$

так как $\cos(k\omega_1 t) = \cos(-k\omega_1 t)$, $\sin(k\omega_1 t) = -\sin(-k\omega_1 t)$.

Модуль спектра сигнала характеризует *спектр амплитуд* этого сигнала:

$$|a(jk\omega_1)| = a(k\omega_1) = \sqrt{a_k^2 + b_k^2}. \quad (2.9)$$

Спектр фаз определяется выражением

$$\varphi(k\omega_1) = -\operatorname{arctg} \frac{b_k}{a_k}. \quad (2.10)$$

Подставим выражение (2.8) в (2.6). В результате получим *ряд Фурье в тригонометрической форме*:

$$\begin{aligned} u(t) &= \frac{1}{2} \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_1 t} - \frac{1}{2} j \sum_{k=-\infty}^{\infty} b_k e^{jk\omega_1 t} = \frac{a_0 - jb_0}{2} + \sum_{k=1}^{\infty} a_k \left(\frac{e^{jk\omega_1 t} + e^{-jk\omega_1 t}}{2} \right) + \\ &+ \sum_{k=1}^{\infty} b_k \left(\frac{e^{jk\omega_1 t} - e^{-jk\omega_1 t}}{2j} \right) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t) = \\ &= \frac{a_0}{2} + \sum_{k=1}^{\infty} a(k\omega_1) \cos[k\omega_1 t + \varphi(k\omega_1)]. \end{aligned} \quad (2.11)$$

Тригонометрические функции, определяющие ряд Фурье (2.11), являются известной ортонормированной системой. Возможно также представление сигналов по системе функций Хаара, по системам базисных многочленов Котельникова, Чебышева, Лаггера, Лежандра и др.

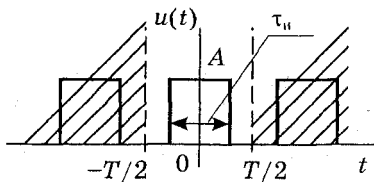


Рис.2.1. Последовательность прямоугольных импульсов

Пример 2.1. Вычислить коэффициенты ряда Фурье для последовательности прямоугольных импульсов, приведенной на рис.2.1. Интервал разложения равен $[-T/2, T/2]$.

Решение. Постоянная составляющая

$$\frac{a_0}{2} = \frac{1}{T} \int_{-T/2}^{T/2} A dt = \frac{A\tau_u}{T}.$$

Коэффициенты $b_k = 0$, а коэффициенты

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} A \cos k\omega_1 t dt = \frac{2A}{k\omega_1 T} \int_{-k\omega_1 \tau_u / 2}^{k\omega_1 \tau_u / 2} \cos x dx = \frac{2A\tau_u}{T} \frac{\sin \frac{k\omega_1 \tau_u}{2}}{k\omega_1 \tau_u / 2}.$$

При вычислении интеграла произведена замена переменной $k\omega_1 t = x$, $dt = dx/k\omega_1$.

Подставляя вычисленные значения для коэффициентов ряда Фурье в формулу (2.11), получим

$$u(t) = A \frac{\tau_{\text{и}}}{T} + \frac{2A\tau_{\text{и}}}{T} \sum_{k=1}^{\infty} \frac{\sin k\omega_1 \tau_{\text{и}} / 2}{k\omega_1 \tau_{\text{и}} / 2} \cos k\omega_1 t. \quad (2.12)$$

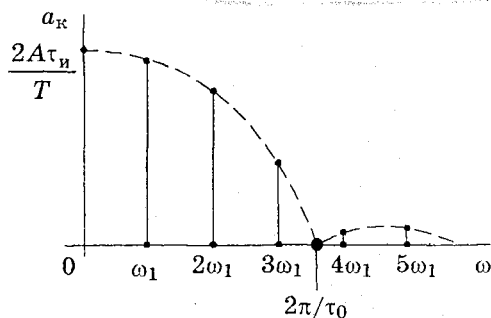


Рис.2.2. График спектра амплитуд

На рис.2.2 приведен график спектра амплитуд, построенный по результатам вычислений. Спектр дискретный, т.е. существует только для дискретных значений частот, которые соответствуют числам $k = 0, 1, 2, 3, \dots$. В промежутках между этими значениями спектр равен нулю.

Обратимся к ряду Фурье (2.6) в комплексной форме и запишем выра-

жение для периодической функции времени. Под *периодической* понимают такую функцию, которая удовлетворяет уравнению:

$$\begin{aligned} u(t) = u(t + mT) &= \frac{1}{2} \sum_{k=-\infty}^{\infty} a(jk\omega_1) e^{jk\omega_1(t+mT)} = \frac{1}{2} \sum_{k=-\infty}^{\infty} a(jk\omega_1) e^{jk\omega_1 t} e^{jm k\omega_1 T} = \\ &= \frac{1}{2} \sum_{k=-\infty}^{\infty} a(jk\omega_1) e^{jk\omega_1 t}. \end{aligned}$$

Числа k, m являются целыми, а $\omega_1 T = 2\pi$. Поэтому $e^{jkm\omega_1 T} = 1$ и разложение функции на интервале $[-T/2, T/2]$ совпадает с разложением этой же функции на других интервалах $[T/2, 3T/2]$, $[-T/2, -3T/2]$ и т.д. Другими словами, в ряд Фурье раскладываются периодические функции времени.

2.2. Спектральная плотность сигнала

Пусть период сигнала (см. рис.2.1) стремится к бесконечности. При увеличении периода T частота $\omega_1 = 2\pi/T$ уменьшается до бесконечно малой величины

$$\lim_{T \rightarrow \infty} \omega_1 = \lim_{T \rightarrow \infty} 2\pi/T = d\omega.$$

Расстояние между спектральными компонентами при этом уменьшается до бесконечно малой величины, а дискретные компоненты $k\omega_1$ превращаются в текущие значения частоты ω (см. рис.2.2). Интервал разложения $[-T/2, T/2]$ стремится к бесконечной величине. С учетом этого выражение для ряда Фурье в комплексной форме перепишем в виде:

$$u(t) = \lim_{T \rightarrow \infty} \frac{1}{2} \sum_{k=-\infty}^{\infty} a(jk\omega_1) e^{jk\omega_1 t} = \lim_{T \rightarrow \infty} \frac{1}{2} \sum_{k=-\infty}^{\infty} \left[\frac{2}{T} \int_{-T/2}^{T/2} u(t) e^{-jk\omega_1 t} dt \right] e^{jk\omega_1 t} =$$

$$\lim_{T \rightarrow \infty} \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \left(\int_{-T/2}^{T/2} u(t) e^{-jk\omega_1 t} dt \right) \frac{2\pi}{T} e^{jk\omega_1 t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt \right) e^{j\omega t} d\omega. \quad (2.13)$$

Выпишем интеграл, который записан в скобках выражения (2.13), и обозначим его символом $S(j\omega)$.

$$S(j\omega) = \int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt. \quad (2.14)$$

Тогда выражение (2.13) запишется более компактно:

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega. \quad (2.15)$$

Выражения (2.14) и (2.15) называются соответственно *прямым* и *обратным преобразованиями Фурье*. Величину $S(j\omega)$ называют *спектральной плотностью*. Она является комплексной величиной и имеет размерность [В/Гц].

С учетом формулы Эйлера перепишем выражение (2.14)

$$\begin{aligned} S(j\omega) &= \int_{-\infty}^{\infty} u(t) [\cos \omega t - j \sin \omega t] dt = \int_{-\infty}^{\infty} u(t) \cos \omega t dt - j \int_{-\infty}^{\infty} u(t) \sin \omega t dt = \\ &= A(\omega) - jB(\omega) = |S(j\omega)| e^{j\varphi(\omega)}, \end{aligned}$$

где

$$A(\omega) = \int_{-\infty}^{\infty} u(t) \cos \omega t dt, \quad B(\omega) = \int_{-\infty}^{\infty} u(t) \sin \omega t dt,$$

$$|S(j\omega)| = \sqrt{A^2(\omega) + B^2(\omega)}, \quad \varphi(\omega) = -\arctg \frac{B(\omega)}{A(\omega)}.$$

Модуль $|S(j\omega)|$ называется *спектральной плотностью амплитуд* или *спектром непериодического сигнала*. Фазовой характеристикой называется величина $\varphi(\omega)$.

Пример 2.2. Рассчитать спектральную плотность задержанного во времени одиночного прямоугольного импульса с длительностью τ_n и амплитудой A .

Решение. Спектральная плотность задержанного на время t_0 импульса произвольной формы равна:

$$S(j\omega) = \int_{-\infty}^{\infty} u(t - t_0) e^{-j\omega t} dt = e^{-j\omega t_0} \int_{-\infty}^{\infty} u(x) e^{-j\omega x} dx = S_0(j\omega) e^{-j\omega t_0},$$

$$S_0(j\omega) = \int_{-\infty}^{\infty} u(x)e^{-j\omega x} dx, \quad t - t_0 = x, \quad dt = dx.$$

Вычисления дают следующий результат

$$S(j\omega) = Ae^{-j\omega t_0} \int_0^{\tau_n} e^{-j\omega t} dt = \left(A\tau_n e^{-j\omega \tau_n / 2} \frac{\sin \frac{\omega \tau_n}{2}}{\omega \tau_n / 2} \right) e^{-j\omega t_0}. \quad (2.16)$$

Пример 2.3. Вычислить спектральную плотность кодированного сигнала

$$u(t) = \sum_{k=0}^{N-1} a_k u_0(t - k\tau_n),$$

где $a_k = \pm 1$, $u_0(t)$ прямоугольный импульс с амплитудой равной A и длительностью τ_n .

Решение. Применим формулу (2.14)

$$S(j\omega) = \int_{-\infty}^{\infty} \sum_{k=0}^{N-1} a_k u_0(t - k\tau_n) e^{-j\omega t} dt.$$

Производя замену переменной $t - k\tau_n = x$, $dt = dx$ и, принимая во внимание формулу (2.16), получим

$$S(j\omega) = \sum_{k=0}^{N-1} a_k \left(\int_0^{\tau_n} u_0(x) e^{-j\omega x} dx \right) e^{-j\omega k\tau_n} = A\tau_n \frac{\sin \frac{\omega \tau_n}{2}}{\omega \tau_n / 2} e^{-j\omega \tau_n / 2} \sum_{k=0}^{N-1} a_k e^{-j\omega k\tau_n}.$$

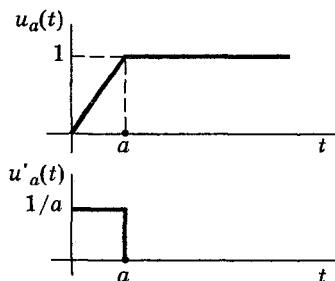


Рис.2.3. Близкая к единичной функция и ее производная

Пример 2.4. Вычислить спектральную плотность периодического сигнала, записанного в виде ряда Фурье.

Решение. При решении этой задачи следует воспользоваться основными сведениями о δ -функции. Возьмем функцию $u_a(t)$, близкую к единичной и вычислим производную $u'_a(t)$. Функция $u_a(t)$ и ее производная показаны на рис.2.3.

Под δ -функцией понимают предел производной $u'_a(t)$ при $a \rightarrow 0$.

$$\delta(t) = \lim_{a \rightarrow 0} u'_a(t) = \lim_{a \rightarrow 0} \frac{u_a(t-a) - u_a(0)}{a} = \lim_{a \rightarrow 0} \frac{1}{a}, \quad a \geq t \geq 0.$$

Здесь параметр a определяет длительность переднего фронта функции $u_a(t)$, изменяющегося по линейному закону (рис.2.3).

Рассмотрим основные свойства δ -функции. При $a \rightarrow 0$ δ -функция стремится к бесконечности.

$$\int_{-\infty}^{\infty} \delta(t) dt = 1,$$

т.е. площадь, ограниченная δ -функцией равна 1.

Полезным для расчетов является *фильтрующее* свойство δ -функции, которое заключается в следующем. Интеграл от произведения некоторой функции $u(t)$ на δ -функцию равен значению этой функции при t , для которого $\delta(t) \neq 0$.

Например:

$$\int_{-\infty}^{\infty} u(t) \delta(t) dt = u(0); \quad \int_{-\infty}^{\infty} u(t - \tau) \delta(t) dt = u(-\tau).$$

Спектральная плотность δ -функции

$$S(j\omega) = \int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = 1.$$

Обратное преобразование Фурье от спектральной плотности δ -функции дает следующий важный результат:

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega. \quad (2.17)$$

В частотной области выражение для δ -функции записывается в виде

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} dt. \quad (2.18)$$

Вычислим спектральную плотность, подставляя выражение (2.11) в (2.14):

$$\begin{aligned} S(j\omega) = \int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt = & \frac{a_0}{2} \int_{-\infty}^{\infty} e^{-j\omega t} dt + \sum_{k=1}^{\infty} (a_k \int_{-\infty}^{\infty} \cos(k\omega_1 t) e^{-j\omega t} dt + \\ & + b_k \int_{-\infty}^{\infty} \sin(k\omega_1 t) e^{-j\omega t} dt). \end{aligned}$$

Принимая во внимание формулу (2.18), а также формулы Эйлера

$$\cos x = \frac{1}{2}(e^{jx} + e^{-jx}), \quad \sin x = \frac{1}{2j}(e^{jx} - e^{-jx}),$$

получим окончательное выражение для спектральной плотности:

$$\begin{aligned} S(j\omega) = & a_0 \pi \delta(0) + \pi \sum_{k=1}^{\infty} \{ a_k [\delta(k\omega_1 - \omega) + \delta(-k\omega_1 - \omega)] + \\ & + j b_k [-\delta(k\omega_1 - \omega) + \delta(-k\omega_1 - \omega)] \}. \end{aligned}$$

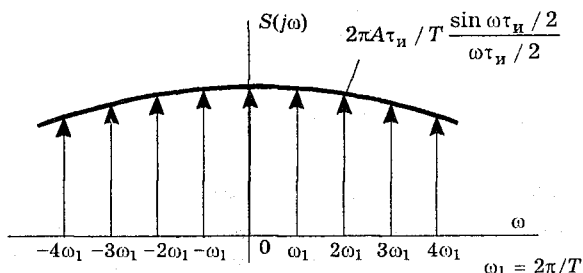


Рис.2.4. Спектральная плотность периодической последовательности импульсов

Для периодической последовательности импульсов, приведенной на рис.2.1, спектральная плотность определяется формулой при $a_0 =$

$$a_k = 2A(\tau_n/T) \times \left(\sin \frac{k\omega_1 \tau_n}{2} \right) / \frac{k\omega_1 \tau_n}{2},$$

$$b_k = 0.$$

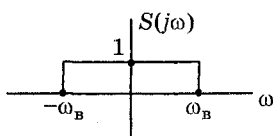


Рис.2.5. Спектральная плотность сигнала

Эта спектральная плотность показана на рис.2.4.

Пример 2.5. По спектральной плотности, вид которой приведен на рис.2.5, вычислить выражение для сигнала $u(t)$.

Решение. Спектральная плотность сигнала ограничена значениями частоты, равными $-\omega_B, \omega_B$.

Сигнал

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} e^{j\omega t} d\omega = 2F_B \frac{\sin \omega_B t}{\omega_B t}. \quad (2.19)$$

2.3 Спектральная плотность мощности детерминированного сигнала

Под энергией сигнала $u(t)$ понимают величину

$$E = \int_{-\infty}^{\infty} u^2(t) dt.$$

Если сигнал имеет конечную длительность T , т.е. не равен нулю на отрезке времени $[-T/2, T/2]$, то его энергия равна

$$E = \int_{-T/2}^{T/2} u^2(t) dt.$$

Запишем выражение для энергии сигнала, используя формулу (2.15):

$$E = \int_{-\infty}^{\infty} u^2(t) dt = \int_{-\infty}^{\infty} u(t) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega \right) dt =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(t) S(j\omega) e^{j\omega t} d\omega dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(j\omega)|^2 d\omega = \int_{-\infty}^{\infty} |S(j\omega)|^2 df,$$

где

$$S^*(j\omega) = \int_{-\infty}^{\infty} u(t) e^{j\omega t} dt = |S(j\omega)| e^{-j\varphi(\omega)},$$

$$S(j\omega) S^*(j\omega) = |S(j\omega)| e^{j\varphi(\omega)} \cdot |S(j\omega)| e^{-j\varphi(\omega)} = |S(j\omega)|^2.$$

Полученное равенство называют *равенством Парсеваля*. Оно определяет энергию сигнала через временную функцию или спектральную плотность энергии, которая равна $|S(j\omega)|^2$.

Рассмотрим сигнал, существующий на ограниченном интервале времени $[-T/2, T/2]$. К такому сигналу применимо равенство Парсеваля. Следовательно,

$$\int_{-T/2}^{T/2} u^2(t) dt = \int_{-\infty}^{\infty} |S(j\omega)|^2 df.$$

Разделим левую и правую части равенства на время, равное T и устремим его к бесконечности:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} u^2(t) dt = \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{|S(j\omega)|^2}{T} df = \int_{-\infty}^{\infty} G(\omega) df,$$

где

$$G(\omega) = \lim_{T \rightarrow \infty} \frac{|S(j\omega)|^2}{T}. \quad (2.20)$$

С увеличением T энергия и величина $|S(j\omega)|^2$ незатухающих сигналов возрастают, однако отношение $|S(j\omega)|^2/T$ может стремиться к определенному пределу. Этот предел называется *спектральной плотностью мощности* $G(\omega)$.

2.4. Автокорреляционная функция

Под автокорреляционной функцией сигнала $u(t)$ понимают величину

$$R(\tau) = \int_{-\infty}^{\infty} u(t) u(t + \tau) dt, \quad (2.21)$$

где τ — некоторый постоянный временной сдвиг функции $u(t)$.

Автокорреляционная функция имеет следующие свойства.

1. Значение автокорреляционной функции при сдвиге $\tau = 0$ равно энергии сигнала E , т.е.

$$R(0) = \int_{-\infty}^{\infty} u(t)u(t)dt = \int_{-\infty}^{\infty} u^2(t)dt = E.$$

2. Автокорреляционная функция при сдвигах $\tau \neq 0$ меньше энергии сигнала:

$$R(\tau) < R(0) = E.$$

3. Автокорреляционная функция является четной функцией, т.е.

$$R(\tau) = R(-\tau).$$

В справедливости свойств 2 и 3 убедимся на следующем примере.

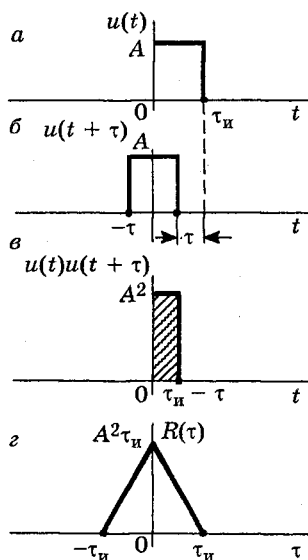


Рис.2.6. К вычислению автокорреляционной функции прямоугольного видеопульса:

- a* — прямоугольный импульс;
- б* — задержанный во времени прямоугольный импульс;
- в* — произведение импульсов;
- г* — автокорреляционная функция

Пример 2.6. Вычислить автокорреляционную функцию сигнала, показанного на рис. 2.6, *a*.

Решение. Автокорреляционная функция

$$R(\tau) = \int_{-\infty}^{\infty} u(t)u(t+\tau)dt$$

определяется интегралом от произведения функции $u(t)$ на ее сдвинутую на время $t = -\tau$ копию. Время сдвига находим из уравнения $t + \tau = 0$. График функции $u(t + \tau)$ приведен на рис.2.6, *б*. Площадь, определяемая графиком произведения функций $u(t) \cdot u(t + \tau)$ (рис.2.6, *в*), равна

$$A^2(\tau_n - \tau) = R(\tau) = \int_{-\infty}^{\infty} u(t)u(t+\tau)dt.$$

Отсюда

$$R(\tau) = A^2\tau_n(1 - \tau/\tau_n).$$

Функция $R(\tau)$ определяется уравнением прямой (рис.2.6, *г*). Функция максимальна при $\tau = 0$ и равна нулю при $\tau = \tau_n$. При значениях $\tau_n \geq \tau \geq 0$ $R(\tau) < R(0)$. Таким образом убеждаемся в справедливости 2-го свойства.

Чтобы убедиться в справедливости 3-го свойства, аналогично вычислим функцию для отрицательных значений τ :

$$R(-\tau) = \int u(t)u(t-\tau)dt = A^2\tau_n(1-\tau/\tau_n).$$

Окончательное выражение для автокорреляционной функции

$$R(\tau) = A^2\tau_n(1-|\tau|/\tau_n).$$

Покажем связь автокорреляционной функции $R(\tau)$ со спектральной плотностью энергии сигнала $|S(j\omega)|^2$. Выразим сдвинутый во времени сигнал $u(t+\tau)$ через его спектральную плотность $S(j\omega)$.

$$u(t+\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)e^{j\omega(t+\tau)}d\omega. \quad (2.22)$$

Подставим выражение (2.22) в выражение (2.21). В результате получим:

$$\begin{aligned} R(\tau) &= \int_{-\infty}^{\infty} u(t) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)e^{j\omega(t+\tau)}d\omega \right) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) \left(\int_{-\infty}^{\infty} u(t)e^{j\omega t}dt \right) e^{j\omega\tau}d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)S^*(j\omega)e^{j\omega\tau}d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(j\omega)|^2 e^{j\omega\tau}d\omega. \end{aligned} \quad (2.23)$$

Нетрудно убедиться также в справедливости равенства

$$|S(j\omega)|^2 = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau. \quad (2.24)$$

Для сигналов с неограниченной энергией автокорреляционная функция:

$$B(\tau) = \lim_{T \rightarrow \infty} \frac{R(\tau)}{T} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} u(t)u(t+\tau)dt.$$

Учитывая это, находим, что

$$B(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{|S(j\omega)|^2}{T} e^{j\omega\tau}d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)e^{j\omega\tau}d\omega. \quad (2.25)$$

Спектральная плотность мощности находится преобразованием автокорреляционной функции по формуле:

$$G(\omega) = \int_{-\infty}^{\infty} B(\tau)e^{-j\omega\tau}d\tau. \quad (2.26)$$

Пример 2.7. Вычислить автокорреляционную функцию кодированного сигнала $u(t) = \sum_{k=0}^2 a_k u_0(t - k\tau_n)$, если $u_0(t)$ — прямо-

угольный импульс с амплитудой A и длительностью $\tau_{\text{и}}$. Кодовая последовательность $\{a_k\} = 1, 1, -1$ (рис.2.7, а).

Решение. Подставляя выражение для сигнала в формулу (2.21), получим

$$\begin{aligned} R(\tau) &= \int_{-\infty}^{\infty} \sum_{k=0}^2 a_k u_0(t - k\tau_{\text{и}}) \times \\ &\times \sum_{m=0}^2 a_m u_0(t - m\tau_{\text{и}} + \tau) dt = \\ &= \sum_{k=0}^2 \sum_{m=0}^2 a_k a_m \int_{-\infty}^{\infty} u_0(t - k\tau_{\text{и}}) \times \\ &\times u_0(t - m\tau_{\text{и}} + \tau) dt. \end{aligned}$$

Произведем замену переменной $t - k\tau_{\text{и}}$ на x .

$$\begin{aligned} R(\tau) &= \sum_{k=0}^2 \sum_{m=0}^2 a_k a_m \int_{-\infty}^{\infty} u_0(x) u_0[x + \\ &+ (k - m)\tau_{\text{и}} + \tau] dx. \end{aligned}$$

Обозначим $k - m = q$ и, перейдем от дискретных переменных k, m к переменным k, q . В результате получим:

$$R(\tau) = \sum_{q=-2}^2 \sum_{k=0}^2 a_k a_{k+q} R_0(q\tau_{\text{и}} + \tau),$$

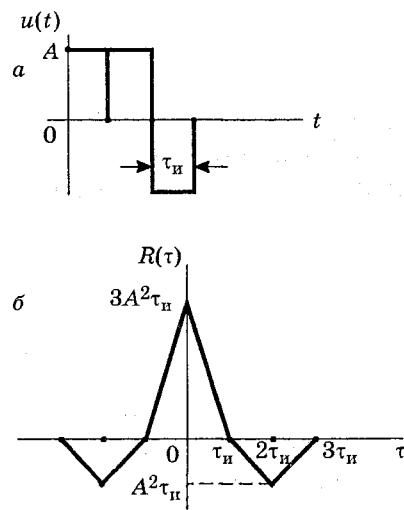


Рис.2.7. К вычислению автокорреляционной функции кодированного сигнала:
а — кодированный сигнал;
б — автокорреляционная функция сигнала

где

$$R_0(\tau) = \int_{-\infty}^{\infty} u_0(x) u_0(x + \tau) dx.$$

График автокорреляционной функции показан на рис.2.7, б.

2.5. Комплексное и квазигармоническое представление узкополосных сигналов

При анализе прохождения сигнала через блоки системы связи его удобно выразить в виде комплексной функции времени

$$\dot{W}(t) = u(t) + ju^*(t) = \vartheta(t)e^{j\theta(t)}, \quad (2.27)$$

где $u(t) = \text{Re } \dot{W}(t)$ — действительная часть, $u^*(t) = \text{Im } \dot{W}(t)$ — мнимая часть, $\vartheta(t)$ — огибающая, $\theta(t)$ — мгновенная фаза сигнала.

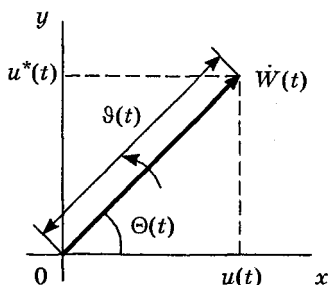


Рис.2.8. Графическое представление комплексного сигнала

Графическое представление сигнала приведено на рис.2.8. Поскольку фаза зависит от t , вектор $\dot{W}(t)$ вращается относительно точки 0. Условно считают, что вращение вектора направлено против часовой стрелки.

Сигнал $\dot{W}(t)$ называется *аналитическим*, если функции $u(t)$ и $u^*(t)$ образуют *пару преобразований Гильберта*:

$$u^*(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(\tau)}{t - \tau} d\tau. \quad (2.28)$$

$$u(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u^*(\tau)}{t - \tau} d\tau. \quad (2.29)$$

Функция $u^*(t)$ является *сопряженной* по отношению к функции $u(t)$. Если функции $u(t)$, $u^*(t)$ составляют пару преобразований Гильберта, то огибающая и фаза сигнала однозначно определяются по формулам:

$$\vartheta(t) = \sqrt{u^2(t) + u^{*2}(t)}, \quad \theta(t) = \arctg \frac{u^*(t)}{u(t)}.$$

Производная от фазы сигнала определяет *мгновенную частоту*

$$\omega(t) = d\theta(t) / dt.$$

По мгновенной частоте находится *мгновенная фаза сигнала*:

$$\theta(t) = \int \omega(t) dt.$$

Часто фазу сигнала представляют в виде:

$$\theta(t) = \omega_0 t + \varphi(t) + \varphi_0 = \omega_0 t + \varphi(t), \quad (2.30)$$

где $\varphi(t) = \varphi(t) + \varphi_0$; ω_0 — постоянная частота, которая называется *несущей*; $\varphi(t)$ — переменная составляющая, а φ_0 — постоянная составляющая фазы. Подставляя (2.30) в (2.27), получим

$$\begin{aligned} \dot{W}(t) &= \vartheta(t) \cos \theta(t) + j\vartheta(t) \sin \theta(t) = \\ &= \vartheta(t) \cos[\omega_0 t + \varphi(t)] + j\vartheta(t) \sin[\omega_0 t + \varphi(t)]. \end{aligned} \quad (2.31)$$

Действительная часть последнего выражения:

$$\begin{aligned} u(t) &= \operatorname{Re} \dot{W}(t) = \vartheta(t) \cos[\omega_0 t + \varphi(t)] = \vartheta(t) \cos \varphi(t) \cos \omega_0 t - \\ &- \vartheta(t) \sin \varphi(t) \sin \omega_0 t = A(t) \cos \omega_0 t - B(t) \sin \omega_0 t, \end{aligned} \quad (2.32)$$

где $A(t) = \vartheta(t) \cos \varphi(t)$, $B(t) = \vartheta(t) \sin \varphi(t)$ — *квадратурные составляющие*.

Математическим выражением (2.32) пользуются при изучении частотно-избирательных цепей с ограниченной шириной полосы пропускания. Если эта ширина не велика по сравнению с несущей частотой ω_0 , то выходной сигнал называется *узкополосным*. Узкополосный сигнал является *квазигармоническим*, т.е. "почти" гармоническим колебанием, у которого огибающая и фаза являются медленно меняющимися функциями времени.

Пример 2.8. Для функции $u(t) = \cos \omega_0 t$ найти сопряженную по Гильберту функцию $u^*(t)$.

Решение. Применяем формулу (2.28).

$$u^*(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\cos \omega_0 t}{t - \tau} d\tau.$$

Обозначим $t - \tau = \lambda$. Тогда

$$\begin{aligned} u^*(t) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\cos \omega_0(t - \lambda)}{\lambda} d\lambda = \frac{1}{\pi} \int_{-\infty}^{\infty} \cos \omega_0 t \frac{\cos \omega_0 \lambda}{\lambda} d\lambda + \\ &+ \frac{1}{\pi} \int_{-\infty}^{\infty} \sin \omega_0 t \frac{\sin \omega_0 \lambda}{\lambda} d\lambda = \frac{1}{\pi} \cos \omega_0 t \int_{-\infty}^{\infty} \frac{\sin \omega_0 \lambda}{\lambda} d\lambda + \\ &\frac{1}{\pi} \sin \omega_0 t \times \int_{-\infty}^{\infty} \frac{\sin \omega_0 \lambda}{\lambda} d\lambda = \sin \omega_0 t. \end{aligned}$$

При вычислении использованы следующие табличные интегралы:

$$\int_{-\infty}^{\infty} \frac{\cos \omega_0 \lambda}{\lambda} d\lambda = 0, \quad \int_{-\infty}^{\infty} \frac{\sin \omega_0 \lambda}{\lambda} d\lambda = \pi.$$

Из приведенного примера видно, что сопряженный по Гильберту сигнал находится путем сдвига фазы на $\pi/2$ у исходного сигнала $u(t) = \cos \omega_0 t$.

2.6. Геометрические представления сигналов

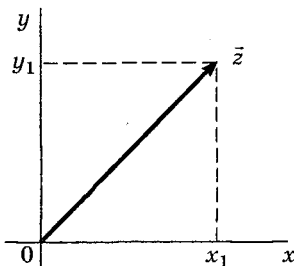


Рис.2.9. Двумерный вектор

При решении ряда задач, связанных с анализом и преобразованием сигналов, целесообразно отображать эти сигналы векторами некоторого векторного пространства. Это, как правило, позволяет упростить решение. Введем некоторые понятия, полезные для последующего анализа.

Совокупность двух чисел x_1, y_1 определяет координаты вектора $\vec{z}$ в двумерном пространстве (рис.2.9). Длина

вектора $\vec{z}$ равна его норме

$$d = \|\vec{z}\| = \sqrt{x_1^2 + y_1^2}. \quad (2.33)$$

Расстояние между векторами $\vec{z}_1$ и $\vec{z}_2$ (см. рис.2.10)

$$d(\vec{z}_1, \vec{z}_2) = \|\vec{z}_1 - \vec{z}_2\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (2.34)$$

Скалярное произведение двух векторов

$$(\vec{z}_1 \cdot \vec{z}_2) = x_1 x_2 + y_1 y_2, \quad (2.35)$$

$$(\vec{z}_1 \cdot \vec{z}_2) = \|\vec{z}_1\| \cdot \|\vec{z}_2\| \cos \alpha. \quad (2.36)$$

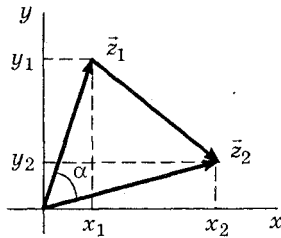


Рис.2.10. Расстояние между двумя векторами

Формула (2.35) определяет скалярное произведение через проекции векторов $\vec{z}_1$, $\vec{z}_2$ на оси x и y . Формула (2.36) дает выражение для скалярного произведения через произведение длин двух векторов на косинус угла между ними.

Условие ортогональности векторов

$$(\vec{z}_1 \cdot \vec{z}_2) = x_1 x_2 + y_1 y_2 = 0. \quad (2.37)$$

Свойства n -мерного пространства являются обобщением свойств двумерного. Поэтому, зная формулы (2.33-2.37) для двумерного пространства, легко написать аналогичные формулы для n -мерного. Например, длина n -мерного вектора $\vec{u}(u_1, u_2, \dots, u_n)$ определяется нормой

$$d = \|\vec{u}\| = \sqrt{\sum_{i=1}^n u_i^2}.$$

Расстояние между двумя n -мерными векторами

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{\sum_{i=1}^n (u_i - v_i)^2}.$$

Скалярное произведение двух n -мерных векторов

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^n u_i v_i.$$

Развитием понятия векторного пространства является функциональное пространство. Норма функции $u(t)$ определяется числом

$$\|u(t)\| = \sqrt{\int_{-T/2}^{T/2} u(t) dt}, \quad (2.38)$$

где T — интервал времени, на котором определена функция $u(t)$. Расстояние между функциями $u(t)$ и $v(t)$ равно норме разности

$$\|u(t) - v(t)\| = \sqrt{\int_{-T/2}^{T/2} [u(t) - v(t)]^2 dt}.$$

Скалярное произведение функций $u(t)$ и $v(t)$ определяется как число

$$(u, v) = \int_{-T/2}^{T/2} u(t)v(t)dt. \quad (2.39)$$

Функции $u(t)$ и $v(t)$ ортогональны, если

$$\int_{-T/2}^{T/2} u(t)v(t)dt = 0.$$

Пространство функций со скалярным произведением (2.39) и нормой (2.38) называется *функциональным пространством Гильберта*.

Рассмотрим скалярное произведение спектральной плотности сигнала $v(t)$ на сдвинутое по частотной оси зеркальное отображение спектральной плотности сигнала $u(t)$:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(jz) S_u[j\omega - z] dz = \int_{-\infty}^{\infty} u(t)v(t)e^{-j\omega t} dt, \quad (2.40)$$

где $S_v(jz)$ спектральная плотность сигнала $v(t)$, $S_u(-jz)$ — зеркальное отображение спектральной плотности сигнала $u(t)$.

В справедливости формулы (2.40) убеждаемся в результате следующих преобразований:

$$\begin{aligned} \int_{-\infty}^{\infty} u(t)v(t)e^{-j\omega t} dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} u(t) \int_{-\infty}^{\infty} [S_v(jz)e^{jzt}] e^{-j\omega t} dz dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(jz) \int_{-\infty}^{\infty} u(t)e^{-j(\omega-z)t} dt dz = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(jz) \cdot S_u[j(\omega - z)] dz. \end{aligned}$$

Выражение (2.40) называется также *сверткой спектров*. Свертку можно также записать и в другом виде, если учесть, что

$$S_u[j(\omega - z)] = S_u^*[j(z - \omega)]:$$

$$\int_{-\infty}^{\infty} u(t)v(t)e^{-j\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(jz) S_u^*[j(z - \omega)] dz. \quad (2.41)$$

При $\omega = 0$ формула (2.40) превращается в известную формулу Рэлея, определяющая скалярное произведение двух функций:

$$\int_{-\infty}^{\infty} u(t)v(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(jz) S_u^*(jz) dz. \quad (2.42)$$

Если $u(t) = v(t)$, то равенство (2.42) превращается в равенство Парсеваля.

2.7. Теорема Котельникова

Для функций с ограниченным спектром В.А. Котельников доказал теорему, которая является теоретической основой построения дискретных систем связи. Ее содержание следующее.

Любая функция времени $u(t)$ с ограниченным спектром полностью определяется своими значениями, отсчитанными в дискретные моменты времени через интервал $\Delta t = 1/2F_B$, где F_B — верхняя граничная частота спектра этой функции.

Ряд, определяющий функцию,

$$u(t) = \sum_{k=-\infty}^{\infty} u(k, \Delta t) \frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)}, \quad (2.43)$$

является разновидностью обобщенного ряда Фурье (2.2). Он построен на базе ортонормированных функций

$$\varphi_k(t) = \sqrt{2F_B} \frac{\sin \omega_B(t - k\Delta t)}{\omega_B(t - k\Delta t)}, \quad (2.44)$$

где k — целые числа в пределах от $-\infty$ до ∞ .

Эти функции удовлетворяют следующим условиям:

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_k(t) \varphi_l(t) dt &= 0, \quad k \neq l \\ \int_{-\infty}^{\infty} \varphi_k^2(t) dt &= 1, \quad k = l. \end{aligned}$$

Коэффициенты a_l ряда вычисляются по формуле

$$a_l = \int_{-\infty}^{\infty} u(t) \varphi_l(t) dt.$$

Для вычисления коэффициентов a_l воспользуемся формулой (2.42). Обращаясь к формуле (2.20), найдем, что спектру с единичной АЧХ в диапазоне частот $[-\omega_B, \omega_B]$ соответствует сигнал

$$2F_B \sin \omega_B t / \omega_B t.$$

Сравнивая этот сигнал с базисной функцией (2.44), определяем ее спектр

$$\frac{1}{\sqrt{2F_B}} e^{-j\omega t}.$$

Его АЧХ меньше в $\sqrt{2F_B}$ раз, а фазочастотная характеристика (ФЧХ) равна $e^{-j\omega \Delta t}$.

$$a_l = \int_{-\infty}^{\infty} u(t) \varphi_l(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\varphi}(jz) \cdot S_u(-jz) dz =$$

$$= \frac{1}{\sqrt{2F_B}} \cdot \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} S_u(-jz) e^{-jz\Delta t} dz = \frac{1}{\sqrt{2F_B}} \cdot \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} S_u(j\omega) e^{j\omega\Delta t} d\omega = \frac{1}{\sqrt{2F_B}} u(\Delta t).$$

Подставляя вычисленное значение для коэффициента a_l обобщенного ряда Фурье и значение базисной функции (2.44) в выражении (2.2), получим *ряд Котельникова* (2.43).

Изучая теорему Котельникова, полезно обратить внимание на следующее.

1. В теореме не оговаривается вид функции $u(t)$. Поэтому она может быть и случайной.

2. Из равенства (2.43) следует, что вся информация о функции $u(t)$ содержится в ее выборочных значениях $u(k, \Delta t)$, так как базисные функции, являясь известными, никаких сведений не несут. Поэтому, передавая выборочные значения по каналу связи, сохраняем всю необходимую информацию о сообщении. Можно передавать выборочные значения непосредственно, а можно перевести их в числа.

3. Теорема определяет восстановление сообщения на приемном конце: необходимо умножить выборочные значения на базисные функции и результаты сложить.

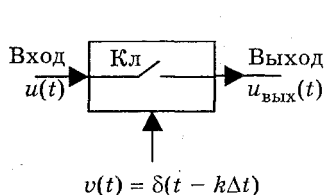


Рис.2.11. Структурная схема дискретизатора

Процесс взятия выборок называется *дискретизацией* (выборки обычно берут через одинаковые интервалы времени $\Delta t = 1/2F_B$). Устройство, с помощью которого ведется дискретизация, называется *дискретизатором*.

Он является частью аналого-цифрового преобразователя. В этом качестве можно использовать управляемый ключ, замыкаемый на непродолжительное время. Структурная схема дискретизатора, выполненная на ключе, приведена на рис.2.11. Временные диаграммы, поясняющие работу дискретизатора, приведены на рис.2.12. При поступлении

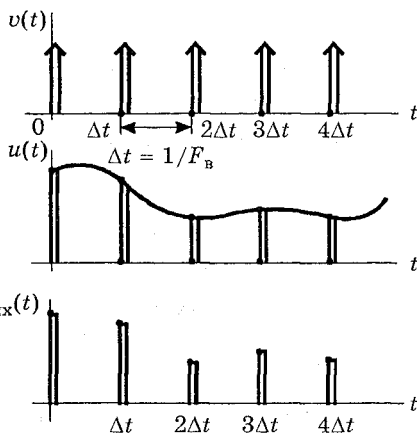


Рис.2.12. Временные диаграммы, поясняющие работу дискретизатора

управляющих импульсов $v(t)$ выход ключа соединяется со входом устройства и на выходе образуются выборочные значения. При восстановлении в приемнике непрерывной функции нужно в каждый момент времени $k\Delta t$, $k = 0, 1, 2, 3 \dots$ генерировать базисную функцию. Эта операция осуществляется с помощью фильтра нижних частот, АЧХ которого постоянна в диапазоне $[-\omega_B, \omega_B]$ и равна $1/\sqrt{2F_B}$, а ФЧХ равна $e^{-jk\omega\Delta t}$. Базисная функция $\varphi_0(t)$ для $k = 0$ показана на рис.2.13.

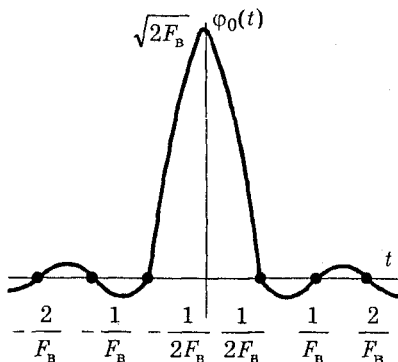


Рис.2.13. Базисная функция

Рассмотрим реакцию фильтра нижних частот (рис.2.14) на входной δ -импульс. Коэффициент передачи $K(j\omega)$ этого фильтра в диапазоне частот $[-\omega_B, \omega_B]$ равен $1/\sqrt{2F_B}$. Реакция любого линейного четырехполюсника на входной δ -импульс называется *импульсной характеристикой* этого четырехполюсника $h(t)$. Вычислим спектральную плотность сигнала на выходе фильтра:

$$S_{\text{ВЫХ}}(j\omega) = S_{\text{ВХ}}(j\omega) \cdot K(j\omega), \quad (2.45)$$

где $S_{\text{ВХ}}(j\omega)$, $S_{\text{ВЫХ}}(j\omega)$ — спектральные плотности сигналов на входе и выходе фильтра.

Спектральная плотность входного сигнала:

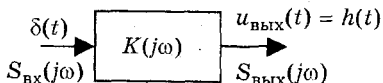


Рис.2.14. Схема линейного четырехполюсника (фильтра нижних частот)

$$S_{\text{ВХ}}(j\omega) = \int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = e^0 = 1.$$

Подставляя полученное выражение в формулу (2.45), получим:

$$S_{\text{ВЫХ}}(j\omega) = K(j\omega).$$

Найдем сигнал на выходе фильтра, применяя обратное преобразование Фурье (2.15):

$$u_{\text{ВЫХ}}(t) = h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(j\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} e^{j\omega t} d\omega = \sqrt{2F_B} \frac{\sin \omega_B t}{\omega_B}.$$

Процесс восстановления сигнала по его выборкам поясняется рис.2.15.

Для реальных сигналов, которые начинаются в момент времени $t = 0$ и заканчиваются в момент времени $T = (N - 1)\Delta t$, ряд Котельникова записывается в виде

$$u(t) = \sum_{k=0}^{N-1} u(k\Delta t) \frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)}. \quad (2.46)$$

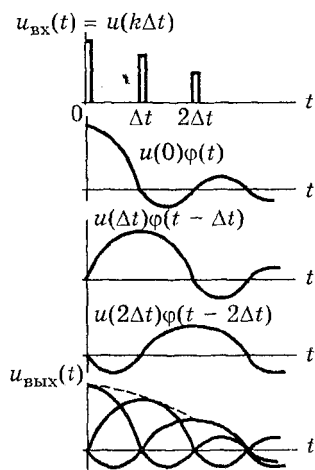


Рис.2.15. Временные диаграммы, поясняющие восстановление сигнала

тами f_1 и f_2 , (рис.2.16) и записывается в виде

$$u(t) = \vartheta(t) \cos[2\pi f_0 t + \varphi(t)], \quad (2.47)$$

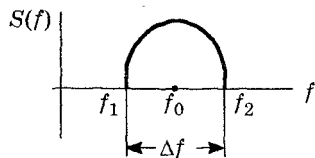


Рис.2.16. Спектр узкополосного сигнала

Энергия таких сигналов почти полностью сосредоточена в указанном интервале времени и в диапазоне частот $[-\omega_B, \omega_B]$. Число отсчетов в интервале T называется базой сигнала. База равна

$$T/\Delta t = 2F_B T.$$

Если энергия сигнала не будет сосредоточена в указанных временном и частотном интервалах, то формула (2.46) окажется приближенной. При этом возникает погрешность

$$\varepsilon = \Delta E/E,$$

где E — энергия сигнала, ΔE — энергия за пределами полосы $\mp \omega_B$.

$$\Delta E = 2 \int_{2\pi F_B}^{\infty} S(j\omega) d\omega, \quad E = \int_{-\infty}^{\infty} |S(j\omega)|^2 d\omega,$$

где $S(j\omega)$ — спектральная плотность сигнала.

Если функция $u(t)$ имеет относительно узкий спектр, ограниченный частотами f_1 и f_2 , (рис.2.16) и записывается в виде

где $f_0 = (f_1 + f_2)/2$; $\vartheta(t)$ — огибающая, $\varphi(t)$ — начальная фаза узкополосного колебания, то ее можно разложить в ряд

$$u(t) = \sum_{k=1}^{n_1} (\varphi'_k(t) \cdot \vartheta_k \cos(2\pi f_0 t + \varphi_k)). \quad (2.48)$$

Выражения, входящие в последнюю формулу равны:

$$\varphi'_k(t) = \frac{\sin \pi \Delta f(t - k\Delta t)}{\pi \Delta f(t - k\Delta t)}; \quad \Delta f = f_2 - f_1;$$

$$\vartheta_k = \vartheta(k\Delta t); \quad \varphi_k = \varphi(k\Delta t); \quad 2n_1 = 2\Delta f T,$$

T — длительность функции.

Функция $u(t)$ полностью определяется значениями ее амплитуды и фазы, взятыми через интервалы времени $\Delta t = 1/\Delta f$. Общее количество выборочных значений $2n_1 = 2\Delta f T$.

Контрольные вопросы

1. Что Вы понимаете под детерминированным сигналом?
2. Как записываются выражения для сигнала с использованием рядов Фурье в комплексной и действительной формах?
3. В чем отличие между спектром сигнала и спектральной плотностью?
4. Как связаны между собой временная функция сигнала $u(t)$ и его спектральная плотность $S(j\omega)$?
5. В каком соотношении находятся спектральная плотность мощности и автокорреляционная функция сигнала с неограниченной энергией?
6. Что Вы понимаете под аналитическим сигналом?
7. Какой сигнал называется узкополосным, полосовым?
8. Какому условию удовлетворяют ортогональные сигналы?
9. Чему равно расстояние между двумя функциями $v(t)$ и $u(t)$?
10. Как формулируется теорема Котельникова?

Глава 3. МЕТОДЫ ФОРМИРОВАНИЯ И ПРЕОБРАЗОВАНИЯ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ В СИСТЕМАХ СВЯЗИ

3.1. Определения линейной, нелинейной и параметрической систем

Структурная схема системы связи состоит из отдельных блоков, между которыми заданы определенные связи. Блоки, как и система в целом, имеют входы и выходы.

При подаче сигнала $u_{\text{вх}}(t)$ на вход блока на его выходе формируется сигнал $u_{\text{вых}}(t)$.

Операция L преобразования входного сигнала, в результате которой формируется определенный выходной сигнал, называется оператором

Оператор вместе с условиями и ограничениями, которые накладываются на входной и выходной сигналы, определяет *математическую модель блока (системы)*.

Система (или блок), у которой выходное напряжение не зависит от момента времени, с которого начинается входной сигнал, называется *стационарной*. Показатели качества и параметры стационарной системы не зависят от времени. Если указанное условие не выполняется, система называется *нестационарной*.

Под *линейной* понимают такую систему, оператор которой удовлетворяет условиям:

$$L[u_{\text{вх1}}(t) + u_{\text{вх2}}(t)] = L[u_{\text{вх1}}(t)] + L[u_{\text{вх2}}(t)], \quad (3.1)$$

$$L[\alpha u_{\text{вх}}(t)] = \alpha L u_{\text{вх}}(t), \quad (3.2)$$

где α — произвольное число.

Аналогичное определение дается для линейного блока или устройства.

Если условия (3.1), (3.2) не выполняются, то система (блок, устройство) называется *нелинейной (нелинейным)*.

Равенство (3.1) выражает свойство *аддитивности*, а равенство (3.2) — свойство *однородности*. Свойство аддитивности заключается в том, что реакция на сумму сигналов равна сумме реакций на каждый из этих сигналов в отдельности.

Под *параметрической* понимают такую систему, параметры которой изменяются во времени. Параметрические системы делятся на линейные и нелинейные. Отклик линейной параметрической цепи на входной сигнал равен

$$u_{\text{вых}}(t) = \theta[z(t)] \cdot u_{\text{вх}}(t),$$

где $\theta[z(t)]$ — оператор, зависимый от времени и выражающий параметрическое преобразование.

Оператор $\theta[z(t)]$ можно представить в виде полинома по некоторым функциям времени $\varphi_n(t)$:

$$u_{\text{вых}}(t) = \Phi(t)u_{\text{вх}}(t), \quad (3.3)$$

где $\Phi(t) = q_0 + q_1\varphi_1(t) + q_2\varphi_2(t) + \dots$; $q_0, q_1, \dots, q_n$ — постоянные множители.

Ниже рассматриваются только линейные параметрические системы.

3.2. Преобразования спектров сигналов в параметрических и нелинейных цепях

3.2.1. Линейные преобразования дискретного спектра

Вычислительный или экспериментальный процесс определения спектрального состава сигнала называется *спектральным анализом*.

Между сигналами и их спектрами существует взаимнооднозначное соответствие, поэтому задача преобразования сигнала может быть сформулирована и как задача преобразования его спектра: при заданных входном спектре сигнала $S_{\text{вх}}(j\omega)$ и операторе преобразования требуется найти спектр выходного сигнала.

Ниже рассмотрены задачи преобразования дискретных спектров, непрерывных и дискретно-непрерывных спектральных плотностей сигналов в линейных, параметрических и нелинейных цепях. Уточним понятия дискретных спектров, непрерывных и дискретно-непрерывных спектральных плотностей.

Во 2-й главе введено понятие дискретного спектра сигнала, как совокупности коэффициентов обобщенного ряда Фурье (2.2). В дальнейшем, если это не будет оговорено дополнительно, под *дискретным* будем понимать спектр, найденный в базисе тригонометрических функций ($\sin x, \cos x$). Этот спектр определяется амплитудами и фазами гармонических сигналов, несущие частоты которых однозначно заданы на оси частот.

Во 2-й главе введено также понятие спектральной плотности, которая связана с сигналом прямым преобразованием Фурье и может быть непрерывной, дискретной и дискретно-непрерывной. Непрерывная спектральная плотность соответствует сигналу $u(t)$, удовлетворяющему условию абсолютной интегрируемости:

$$\int_{-\infty}^{\infty} |u(t)| dt < \infty. \quad (3.4)$$

Дискретная и дискретно-непрерывная спектральные плотности определяются сигналами, которые условию (3.4) не удовлетворяют, однако и для них может быть найдено прямое преобразование Фурье.

При линейном преобразовании спектральные компоненты входного сигнала изменяются по амплитуде и приобретают дополнительный сдвиг по фазе. Это приводит к искажению его формы.

Для входного дискретного спектра выходной спектр является также дискретным. Пусть входной сигнал $u_{\text{вх}}(t)$ на интервале $[-T/2, T/2]$ представлен в виде ряда Фурье

$$u_{\text{вх}}(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a(k\omega_1) \cos[k\omega_1 t + \varphi(k\omega_1)], \quad (3.5)$$

где

$$a(k\omega_1) = \sqrt{a_k^2 + b_k^2}, \quad \varphi(k\omega_1) = -\arctg(b_k / a_k).$$

Ряд (3.5) получен на основе выражения (2.11) с учетом формул (2.9) и (2.10).

При воздействии k -й гармоники входного сигнала $u_{\text{вх}}(t)$ на линейную цепь с коэффициентом передачи $K(k\omega_1)$ на ее выходе появится сигнал с амплитудой $a(k\omega_1) \cdot K(k\omega_1)$ и начальной фазой $\varphi(k\omega_1) = \varphi(k\omega_1) + \varphi_k(k\omega_1)$, где $\varphi_k(k\omega_1)$ — фазовый сдвиг линейной цепи на частоте $k\omega_1$.

Суммируя результат воздействия на линейную систему всех гармоник, включая и постоянную составляющую, получим

$$u_{\text{вых}}(t) = \frac{a_0}{2} K(0) + \sum_{k=1}^{\infty} a(k\omega_1) |K(k\omega_1)| \cos[k\omega_1 t + \varphi(k\omega_1)]. \quad (3.6)$$

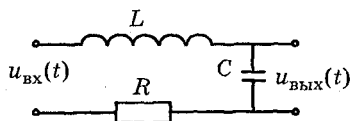


Рис.3.1. Последовательный колебательный контур в режиме передачи напряжения

Пример 3.1. Для последовательного колебательного контура, схема которого показана на рис.3.1, найти выходное напряжение. Входной сигнал определен рядом Фурье (3.5).

Решение. Коэффициент передачи линейной цепи на частоте $k\omega_1$ имеет вид

$$K(k\omega_1) = \frac{Q(\omega_0 / k\omega_1)}{\sqrt{1 + Q^2(k\omega_1 / \omega_0 - \omega_0 / k\omega_1)^2}} e^{j\varphi_k k\omega_1},$$

где $\omega_0 = 1/\sqrt{LC}$ — резонансная частота; Q — добротность контура;

$$\varphi_k(k\omega_1) = -\frac{\pi}{2} - \operatorname{arctg} \frac{k\omega_1 L - 1/k\omega_1 c}{R} = -\frac{\pi}{2} - \operatorname{arctg}[Q(k\omega_1/\omega_0 - \omega_0/k\omega_1)].$$

Подставляя выражение для коэффициента передачи в (3.6), получим:

$$u_{\text{вых}}(t) = \frac{a_0}{2} + Q \sum_{k=1}^{\infty} \frac{a(k\omega_1) \frac{\omega_0}{k\omega_1}}{\sqrt{1 + Q^2 \left(\frac{k\omega_1}{\omega_0} - \frac{\omega_0}{k\omega_1} \right)^2}} \cos[k\omega_1 t + \varphi(k\omega_1) + \varphi_k(k\omega_1)].$$

3.2.2. Линейное преобразование непрерывной спектральной плотности

Пусть оператор линейного преобразования сигнала задан в виде интеграла Дюамеля

$$u_{\text{вых}}(t) = \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) h(t - \tau) d\tau, \quad (3.7)$$

где $h(t)$ — импульсная характеристика линейной цепи. Эта характеристика связана с коэффициентом передачи линейной цепи $K(j\omega)$ обратным преобразованием Фурье:

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(j\omega) e^{j\omega t} d\omega.$$

Спектральную плотность входного сигнала найдем, применяя прямое преобразование Фурье:

$$S_{\text{вх}}(j\omega) = \int_{-\infty}^{\infty} u_{\text{вх}}(t) e^{-j\omega t} dt.$$

Спектральная плотность выходного сигнала

$$\begin{aligned} S_{\text{вых}}(j\omega) &= \int_{-\infty}^{\infty} u_{\text{вых}}(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) h(t - \tau) e^{-j\omega t} dt d\tau = \\ &= \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) e^{-j\omega \tau} d\tau \int_{-\infty}^{\infty} h(x) e^{-j\omega x} dx = S_{\text{вх}}(j\omega) K(j\omega), \end{aligned} \quad (3.8)$$

где $S_{\text{вх}}(j\omega) = \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) e^{-j\omega \tau} d\tau$ — спектральная плотность входного

сигнала; $K(j\omega) = \int_{-\infty}^{\infty} h(x) e^{-j\omega x} dx$ — коэффициент передачи линей-

ной цепи, связанный с ее импульсной характеристикой прямым преобразованием Фурье. При вычислении двойного интеграла использована замена переменной $(t - \tau)$ на x .

Формула (3.8) часто применяется для анализа линейных цепей. Она содержит простую алгебраическую операцию умножения спектральной плотности входного сигнала на коэффициент передачи цепи.

Пример 3.2. Заданы сигнал $u_{\text{вх}}(t)$, его спектральная плотность $S_{\text{вх}}(t)$ и коэффициент передачи линейной цепи

$$K(j\omega) = e^{-j\omega t_0}.$$

Найти спектральную плотность сигнала на выходе цепи и соответствующую этой плотности временную функцию.

Решение. Применяем формулу (3.8):

$$S_{\text{вых}}(j\omega) = S_{\text{вх}}(j\omega)K(j\omega) = S_{\text{вх}}(j\omega)e^{-j\omega t_0}.$$

Выходной сигнал найдем по формуле обратного преобразования Фурье:

$$\begin{aligned} u_{\text{вых}}(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{вых}}(j\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{вх}}(j\omega) e^{-j\omega t_0} e^{j\omega t} d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{вх}}(j\omega) e^{j\omega(t-t_0)} d\omega = u_{\text{вх}}(t-t_0). \end{aligned}$$

Линейная цепь передает входной сигнал без изменения его формы, но со сдвигом во времени на t_0 . Такая линейная цепь называется *идеальной линией задержки*.

3.2.3. Преобразование дискретного спектра параметрической цепью

При параметрическом преобразовании спектр на выходе существенно отличается от спектра на входе за счет появления новых составляющих.

Если функции $\Phi(t)$ и $u_{\text{вх}}(t)$, входящие в формулу (3.3), разложить в ряды Фурье

$$\Phi(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t), \quad (3.9)$$

$$u_{\text{вх}}(t) = \frac{d_0}{2} + \sum_{k=1}^{\infty} (d_k \cos k\omega_1 t + q_k \sin k\omega_1 t), \quad (3.10)$$

то дискретный спектр сигнала на выходе параметрической цепи можно найти перемножив функцию $\Phi(t)$ на функцию $u_{\text{вх}}(t)$ и выделив из произведения постоянную составляющую, первую, вторую, третью и т.д. гармоники.

Задача упрощается, если хотя бы один из рядов (3.9), (3.10), содержит конечное и сравнительно небольшое число членов.

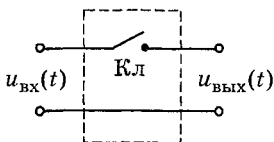


Рис.3.2.
Параметрическая цепь

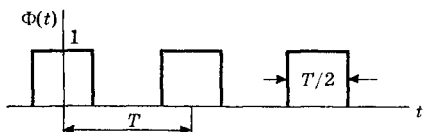


Рис.3.3. Изменение коэффициента передачи параметрической цепи во времени

$$+ \cos(k+1)\omega_1 t] = \frac{A}{\pi} + \frac{A}{2} \cos \omega_1 t + \frac{4A}{3\pi} \cos 2\omega_1 t + \frac{A}{3\pi} \cos 3\omega_1 t + \dots$$

Пример 3.3. В параметрической цепи на рис.3.2 коэффициент передачи изменяется по закону, показанному на рис.3.3. Найти дискретный спектр на выходе параметрической цепи, если частота входного гармонического сигнала $\omega_1 = 2\pi/T$.

Решение. Используя результат, полученный в примере 2.1, запишем

$$\Phi(t) = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{\sin k\pi/2}{k\pi/2} \cos k\omega_1 t,$$

$$u_{\text{вых}}(t) = \Phi(t)A \cos \omega_1 t = \frac{A}{2} \cos \omega_1 t +$$

$$+ \frac{A}{2} \sum_{k=1}^{\infty} \frac{\sin k\pi/2}{k\pi/2} [\cos(k-1)\omega_1 t +$$

3.2.4. Преобразование параметрической цепью спектральной плотности

Пусть для функций $\Phi(t)$ и $u_{\text{вх}}(t)$, входящих в выражение (3.3), можно найти их спектральные плотности. Тогда спектральная плотность на выходе параметрической цепи определяется сверткой (2.40) этих плотностей.

Пример 3.4. Для параметрической цепи, схема которой показана на рис.3.2, найти спектральную плотность сигнала на ее выходе, если на входе действует сигнал $u_{\text{вх}}(t)$, спектральная плотность которого равна $S_{\text{вх}}(j\omega)$. Закон изменения коэффициента передачи показан на рис.3.3. Функция $\Phi(t)$ периодическая, с периодом $T = 1/f_1$.

Решение. Найдем спектральную плотность функции $\Phi(t)$.

Интеграл $\int_{-\infty}^{\infty} |\Phi(t)| dt$ равен ∞ и поэтому условие абсолютной интегрируемости не выполняется. Однако преобразование Фурье для этой функции существует.

Убедимся в этом. Функция $\Phi(t)$ является периодической, определяемой рядом Фурье. Этот ряд, в свою очередь, определяется постоянной составляющей и тригонометрическими функциями типа $\sin x$, $\cos x$, для которых условие абсолютной интегрируемости также не выполняется. Покажем, что для этих функций преобразование Фурье существует.

Считая, что подынтегральная функция $\cos \omega_0 t$ вне интервала $[-\tau/2, \tau/2]$ равна нулю, вычислим предел интеграла при $\tau \rightarrow \infty$:

$$\lim_{\tau \rightarrow \infty} \int_{-\tau/2}^{\tau/2} \cos \omega_0 t e^{-j\omega t} dt = \lim_{\tau \rightarrow \infty} \frac{\tau}{2} \left\{ \frac{\sin[(\omega - \omega_0)\tau/2]}{(\omega - \omega_0)\tau/2} + \frac{\sin(\omega + \omega_0)\tau/2}{(\omega + \omega_0)\tau/2} \right\} = \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)], \quad (3.11)$$

так как

$$\lim_{\tau \rightarrow \infty} \frac{\tau}{2\pi} \frac{\sin[(\omega - \omega_0)\tau/2]}{(\omega - \omega_0)\tau/2} = \delta(\omega - \omega_0).$$

Аналогично убеждаемся в том, что преобразование Фурье существует и для функции $\sin \omega_0 t$:

$$\int_{-\infty}^{\infty} \sin \omega_0 t e^{-j\omega t} dt = j\pi[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]. \quad (3.12)$$

Вычислим предел

$$\lim_{\tau \rightarrow \infty} \int_{-\tau}^{\tau} \frac{a_0}{2} e^{-j\omega t} dt,$$

считая, что подынтегральное выражение равно нулю за пределами интервала $[-\tau, \tau]$.

$$\lim_{\tau \rightarrow \infty} \int_{-\tau}^{\tau} \frac{a_0}{2} e^{-j\omega t} dt = \pi a_0 \lim_{\tau \rightarrow \infty} \frac{\tau}{2\pi} \frac{\sin \omega \tau}{\omega \tau} = \pi a_0 \delta(\omega), \quad (3.13)$$

где

$$\delta(\omega) = \lim_{\tau \rightarrow \infty} \frac{\tau}{2\pi} \frac{\sin \omega \tau}{\omega \tau}.$$

Функция $\Phi(t)$, записанная в виде ряда Фурье,

$$\Phi(t) = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{\sin k\pi/2}{k\pi/2} \cos k\omega_1 t.$$

Спектральную плотность этой функции вычислим с применением формул (3.11) и (3.13)

$$\begin{aligned} S_{\Phi}(j\omega) &= \pi\delta(0) + \pi \sum_{k=1}^{\infty} \frac{\sin k\pi/2}{k\pi/2} [\delta(\omega - k\omega_1) + \delta(\omega + k\omega_1)] = \\ &= \pi \sum_{k=-\infty}^{\infty} \frac{\sin k\pi/2}{k\pi/2} \delta(\omega - k\omega_1). \end{aligned} \quad (3.14)$$

Применяя формулу (2.40), найдем спектральную плотность сигнала на выходе параметрической цепи

$$S_{\text{ВЫХ}}(j\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{ВХ}}(jz) S_{\Phi}[j(\omega - z)] dz = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{ВХ}}(jz) \pi \sum_{k=-\infty}^{\infty} \frac{\sin k\pi/2}{k\pi/2} \times$$

$$\times \delta[(\omega - k\omega_1 - z)] = \frac{1}{2} \sum_{k=-\infty}^{\infty} \frac{\sin k\pi/2}{k\pi/2} S_{\text{вх}}[j(\omega - k\omega_1)]. \quad (3.15)$$

3.2.5. Безынерционное преобразование дискретного спектра

Рассмотрим задачу преобразования спектра нелинейной системой. Эта задача разбивается на две: преобразования спектра безынерционной и преобразования спектра линейной частями систем. При этом нелинейная система разбивается на безынерционную и инерционную части так, как показано на рис.3.4.



Рис.3.4. Структурная схема нелинейной системы:

- 1 — безынерционная часть;
- 2 — линейная (инерционная) часть

Подобное разбиение оказывает оправданным потому, что значительное число преобразований сигналов содержит эти части.

К числу устройств, явно содержащих инерционную и безынерционную части, относятся выпрямители, умножители частоты, модуляторы, детекторы и др.

Математическое описание нелинейной систем содержит таким образом описание безынерционной и линейной частей.

Безынерционная система описывается нелинейным уравнением, которое не содержит явно времени, при этом за изменением входного сигнала $u_{\text{вх}}(t)$ мгновенно следует изменение выходного $u_{\text{вых}}(t)$.

В отличие от линейной, оператор L безынерционной системы не содержит членов, соответствующих операциям дифференцирования, интегрирования и временного сдвига.

Задача безынерционного преобразования спектра сигнала формируется следующим образом. Для заданного спектра на входе и заданном операторе нелинейного преобразования L найти выходной спектр.

При нелинейном преобразовании выходной спектр, также как и при параметрическом, существенно отличается от входного за счет появления новых спектральных составляющих.

Для входного дискретного спектра рассмотрим несколько случаев преобразования.

1. Нелинейный оператор задан полиномом:

$$L[u_{\text{вх}}(t)] = c_0 + c_1 u_{\text{вх}}(t) + c_2 u_{\text{вх}}^2(t) + \dots \quad (3.16)$$

При дискретном спектре входного сигнала, его выражение во временной области записывается или в виде ряда Фурье или в виде суммы некоторого числа гармонических колебаний. Сложность решения задачи определяется количеством удерживаемых членов полинома (3.16) и числом гармонических составляющих входного сигнала. Решение задачи сводится к подстановке выра-

жения для входного сигнала в формулу (3.16) и приведения результата к сумме постоянной и гармонических составляющих.

Пример 3.5. Для заданного входного сигнала

$$u_{\text{вх}}(t) = u_1 \cos(\omega_1 t + \varphi_1) + u_2 \cos(\omega_2 t + \varphi_2) \quad (3.17)$$

и оператора нелинейного преобразования типа (3.16) найти дискретный спектр сигнала на выходе нелинейного безынерционного преобразователя.

Решение. Сигнал (3.17) называется *бигармоническим*. Выходной сигнал получим, подставив формулу (3.17) в (3.16).

$$\begin{aligned} u_{\text{вых}}(t) = & c_0 + c_1[u_1 \cos(\omega_1 t + \varphi_1) + u_2 \cos(\omega_2 t + \varphi_2) + \\ & + c_2[u_1 \cos(\omega_1 t + \varphi_1) + u_2 \cos(\omega_2 t + \varphi_2)]^2 = c_0 + \frac{c_2}{2} u_1^2 + \frac{c_2}{2} u_2^2 + \\ & + c_1 u_1 \cos(\omega_1 t + \varphi_1) + c_1 u_2 \cos(\omega_2 t + \varphi_2) + 0,5 c_2 u_1^2 \cos(2\omega_1 t + 2\varphi_1) + \\ & + 0,5 c_2 u_2^2 \cos(2\omega_2 t + 2\varphi_2) + c_2 u_1 u_2 \cos[(\omega_1 - \omega_2)t + \varphi_1 - \varphi_2] + \\ & + c_2 u_1 u_2 \cos[(\omega_1 + \omega_2)t + \varphi_1 + \varphi_2]. \end{aligned}$$

Проанализируем полученный результат.

На выходе содержится гораздо большее число гармонических составляющих, чем на входе. Кроме того, на выходе присутствует постоянная составляющая, которой нет на входе. Гармонические составляющие на выходе имеют частоты, равные $n\omega_1$, и $m\omega_2$, где n, m — целые числа, а также частоты, равные $n\omega_1 \pm m\omega_2$. Гармонические составляющие с частотами $n\omega_1 \pm m\omega_2$ называются *комбинационными колебаниями*. Число $p = |n| + |m|$ называется *порядком комбинационного колебания*. В рассматриваемом примере $p = 2$.

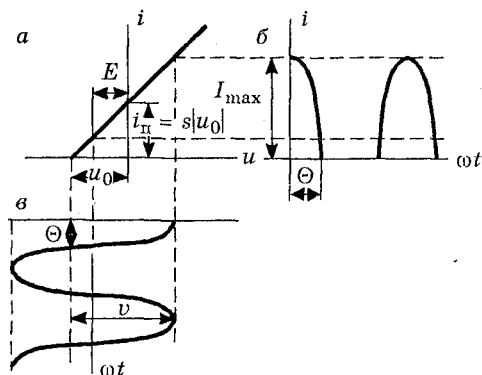


Рис.3.5. Диаграммы, поясняющие метод угла отсечки

Обогащение спектра выходного сигнала новыми гармоническими составляющими является характерным свойством нелинейной цепи.

2. Нелинейный оператор задается в виде *ку-сочно-линейной функции*.

Пусть характеристика нелинейного элемента определяется выражением (рис.3.5).

$$i = \begin{cases} 0 & \text{при } u < u_0, \\ i_n + su & \text{при } u > u_0. \end{cases}$$

На вход нелинейного элемента подается напряжение

$$u = E + V \cos \omega t. \quad (3.18)$$

Амплитуда входного колебания V больше разности $|u_0 - E|$.

При выполнении этого условия во входной цепи происходит ограничение гармонического колебания. Угол θ называется углом отсечки. Импульсы тока на выходе нелинейного элемента определяются следующими уравнениями:

$$i = \begin{cases} i_n + su, & 0 < \omega t < \theta, \\ 0, & \theta < \omega t < 2\pi - \theta, \\ i_n + su & 2\pi - \theta < \omega t < 2\pi. \end{cases} \quad (3.19)$$

Подставляя (3.18) в (3.19), получим

$$i = i_n + sE + sV \cos \omega t, \quad (3.20)$$

$$0 = i_n sE + sV \cos \theta. \quad (3.21)$$

Вычтем из левой части уравнения (3.20) левую часть уравнения (3.21). То же сделаем и с правыми частями. В результате получим

$$i = sV(\cos \omega t - \cos \theta). \quad (3.22)$$

Из (3.22) определяем

$$\cos \theta = -(E - u_0) / V.$$

Максимальное значение тока

$$I_{\max} = sV(1 - \cos \theta).$$

Найдем амплитуду тока первой гармоники, применяя разложение периодической последовательности импульсов (рис.3.5, б) в ряд Фурье (2.1):

$$I_1 = \frac{2}{T} \int_{-T/2}^{T/2} i(\omega t) \cos \omega t dt = \frac{2}{\omega t} \int_{-\pi}^{\pi} i(\varphi) \cos \varphi d\varphi = \frac{sV}{\pi} (\theta - \sin \theta \cos \theta).$$

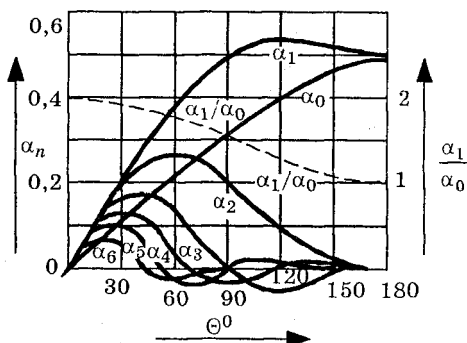


Рис.3.6. Зависимость коэффициентов Берга от угла отсечки

Нормированное значение тока

$$\frac{I_1}{I_{\max}} = \frac{\theta - \sin \theta \cos \theta}{\pi(1 - \cos \theta)} = \alpha_1(\theta).$$

Аналогично находятся отношения $I_0/I_{\max} = \alpha_0(\theta)$, $I_2/I_{\max} = \alpha_2(\theta)$ и т.д. Коэффициенты $\alpha_0(\theta)$, $\alpha_1(\theta)$, $\alpha_2(\theta)$, ... — называются коэффициентами Берга. График зависимостей $\alpha_0(\theta)$, $\alpha_1(\theta)$, $\alpha_2(\theta)$, ... приведен на рис.3.6. Он позволяет найти

значения амплитуд гармоник при заданном угле отсечки.

Пример 3.6. Для заданного значения тока $I_{\max}$ и угла отсечки $\theta = 60^\circ$ найти амплитуды первой и второй гармоник на выходе нелинейного преобразователя.

Решение. Пользуясь графиком для коэффициентов Берга, находим $\alpha_1 = 0,4$, $\alpha_2 = 0,26$.

Амплитуды первой и второй гармоник равны

$$I_1 = I_{\max} \alpha_1 = 0,4 I_{\max}, \quad I_2 = I_{\max} \alpha_2 = 0,26 I_{\max}.$$

3. Нелинейный оператор задан экспоненциальной функцией. Если характеристика нелинейного элемента аппроксимируется экспонентой или экспоненциальным полиномом, то задача анализа упрощается.

Например, если нелинейное уравнение преобразователя равно $e^{a \cos x}$, то это уравнение можно записать в виде следующего ряда Фурье:

$$e^{a \cos x} = I_0(a) + 2I_1(a) \cos x + 2I_2(a) \cos 2x \dots \quad (3.23)$$

Аналогично функцию $e^{a \sin x}$ можно записать в виде ряда

$$e^{a \sin x} = I_0(a) + 2I_1(a) \cos x + 2I_4(a) \cos 4x + 2I_1(a) \sin x + 2I_3(a) \sin 3x + \dots \quad (3.24)$$

Коэффициентами рядов Фурье в выражениях (3.23), (3.24) являются модифицированные функции Бесселя k -го индекса. Это табличные функции, которые приводятся в справочниках по математике.

Пример 3.7. Для заданной характеристики нелинейного элемента

$$i = f(u) = i_n e^{au}$$

и входного напряжения

$$u = V \cos \omega t$$

найти ток на выходе нелинейного элемента.

Решение. Применяем формулу (3.23)

$$i = i_n e^{aV \cos \omega t} = i_n [I_0(aV) + 2I_1(aV) \cos \omega t + 2I_2(aV) \cos 2\omega t + \dots].$$

3.2.6. Безынерционное преобразование спектральной плотности

Перейдем к рассмотрению случая, когда на входе, задана спектральная плотность сигнала. Найдем спектральную плотность на выходе, используя полином (3.16):

$$S_{\text{вых}}(j\omega) = \int_{-\infty}^{\infty} [c_0 + c_1 u_{\text{вх}}(t) + c_2 u_{\text{вх}}^2(t) + \dots] e^{-j\omega t} dt = c_0 \int_{-\infty}^{\infty} e^{-j\omega t} dt +$$

$$\begin{aligned}
& + c_1 \int_{-\infty}^{\infty} u_{\text{вх}}(t) e^{-j\omega t} dt + c_2 \int_{-\infty}^{\infty} u_{\text{вх}}^2(t) e^{-j\omega t} dt + \dots = \\
& = 2\pi c_0 \delta(\omega) + c_1 S_{\text{вх}}(j\omega) + c_2 S_n(j\omega) + \dots,
\end{aligned} \quad (3.25)$$

где $\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega t} dt$ — δ -функция; $S_{\text{вх}}(j\omega) = \int_{-\infty}^{\infty} u_{\text{вх}}(t) e^{-j\omega t} dt$ — непрерывная спектральная плотность входного сигнала; $S_n(j\omega)$ спектральная плотность произведения двух одинаковых сигналов $u(t)$.

Проанализируем полученное выражение.

Спектральная плотность выходного сигнала дискретно-непрерывная, хотя спектральная плотность входного сигнала дискретных составляющих не имеет. Спектральная плотность, соответствующая квадратичному члену полинома (3.16), находится как спектр произведения двух одинаковых входных сигналов. Если бы полином содержал кубический член, то соответствующая спектральная плотность была бы найдена по свертке спектральной плотности входного сигнала $S_{\text{вх}}(j\omega)$ и найденной спектральной плотности произведения $u_{\text{вх}}(t) \cdot u_{\text{вх}}(t)$.

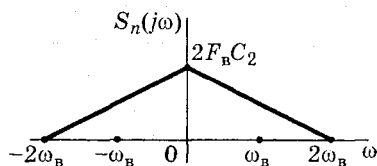


Рис.3.7. Свертка спектров

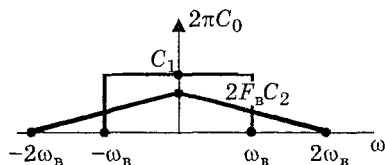


Рис.3.8. Спектральная плотность выходного сигнала

Пример 3.8. Вычислить спектральную плотность на выходе безынерционной части нелинейной системы. Нелинейный оператор задан полиномом второй степени (3.16), а входной сигнал

$$u_{\text{вх}}(t) = 2F_B \frac{\sin \omega_B t}{\omega_B t}.$$

Спектральную плотность сигнала найдем с использованием результата, полученного в примере 2.5.

$$S_{\text{вх}}(j\omega) = 1, \quad -\omega_B \leq \omega \leq \omega_B.$$

Свертка спектров

$$S_{\Pi}(j\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{вх}}(jz) S_{\text{вх}}[j(\omega - z)] dz =$$

$$= \frac{1}{2\pi} \int_{-\omega_B + \omega}^{\omega_B} dz = \frac{1}{2\pi} (2\omega_B + \omega), \quad \omega > 0.$$

График свертки показан на рис.3.7.

На рис.3.8. показаны составляющие спектральной плотности выходного сигнала.

3.3. Преобразование частоты.

Модуляция и детектирование

3.3.1. Преобразование частоты

Под преобразованием частоты понимают процесс дискретного сдвига спектра (спектральной плотности) по оси частот с сохранением его формы.

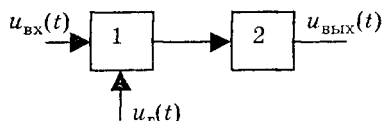


Рис.3.9. Структурная схема преобразователя частоты:

- 1 — перемножитель;
2 — полосовой фильтр

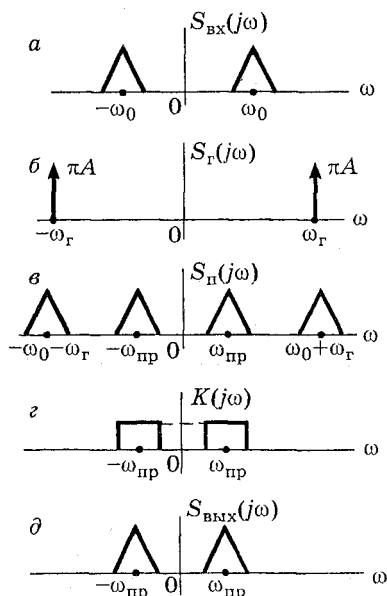


Рис.3.10. Спектральные диаграммы, поясняющие работу преобразователя

Преобразование частоты применяют для размещения спектра сигнала в заданном участке диапазона частот канала связи, а также для повышения чувствительности и избирательности приемников супергетеродинного типа. Принцип преобразования поясняется рис.3.9, 3.10.

Сигнал на входе преобразователя

$$u_{\text{вх}}(t) = u[t, b(t)],$$

зависимый от времени и первичного сигнала $b(t)$, умножается на сигнал гетеродина

$$u_r(t) = A_r \cos(\omega_r t + \varphi_r),$$

а затем фильтруется полосовым фильтром.

Входной сигнал может быть модулированным (непрерывно или дискретно) по амплитуде, фазе, несущей частоте.

Пусть спектральная плотность любого модулированного сигнала состоит из спектральных компонент, сконцентрированных в некоторой полосе около частот $\pm\omega_0$ (рис.3.10, а):

$$\begin{aligned} S_{\text{вх}}(j\omega) &= \int_{-\infty}^{\infty} u_{\text{вх}}(t) e^{-j\omega t} dt = \\ &= S[j(\omega + \omega_0)] + S[j(\omega - \omega_0)]. \end{aligned}$$

Вычислим спектральную плотность сигнала гетеродина, используя выражение (2.18) для δ -функции

$$\begin{aligned}
 S_{\Gamma}(j\omega) &= A \int_{-\infty}^{\infty} \cos(\omega_{\Gamma}t + \varphi_{\Gamma}) e^{-j\omega t} dt = A \int_{-\infty}^{\infty} \frac{e^{j(\omega_{\Gamma}t + \varphi_{\Gamma})} + e^{-j(\omega_{\Gamma}t + \varphi_{\Gamma})}}{2} e^{-j\omega t} dt = \\
 &= \pi A [e^{j\varphi_{\Gamma}} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j(\omega_{\Gamma} - \omega)t} dt + e^{-j\varphi_{\Gamma}} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j(\omega_{\Gamma} - \omega)t} dt] = \\
 &= \pi A [e^{j\varphi_{\Gamma}} \delta(\omega_{\Gamma} - \omega) + e^{-j\varphi_{\Gamma}} \delta(\omega_{\Gamma} + \omega)].
 \end{aligned}$$

Полагая что $\varphi_{\Gamma} = 0$, получим

$$S_{\Gamma}(j\omega) = \pi A [\delta(\omega_{\Gamma} - \omega) + \delta(\omega_{\Gamma} + \omega)]. \quad (3.26)$$

Спектральная плотность гармонического косинусоидального колебания с нулевой начальной фазой (рис.3.10, б) определяется произведением амплитуды этого колебания увеличенной в π раз на сумму двух δ -функций, расположенных в точках частотной оси $\omega = \pm\omega_{\Gamma}$. δ -функции в дальнейшем будем условно изображать стрелками.

Спектральную плотность произведения сигналов вычислим по формуле (2.40):

$$\begin{aligned}
 S_{\Pi}(j\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{вх}}(jz) S_{\Gamma}[j(\omega - z)] dz = \frac{1}{2\pi} \int_{-\infty}^{\infty} \{S[j(z + \omega_0)] + S[j(z - \omega_0)]\} \\
 &\{A\pi[\delta(\omega + \omega_{\Gamma} - z) + \delta(\omega - \omega_{\Gamma} - z)]\} dz = \frac{A}{2} \{S[j(\omega + \omega_{\Gamma} + \omega_0)] + \\
 &+ S[j(\omega + \omega_{\Gamma} + \omega_0)] + S[j(\omega - \omega_{\Gamma} + \omega_0)] + S[j(\omega - \omega_{\Gamma} - \omega_0)]\}, \quad (3.27)
 \end{aligned}$$

где $\omega_{\Gamma} - \omega_0 = \omega_{\text{пр}}$ — промежуточная частота; $S_{\text{вх}}(j\omega)$, $S_{\Gamma}(j\omega)$ — спектральные плотности входного сигнала и гетеродина соответственно.

Поясним вычисление произведения $S_{\Pi}(j\omega)$, используя спектральные диаграммы на рис.3.10.

Спектральная плотность $S_{\Pi}(j\omega)$ вычисляется:

1) перемножением сдвинутых относительно друг друга плотностей $S_{\text{вх}}(jz)$ и $S_{\Gamma}[j(\omega - z)]$ величину ω ;

2) интегрированием результата перемножения.

Сдвиг спектральных плотностей находится из уравнения

$$\omega - z = 0,$$

т.е. $\omega = z$. Этому значению частоты ω соответствует значение спектральной плотности $S_{\Pi}(j\omega)$. Изменяя ω , можно получить все интересующие нас значения этой плотности.

Пусть спектральная плотность $S_{\Gamma}(j\omega)$ равномерно сдвигается вправо относительно $S_{\text{вх}}(j\omega)$ (рис.3.10, а, б, в). Вначале произведение $S_{\text{вх}}(jz) \cdot S_{\Gamma}[j(\omega - z)]$ равно нулю. Затем, при сдвигах близких к $\omega_{\text{пр}} = \omega_{\Gamma} - \omega_0$, это произведение становится неравным нулю,

достигая максимума при $\omega = \omega_{\text{пр}}$. При сдвигах близких к $\omega_{\text{пр}} + 2\omega_0 = \omega_0 + \omega_r$, произведение также не равно нулю.

Если произведение не равно нулю, то и интеграл также не равен нулю. Вычисления существенно упрощаются, если одна из спектральных плотностей представлена δ -функцией или их суммой. В этом случае полезно использовать фильтрующее свойство δ -функции. На рис.3.10 функция $S_{\text{вх}}(j\omega)$ является медленно меняющейся по сравнению с $S_r(j\omega)$. Поэтому интеграл равен функции $S_{\text{вх}}(jz)$ при значении такого аргумента z , при котором δ -функция не равна нулю.

Приравнивая к нулю аргументы δ -функций в выражении

$$\text{Ал}[S(\omega + \omega_r - z) + S(\omega - \omega_r - z)],$$

находим эти значения:

$$z = \omega + \omega_r \text{ и } z = \omega - \omega_r.$$

В спектральной плотности произведения, показанной на рис.3.10, *в*, содержится полезный продукт преобразования — спектральные компоненты вблизи значений промежуточной частоты $\omega = \pm \omega_{\text{пр}}$, а также мешающие компоненты вблизи частот $-\omega_0 - \omega_r, -\omega_0 + \omega_r$.

Полезные компоненты (см. рис.3.10, *в*, *г*) проходят на выход полосового фильтра, а мешающие этим фильтром существенно ослабляются. Спектральные компоненты на выходе полосового фильтра (рис.3.10, *д*) определяются выражением

$$S_{\text{вых}}(j\omega) = S_{\text{п}}(j\omega)K(j\omega) = \frac{A}{2} \{S[j(\omega + \omega_{\text{пр}})] + S[j(\omega - \omega_{\text{пр}})]\}.$$

Они с точностью до постоянного множителя $A/2$ совпадают со спектральными компонентами сигнала на его входе, а спектр преобразованного сигнала группируется около значений частот $\omega = \pm \omega_{\text{пр}}$.

Отметим, что, если спектральные плотности перемножаемых сигналов представлены графически, например так, как на рис.3.10, то можно, не прибегая к вычислению свертки, определить значения частот, около которых группируются спектральные составляющие произведения сигналов. Кроме того, если одна из спектральных плотностей представлена δ -функциями, то можно определить и вид функции $S_{\text{п}}(j\omega)$.

Например, по диаграммам 3.10, *а*, *б* находим, что δ -функция на частоте $\omega = -\omega_r$ сдвинута от максимумов функции $S_{\text{вх}}(j\omega)$ соответственно на $\omega_{\text{пр}} = \omega_r - \omega_0$ и $\omega_0 + \omega_r$. Другая δ -функция на частоте $\omega = \omega_r$ сдвинута от тех же максимумов на $-\omega_{\text{пр}}$ и $-\omega_0 - \omega_r$.

Преобразование частоты используется при модуляции и детектировании сигналов.

3.3.2. Модуляция

На рис.3.11 приведены примеры носителей (переносчиков) сообщений. Гармонический сигнал $u(t, A, \omega_0, \varphi)$ — наиболее часто используемый носитель (рис.3.11, а, е). Он определяется функцией, зависимой от времени t , амплитуды A , частоты ω_0 и фазы φ .

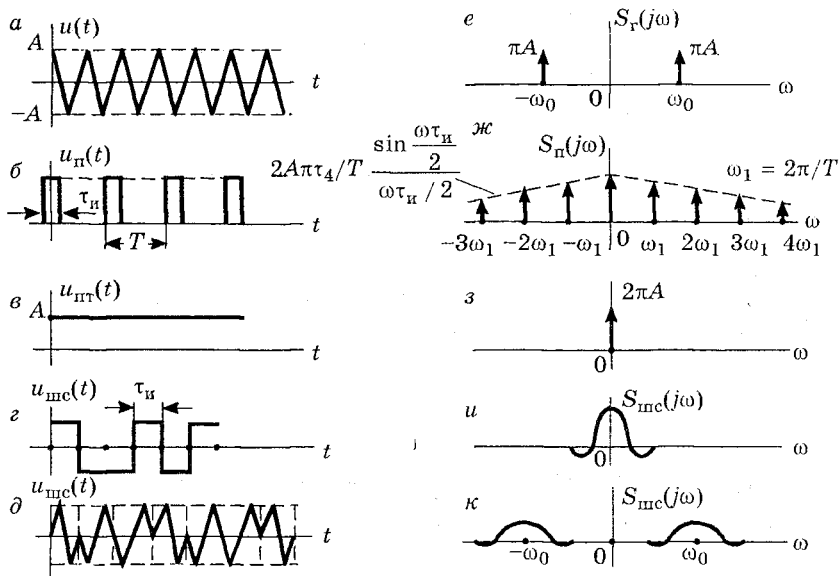


Рис.3.11. Разновидности носителей сообщений

В частотной области гармонический сигнал описывается спектральной плотностью $S_r(j\omega)$, состоящей из двух δ -функций в точках $\omega = \pm\omega_0$. На рис.3.11, б, ж показан другой носитель сообщения — периодическая последовательность видеоимпульсов. Этот носитель описывается функцией, зависимой от времени t , а также от амплитуды A , длительности импульсов $\tau_{\text{и}}$ и периода T , т.е. $u_{\text{п}}(t) = u(t, A, \tau_{\text{и}}, T)$.

Спектральная плотность периодической последовательности импульсов (рис.3.11, ж) состоит из множества δ -функций, расположенных в точках частотной оси $\omega = k\omega_0$, $k = 0, \pm 1, \pm 2, \dots$. Вычисление спектральной плотности приведено в примере 2.4.

Носитель сообщения — постоянный ток описывается постоянной функцией $u_{\text{пт}}(t) = A$. Его спектральная плотность содержит единственную δ -функцию в точке $\omega = 0$ и определяется формулой (3.13.)

В специальных системах связи в качестве носителя сообщений применяются также *сложные шумоподобные сигналы* (ШС).

Известные подклассы этих сигналов можно получить модуляцией направления постоянного тока (рис.3.11, *г*) или модуляцией фазы гармонического носителя (рис.3.11, *д*). Она производится по заранее заданному закону упорядоченной двоичной последовательностью нулей и единиц. Методы дискретной модуляции постоянного тока и гармонического сигнала этой двоичной последовательностью рассмотрены ниже, в п.3.7.3. Это так называемая *псевдослучайная последовательность* (слово "псевдо" означает почти). Математическая модель ШС задается функцией, зависимой от времени, формы единичного импульса $u_0(t)$ и закона формирования двоичной последовательности $q(t)$:

$$u_{\text{шс}}(t) = f[t, u_0(t), q(t)].$$

Соответствующие временным функциям (рис.3.11, *г*, *д*) спектры ШС приведены на рис.3.11, *и*, *к*.

Модулирующий сигнал (непрерывный или дискретный) и параметр носителя определяют вид модуляции.

Классификация различных видов модуляции приведена на рис.3.12.

Модуляцией гармонического носителя непрерывным первичным сигналом получают амплитудную (АМ), амплитудную с одной боковой полосой (АМ ОБП), частотную (ЧМ) и фазовую модуляцию (ФМ). Модуляцией гармонического носителя дискретным модулирующим сигналом (кодовой комбинацией) получают дискретную АМ, дискретную ЧМ, дискретную ФМ, дискретную ОФМ, дискретную ФМ с ОБП, дискретную ОФМ с ОБП, многопозиционную (многократную) фазовую модуляцию, амплитудно-фазовую многократную модуляцию, амплитудно-фазовую многократную модуляцию с ОБП.

Непрерывной модуляцией амплитуды, длительности импульсов, периода повторения периодической последовательности импульсов получают соответственно амплитудно-импульсную АИМ, широтно-импульсную модуляцию (ШИМ) и, связанную с ШИМ фазово-импульсную модуляцию (ФИМ). Частотно-импульсную модуляцию (ЧИМ) получают в результате воздействия на частоту следования импульсов.

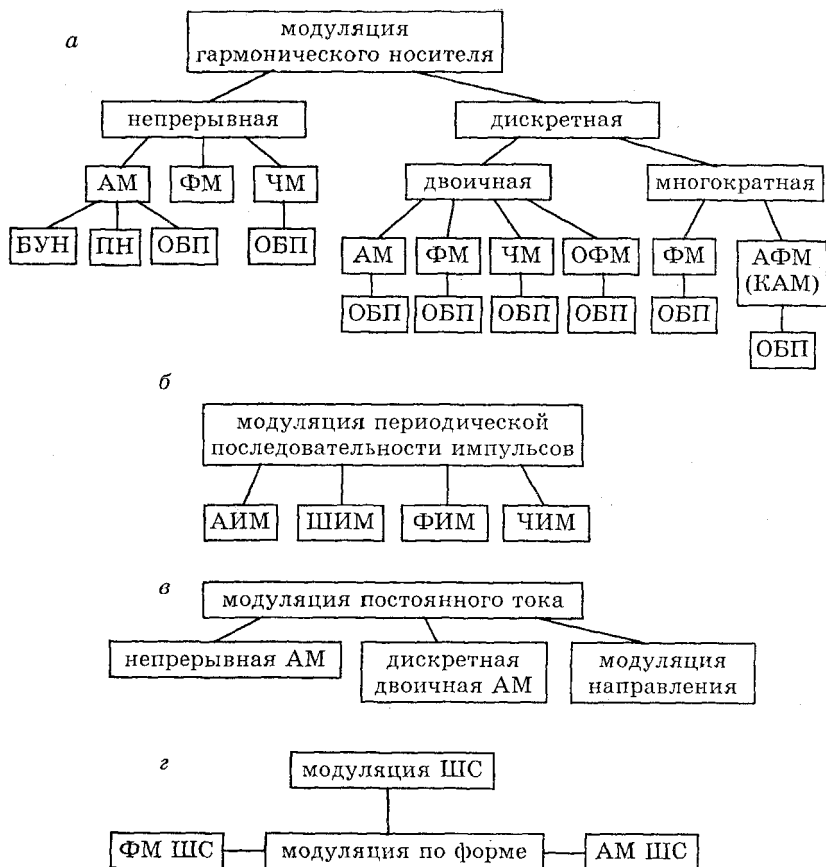


Рис.3.12. Виды модуляций

Изменением амплитуды постоянного тока получают непрерывную и дискретную АМ, а коммутацией знака тока — модуляцию его направления. Модуляцией ШС двоичной последовательностью, определяемой передаваемым кодовым словом над алфавитом 0,1, получают АМ, а определяемой кодовым словом над алфавитом $-1,1$ — ФМ. Последний вид модуляции называется инверсной модуляцией ШС. Если при изменении кодового слова меняется форма ШС, то такая модуляция называется модуляцией ШС по форме. АМ и ФМ ШС являются частными случаями модуляции ШС по форме.

Вид модуляции и связанные с ней преобразования сигналов определяют математическую модель модулятора.

Задача модуляции формулируется следующим образом. Нужно так выбрать оператор L , чтобы его действие привело к формированию требуемого модулированного сигнала:

$$u[t, b(t)] = L[u_n(t, a_1, a_2, \dots, a_n)],$$

где $u_n(t, \bar{a}]$ — носитель сообщения; $a_i, i = 1, 2, \dots, N$ — параметр носителя, который модулируется; $b(t)$ — модулирующий сигнал.

Решение задачи связано с анализом модулированных сигналов и их спектров (спектральных плотностей). С целью отбора допустимых решений вначале задают структуру модулятора (оператор L), входные сигналы и их спектры (спектральные плотности), затем проводят анализ полученных математических выражений для выходного сигнала и его спектра (спектральной плотности).

Спектральная плотность реального модулированного сигнала является случайной функцией

$$S(j\omega) = \int_{-\infty}^{\infty} u[t, b(t)] e^{-j\omega t} dt,$$

так как модулирующий сигнал $b(t)$ является случайным.

Ее вычисление затруднено. Поэтому вместо случайных используют детерминированные функции. Такими функциями являются, например, гармонические функции $\sin x, \cos x$.

3.3.3. Детектирование

Детектированием называется процесс выделения модулирующего сигнала из принимаемого. Название детектирования определяется модулируемым параметром сигнала.

Задача детектирования формулируется следующим образом. Для заданного модулированного сигнала $u[t, b(t)]$, найти преобразование, однозначно определяющее модулирующий сигнал

$$b(t) = L\{u[t, b(t)]\}.$$

Под L здесь следует понимать не только линейное и нелинейное, но и параметрическое преобразование.

Основным методом исследования процессов, происходящих при детектировании, является *анализ*. При этом методе вначале задают оператор преобразования L (или структурную схему детектора), а затем решают задачи преобразования сигналов и их спектров. Полученные результаты анализируют. На этой основе принимают решения о соответствии выбранной схемы и ее параметров заданным требованиям.

При выборе модуляции и демодуляции наряду с задачами анализа решают и *задачи синтеза*. Задачи синтеза сложнее задач анализа. Они формулируются следующим образом. Для заданных сигналов (спектров) на входе и выходе цепи требуется найти

схему и параметры ее элементов. Из-за отсутствия достаточно разработанных математических методов не все задачи синтеза удается довести до конца. Основные понятия синтеза систем и их отдельных частей даны в восьмой главе. Примеры синтеза приведены в шестой главе.

3.4. Амплитудная модуляция гармонического переносчика

3.4.1. Модулирующие сигналы

Физические процессы, происходящие в модуляторах АМ сигналов, аналогичны процессам в преобразователях частоты. Поэтому амплитудную модуляцию можно считать частным случаем этого преобразования.

Модуляцией гармонического носителя непрерывным модулирующим сигналом получают *непрерывный АМ сигнал*, а модуляцией дискретным сигналом — *дискретный АМ сигнал*. Наиболее часто встречающиеся разновидности модулирующих сигналов приведены на рис.3.13.

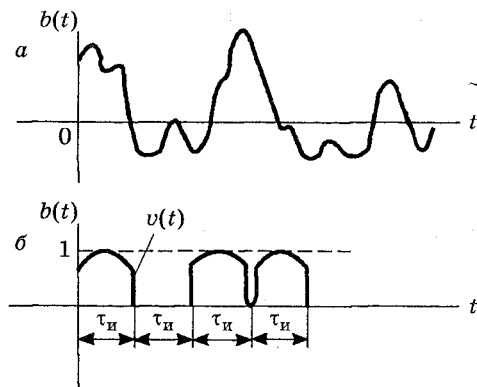


Рис.3.13. Наиболее часто встречающиеся разновидности модулирующих сигналов при АМ

Амплитуда непрерывного сигнала (рис.3.13, а) является случайной величиной и в каждый момент времени принимает одно значение из определенного непрерывного множества. Амплитуда дискретного сигнала определяется или функцией $v(t)$ или равна нулю (рис.3.13, б) в пределах единичного интервала времени. Появление импульса в заданном единичном интервале случайно и определено соответствующим элементом кодового слова — носителя передаваемой информации. Элементы кодового слова при АМ равны нулю или единице, т.е. принадлежат двоичному алфавиту. Поэтому эта разновидность дискретной модуляции называется двоичной.

Для упрощения анализа случайные модулирующие сигналы заменяют детерминированными. При непрерывной АМ модулирующий сигнал часто берут равным

Для упрощения анализа случайные модулирующие сигналы заменяют детерминированными. При непрерывной АМ модулирующий сигнал часто берут равным

$$b(t) = b_0 \cos \Omega t,$$

где b_0 — амплитуда, Ω — угловая частота.

При дискретной модуляции случайное кодовое слово заменяют некоторым заранее заданным двоичным числом.

3.4.2. Амплитудная непрерывная модуляция с подавленной несущей

Эта модуляция отличается тем, что в спектре модулированного сигнала нет спектральной составляющей на несущей частоте.

Модулированный сигнал определяется произведением модулирующего сигнала на гармонический сигнал-переносчик:

$$u_{\text{АМ-ПН}}(t) = b(t) \cos \omega_0 t. \quad (3.28)$$

Структурная схема модулятора весьма проста и состоит из перемножителя.

Найдем спектральные плотности сигналов на входе и выходе модулятора, используя прямое преобразование Фурье. Спектральная плотность модулирующего сигнала (3.14, а, б)

$$S(j\omega) = \int_{-\infty}^{\infty} b(t) e^{-j\omega t} dt.$$

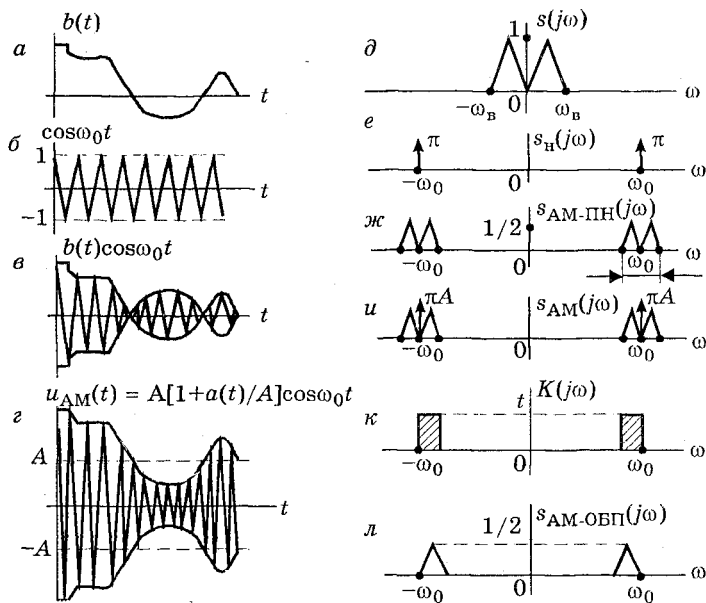


Рис.3.14. Диаграммы, поясняющие процессы в модуляторах АМ сигналов

Спектральная плотность гармонического сигнала — носителя сообщения (рис.3.14, б, в)

$$S_H(y'co) = \int_{-\infty}^{\infty} \cos a^0 t e^{-i\omega t} dt = \pi [5(\omega_0 - \omega) + 5(\omega_0 + \omega)], \quad (3.29)$$

где $5(\omega)$ — дельта функция, определяемая формулой (2.18).

Спектральная плотность амплитудно-модулированного с подавляющей несущей (АМ-ПН) сигнала (рис.3.14, в, ж) на выходе модулятора

$$S_{AM-ПН}(\omega) = \int_{-\infty}^{\infty} b(t) e^{-i\omega t} dt = \int_{-\infty}^{\infty} b(t) e^{-i(\omega_0 \pm \omega) t} dt + i \int_{-\infty}^{\infty} b(t) e^{-i\omega_0 t} dt + i \int_{-\infty}^{\infty} b(t) e^{-i\omega t} dt, \quad (3.30)$$

где $S(\omega)$ — спектральная плотность модулирующего сигнала.

Спектральная плотность АМ-ПН сигнала состоит из двух частей, сдвинутых по оси частот относительно спектральной плотности модулирующего сигнала на величину $\omega_0 = \pm \omega_0$.

Амплитуды спектральных компонент АМ-ПН сигнала, как это следует из выражения (3.30), в два раза меньше амплитуд спектральных компонент модулирующего сигнала. Поэтому, если спектральная составляющая на нулевой частоте модулирующего сигнала равна $5(0)$, то и составляющие спектра АМ-ПН сигнала на частотах $\omega = \pm\omega_0$ будут равны $5(0)/2$. Это означает, что уровень несущей в исследуемом модуляторе определяется уровнем постоянной составляющей модулирующего сигнала. Если эта составляющая равна нулю, то и спектральные составляющие на частотах $\omega = \pm\omega_0$ будут также равны нулю.

При АМ-ПН спектральная плотность модулирующего сигнала как бы переносится из области нулевой частоты в область более высоких частот, равных $\pm\omega_0$. С уменьшением амплитуд спектральных компонент в 2 раза без какого-либо растяжения или сжатия ее по оси частот. Следовательно, ширина спектра АМ-ПН сигнала определяется шириной спектра модулирующего сигнала.

Под *шириной спектра* здесь и в дальнейшем будем понимать частотный интервал, расположенный в области положительных значений частот, в пределах которого модуль спектральной плотности больше некоторого заданного уровня. Например, таким может быть уровень, численно равный или нулю или $0,1 L_{S_{max}}$,

где $|S_{\max}|$ — максимальное значение модуля спектральной плотности.

Задание уровня означает, что часть спектральных составляющих при расчетах не учитывается. На практике считается допустимым отбрасывать высшие частотные составляющие спектральной плотности, суммарная энергия которых не превышает небольшую заданную часть от общей энергии сигнала. Если заданный уровень равен нулю, то ширина спектра АМ-ПН сигнала равна $2\omega_b$, т.е. в 2 раза больше ширины спектра модулирующего сигнала (рис.3.14, ж).

3.4.3. Амплитудная непрерывная модуляция с большим уровнем несущей

Этот АМ сигнал равен сумме АМ-ПН сигнала и сигнала несущей $A\cos\omega_0 t$:

$$\begin{aligned} u_{AM}(t) &= u_{AM-ПН}(t) + A\cos\omega_0 t = b(t)\cos\omega_0 t + A\cos\omega_0 t = \\ &= A[1 + b(t)/A]\cos\omega_0 t. \end{aligned} \quad (3.31)$$

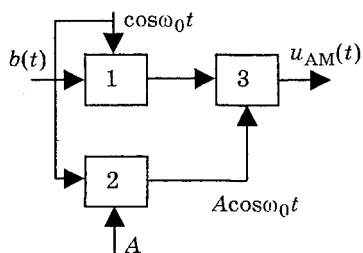


Рис.3.15. Модулятор АМ-сигнала:

1, 2 — умножители;
3 — сумматор

Функция (3.31) определяет структуру модулятора АМ — сигнала (рис.3.15), который содержит два умножителя и сумматор.

Спектральная плотность АМ-сигнала

$$\begin{aligned} S_{AM}(j\omega) &= A \int_{-\infty}^{\infty} [1 + b(t)/A]\cos\omega_0 t e^{-j\omega t} dt = \\ &= S_H(j\omega) + S_{AM-ПН}(j\omega), \end{aligned}$$

где $S_H(j\omega)$ — спектральная плотность несущего колебания; $S_{AM-ПН}(j\omega)$ — спектральная плот-

ность АМ-ПН сигнала.

Спектральная плотность состоит из двух составляющих: спектральной плотности несущего колебания и спектральной плотности АМ-ПН сигнала (рис.3.14 г, и). Поэтому спектральная плотность АМ-сигнала сохраняет все свойства АМ-ПН сигнала. В частности ширина спектра АМ сигнала определяется шириной спектра модулирующего сигнала и равна $2\omega_b$, где ω_b — ширина спектра модулирующего сигнала.

Выберем в качестве модулирующего сигнала простой гармонический сигнал

$$b(t) = b_0 \cos(\Omega t + \theta),$$

где b_0 — амплитуда, Ω — частота, θ — начальная фаза.

Примем также, что начальная фаза гармонического носителя равна φ . С учетом этого, выражение для АМ сигнала (3.31) перепишем в виде

$$\begin{aligned} u_{\text{AM}}(t) &= A[1 + \frac{b_0}{A} \cos(\Omega t + \theta)] \cos(\omega_0 t + \varphi) = A[1 + \\ &+ m \cos(\Omega t + \theta)] \cos(\omega_0 t + \varphi) = A \cos(\omega_0 t + \varphi) + \\ &+ \frac{Am}{2} \cos[(\omega_0 + \Omega)t + \varphi_0 + \theta_0] + \frac{Am}{2} \cos[(\omega_0 - \Omega)t + \varphi_0 - \theta_0], \end{aligned} \quad (3.32)$$

где $m = b_0/A$.

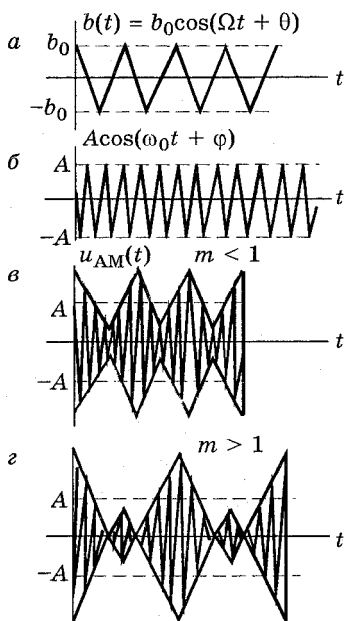


Рис.3.16. Диаграммы, поясняющие работу амплитудного модулятора

Число m называется *индексом амплитудной модуляции*. Индекс амплитудной модуляции выбирают в пределах от 0 до 1. Если этот индекс выбирать больше единицы, то в модуляторе возникнут искажения.

Этот режим называется *перемодуляцией*. Работа АМ модулятора поясняется диаграммами на рис.3.16 и рис.3.17. На рис.3.16, а приведен модулирующий косинусоидальный сигнал, а его спектральная плотность — на рис.3.17, а. На рис.3.16, б показан носитель сообщения. Его спектральная плотность приведена на рис.3.17, б. Модулированный сигнал и его спектральная плотность приведены соответственно на рис.3.16, в и рис.3.17, в.

Для АМ сигнала, записанного в виде (3.32), мощность несущего колебания $P_n = A^2/2$. Мощность боковых спектральных компонент $P_6 = A^2 m^2/8$. Общая мощность АМ сигнала

$$P = \frac{A^2}{2} (1 + m^2/2).$$

Процентное содержание мощности боковых спектральных компонент определяется отношением:

$$\eta(\%) = \frac{2P_6}{P} 100\% = \frac{m^2}{2 + m^2} 100\%.$$

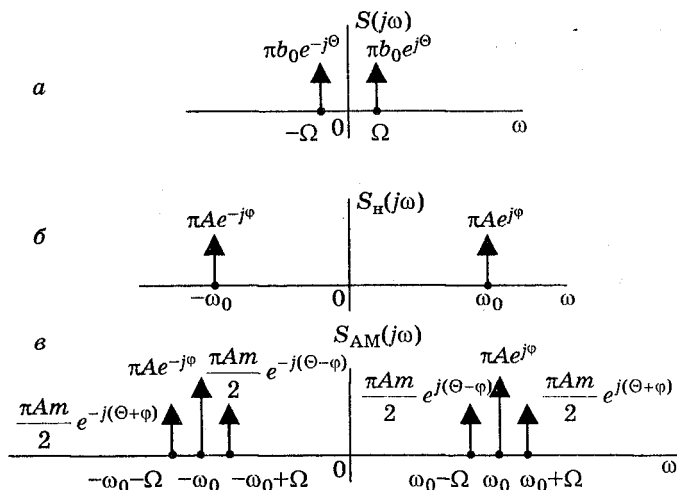


Рис.3.17. Диаграммы, поясняющие работу амплитудного модулятора

При $m = 1$ $\eta \cong 33\%$, а при $m = 0,5$ $\eta \cong 10\%$.

При коэффициенте $m = 1$ около 67% мощности бесполезно тратится на передачу несущей.

Несмотря на этот недостаток амплитудная модуляция с большим уровнем несущей нашла широкое применение в радиовещании.

3.4.4. Дискретная двоичная амплитудная модуляция

Дискретный двоичный амплитудно-модулированный (АМ) сигнал, полученный модуляцией гармонического носителя, пока-

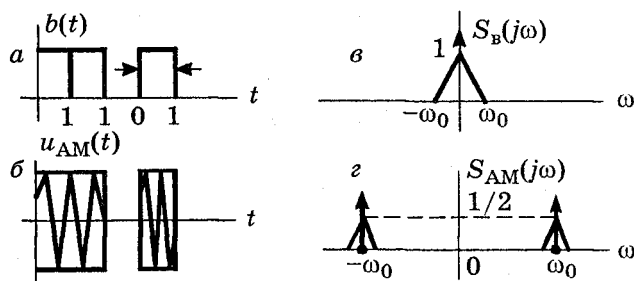


Рис.3.18. Двоичный дискретный сигнал

зан на рис.3.18, б. Этот сигнал определяется произведением дискретного модулирующего сигнала на сигнал переносчик сообщения. Энергия двоичного АМ сигнала в пределах некоторого единичного интервала или не равна или равна нулю. Такие сигналы называются сигналами с *пассивной паузой*.

Модулирующий дискретный двоичный сигнал

$$b(t) = \sum_{k=1}^n b_k v[t - (k-1)\tau_{\text{и}}], \quad (3.33)$$

где b_k k -й элемент двоичного кодового слова, равный 1 или 0;

$v(t)$ — форма единичного элемента модулирующего сигнала.

При замене случайного кодового слова на двоичное число нетрудно вычислить спектральную плотность полученного таким образом модулирующего сигнала.

$$\begin{aligned} S_B(j\omega) &= \int_{-\infty}^{\infty} b(t) e^{-j\omega t} dt = \sum_{k=1}^n b_k \int_{-\infty}^{\infty} v[t - (k-1)\tau_{\text{и}}] e^{-j\omega t} dt = \\ &= S_0(j\omega) \sum_{k=1}^n b_k e^{-j\omega(k-1)\tau_{\text{и}}}, \end{aligned} \quad (3.34)$$

где b_k — двоичное число; $S_0(j\omega) = \int_{-\infty}^{\infty} v(t) e^{-j\omega t} dt$ — спектральная

плотность единичного элемента.

Для прямоугольного единичного элемента

$$v(t) = \begin{cases} V, & \tau_{\text{и}} \geq t \geq 0; \\ 0, & 0 > t > \tau_{\text{и}} \end{cases}$$

спектральная плотность

$$S_0(j\omega) = V\tau_{\text{и}} \frac{\sin \frac{\omega\tau_{\text{и}}}{2}}{\frac{\omega\tau_{\text{и}}}{2}} e^{-j\frac{\omega\tau_{\text{и}}}{2}}. \quad (3.35)$$

Ширина спектральной плотности определяется значением частоты, соответствующей первому нулю этой спектральной плотности: $\Delta\omega = 2\pi/\tau_{\text{и}}$.

Спектральная плотность двоичного дискретного АМ сигнала определяется сверткой соответствующих спектральных плотностей:

$$\begin{aligned} S_{\text{AM}}(j\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_B(jz) S_{\text{H}}[j(\omega - z)] dz = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(jz) \sum_{k=1}^n b_k e^{-jz(k-1)\tau_{\text{и}}} \pi A [\delta(\omega - z + \omega_0) + \delta(\omega - z - \omega_0)] dz = \end{aligned}$$

$$\begin{aligned}
&= \frac{A}{2} \{ S_0[j(\omega + \omega_0)] \sum_{k=1}^n b_k e^{-j(\omega + \omega_0)(k-1)\tau_n} + \\
&\quad + S_0[j(\omega - \omega_0)] \sum_{k=1}^n b_k e^{-j(\omega - \omega_0)(k-1)\tau_n} \}, \quad (3.36)
\end{aligned}$$

где спектральная плотность носителя сообщения $S_H(j\omega)$ определяется формулой, аналогичной (3.26), а спектральная плотность модулирующей функции — формулой (3.34).

Также, как и при непрерывной АМ, спектральная плотность двоичного дискретного АМ сигнала определяется спектральной плотностью двоичного дискретного модулирующего сигнала.

При преобразовании этого сигнала его спектральная плотность переносится из области нижних в область высоких частот без искажений спектральных составляющих.

Для прямоугольного единичного элемента ширина спектра модулирующего сигнала равна $\Delta\omega = 2\pi/\pi_n$, а ширина спектра модулированного сигнала равна $\Delta\omega = 4\pi/\pi_n$.

3.4.5. Амплитудная модуляция с одной боковой полосой

Амплитудную модуляцию с одной боковой полосой (АМ-ОБП) получают полосовой фильтрацией непрерывного или дискретного АМ сигнала.

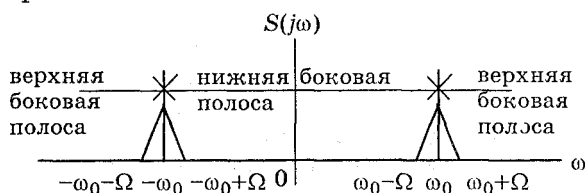


Рис.3.19. Спектральная плотность АМ сигнала

диапазоне частот от $\omega = -\omega_0$ до $\omega = +\omega_0$ (рис.3.19), а с верхней боковой полосой расположены в диапазоне частот от $\omega = -\infty$ до $\omega = -\omega_0$ и от $\omega = \omega_0$ до $\omega = +\infty$.

АМ ОБП сигнал называется также *однополосным АМ сигналом*. Спектральные компоненты АМ ОБП сигнала с нижней боковой полосой расположены в диа-



Рис.3.20. Модулятор АМ-ОБП сигнала:

- 1 — модулятор АМ сигнала;
2 — полосовой фильтр

Структурная схема модулятора АМ-ОБП сигнала включает модулятор АМ сигнала и полосовой фильтр, настроенный на нижнюю или верхнюю полосу (рис.3.20).

Например, спектральная

плотность непрерывного АМ-ОБП сигнала (см. рис.3.14)

$$S_{\text{АМ-ОБП}}(j\omega) = S_{\text{АМ}}(j\omega) \cdot K(j\omega) = [S_{\text{н}}(j\omega) + S_{\text{АМ-ПН}}(j\omega)]K(j\omega), \quad (3.27)$$

где $K(j\omega)$ — коэффициент передачи полосового фильтра.

Подставим в (3.37) выражение (3.30) для спектральной плотности АМ-ПН сигнала. В результате получим:

$$S_{\text{АМ-ОБП}}(j\omega) = \{S_{\text{н}}(j\omega) + \frac{1}{2}S[j(\omega_0 - \omega)] + \frac{1}{2}S[j(\omega_0 + \omega)]\}K(j\omega). \quad (3.38)$$

Кроме того, каждую составляющую спектральной плотности АМ-ПН сигнала в (3.38) запишем в виде спектральных составляющих АМ сигналов с нижней и верхней боковыми полосами:

$$S[j(\omega_0 - \omega)] = S_{\text{нб}}[j(\omega_0 - \omega)] + S_{\text{вб}}[j(\omega_0 - \omega)];$$

$$S[j(\omega_0 + \omega)] = S_{\text{нб}}[j(\omega_0 + \omega)] + S_{\text{вб}}[j(\omega_0 + \omega)].$$

Спектральная плотность АМ-ОБП сигнала

$$\begin{aligned} S_{\text{АМ-ОБП}}(j\omega) = & \{S_{\text{н}}(j\omega) + \frac{1}{2}S_{\text{нб}}[j(\omega_0 - \omega)] + \frac{1}{2}S_{\text{вб}}[j(\omega_0 - \omega)] + \\ & + \frac{1}{2}S_{\text{нб}}[j(\omega_0 + \omega)] + \frac{1}{2}S_{\text{вб}}[j(\omega_0 + \omega)]\}K(j\omega), \end{aligned}$$

Полагая, что полосовой фильтр задерживает спектральные компоненты на несущей частоте и в области высоких частот, найдем:

$$S_{\text{АМ-ОБП}}(j\omega) = \frac{1}{2}\{S_{\text{нб}}[j(\omega_0 - \omega)] + S_{\text{нб}}[j(\omega_0 + \omega)]\}. \quad (3.39)$$

Спектральная плотность АМ-ОБП сигнала содержит часть спектральных компонент АМ сигнала — его нижнюю полосу. Амплитуды спектральных составляющих АМ ОБП сигнала в 2 раза меньше амплитуд спектральных компонент модулирующего сигнала и группируются около значений частот $\omega = \pm\omega_0$. Ширина спектра АМ ОБП сигнала равна ширине спектра модулирующего сигнала, т.е. в 2 раза меньше ширины АМ и АМ-ПН сигналов. При настройке полосового фильтра на верхние боковые полосы аналогично получают АМ ОБП сигнал с верхней боковой полосой.

Сигналы АМ-ОБП сравнительно трудно получить из-за сложной реализации полосового фильтра. Он должен иметь коэффициент передачи с очень резкими переходами из области прозрачности, в которой $|K(j\omega)| \cong 1$ в область непрозрачности, где $|K(j\omega)| = 0$ (см. рис.3.14, κ). Поэтому на практике нашли применение АМ ОБП сигналы с частичным подавлением боковых спектральных компонент. Это подавление позволяет ослабить требования к коэффициенту передачи полосового фильтра и существенно упростить его.

Диаграммы, поясняющие принцип формирования сигнала с частичным подавлением боковых спектральных компонент, приведены на рис.3.21. На рис.3.21, а приведена спектральная плотность АМ-сигнала, а на рис.3.21, б — коэффициент передачи полосового фильтра. Спектральная плотность АМ-ОБП сигнала с частичным подавлением нижних спектральных компонент (рис.3.21, д) равна сумме отфильтрованных спектральных компонент нижней (рис.3.21, в) и верхней (рис.3.21, з) боковых полос АМ сигнала.

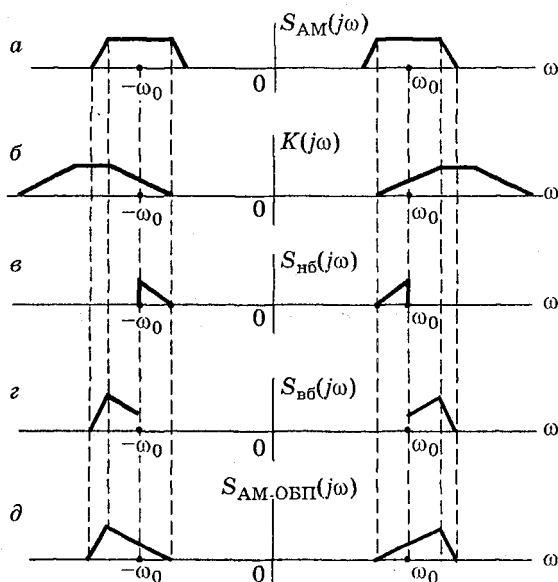


Рис.3.21. АМ-ОБП сигнал с частичным подавлением одной боковой полосы

Амплитудно-частотная характеристика полосового фильтра (рис.3.21, б) не имеет резких переходов и выбирается так, чтобы спектральная плотность сигнала на выходе полосового фильтра (рис.3.21, д) удовлетворяла условию:

$$S_{нб}[j(-\omega_0 + \omega)] + S_{вб}[j(\omega_0 + \omega)] = cS(j\omega), \quad (3.40)$$

где $S[j\omega]$ — спектральная плотность модулирующего сигнала; c — постоянная величина.

Необходимость выполнения условия (3.40) обоснована ниже при анализе работы когерентного детектора.

При этом доказывается, что, если условие (3.40) не выполняется, то при когерентном детектировании возникают искажения.

3.4.6. Векторное представление непрерывных АМ сигналов

Кроме временного и спектрального описания непрерывных АМ сигналов полезно использовать и их векторное представление (рис.3.22).

Непрерывный АМ сигнал, представленный комплексной функцией времени является вектором ОВ на рис.3.22, а.

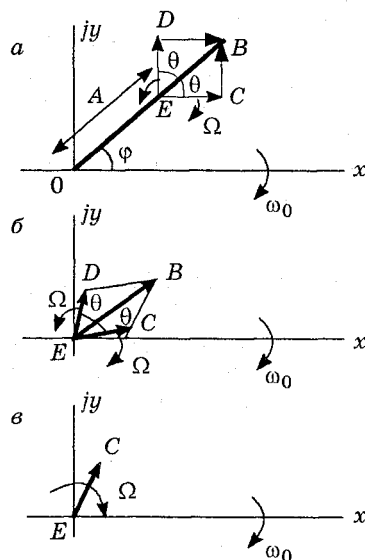


Рис.3.22. Векторные диаграммы непрерывных АМ сигналов

EC, вращающийся по отношению к точке E с угловой скоростью Ω . Временное выражение для АМ-ПН ОБП сигнала определяется третьим слагаемым выражения (3.32).

Действительная часть этой функции является АМ сигналом и записывается в виде (3.32). Таким образом этот сигнал равен проекции вектора OB на ось x. Векторы EC, ED (рис.3.22, а) вращаются с угловой скоростью Ω относительно точки E по направлениям, показанным стрелками. Вектор OB неподвижен на плоскости, а оси x, y вращаются относительно этого вектора с угловой скоростью ω_0 .

Векторная диаграмма для АМ-ПН сигнала приведена на рис.3.22, б. Она отличается от векторной диаграммы АМ сигнала тем, что вектор OE равен нулю. Временное выражение для АМ-ПН сигнала равно сумме второго и третьего слагаемых выражения (3.32). Векторная диаграмма АМ-ОБП сигнала с нижней боковой полосой приведена на рис.3.22, в. Она содержит всего один вектор

3.5. Частотная модуляция непрерывного гармонического переносчика

3.5.1. Непрерывный частотно-модулированный сигнал с малым индексом модуляции

Запишем сигнал в виде комплексной функции времени:

$$\dot{W}(t) = \vartheta(t)e^{j\theta(t)}, \quad (3.41)$$

где $\vartheta(t)$ — огибающая; $\theta(t)$ мгновенная фаза сигнала. Действительная часть выражения (3.41)

$$\text{Re } \dot{W}(t) = \vartheta(t) \cos \theta(t)$$

является обобщением косинусоидального сигнала с изменяющейся во времени амплитудой и фазой.

Производная от мгновенной фазы определяет *мгновенную частоту* этого сигнала:

$$\omega(t) = d\theta(t) / dt.$$

Соответственно, мгновенная частота определяет и *мгновенную фазу*:

$$\theta(t) = \int \omega(t) dt.$$

Модуляция, при которой фаза $\theta(t)$ изменяется в зависимости от модулирующего сигнала $b(t)$, называется *угловой*.

Пусть при передаче информации частота сигнала изменяется по закону:

$$\omega(t) = \omega_0 + 2\pi K b(t),$$

где K — коэффициент пропорциональности с размерностью Гц/ В.

В этом случае фаза

$$\theta(t) = \int \omega(t) dt = \omega_0 t + 2\pi K \int b(t) dt. \quad (3.42)$$

Разновидность угловой модуляции, при которой частота сигнала изменяется пропорционально значению модулирующего сигнала, называется *частотной модуляцией* (ЧМ).

Если при передаче информации мгновенная фаза изменяется в соответствии с выражением

$$\theta(t) = \omega_0 t + 2\pi K_{\text{фм}} b(t), \quad (3.43)$$

то, эта разновидность угловой модуляции называется *фазовой модуляцией* (ФМ).

Огибающая сигнала $\mathcal{Q}(t)$ при угловых способах модуляции остается постоянной величиной. Обозначим ее буквой A . Выражения для ЧМ и ФМ сигналов запишем в виде:

$$u_{\text{чм}}(t) = A \cos[\omega_0 t + 2\pi K \int b(t) dt], \quad u_{\text{фм}}(t) = A \cos[\omega_0 t + 2\pi K_{\text{фм}} b(t) dt].$$

Сравнивая их, находим, что различие между ЧМ и ФМ сигналами незначительно.

Исследуем ЧМ сигнал при условии, что модулирующий сигнал $b(t) = b_0 \cos \Omega t$.

Его фазу определим из (3.42), полагая, что

$$\omega(t) = \omega_0 + 2\pi K b_0 \cos \Omega t = \omega_0 + \Delta \omega \cos \Omega t.$$

где $\Delta \omega = 2\pi K b_0$ — *девиация частоты*.

$$\theta(t) = \omega_0 t + \frac{2\pi K b_0}{\Omega} \sin \Omega t = \omega_0 t + \frac{\Delta \omega}{\Omega} \sin \Omega t = \omega_0 t + m_{\text{чм}} \sin \Omega t, \quad (3.44)$$

где $m_{\text{чм}} = \Delta \omega / \Omega$ — *индекс частотной модуляции*.

Подставив (3.44) в (3.41), получим выражение для ЧМ сигнала в комплексной форме

$$\dot{W}_{\text{чм}}(t) = A e^{j\omega_0 t} e^{jm_{\text{чм}} \sin \Omega t}. \quad (3.45)$$

При малом индексе модуляции ($m_{\text{чм}} \ll 1$) используем разложение

$$e^x \cong 1 + x,$$

и запишем выражение для ЧМ сигнала в виде

$$\begin{aligned} W_{\text{ЧМ}}(t) &= A[\cos \omega_0 t + j \sin \omega_0 t][1 + jm_{\text{ЧМ}} \sin \Omega t] = \\ &= A[\cos \omega_0 t - m_{\text{ЧМ}} \sin \Omega t \sin \omega_0 t + j(\sin \omega_0 t + m_{\text{ЧМ}} \sin \Omega t \cos \omega_0 t)]. \end{aligned}$$

Действительная часть комплексной функции определяет ЧМ сигнал во временной области:

$$\begin{aligned} u_{\text{ЧМ}}(t) &= \text{Re } W(t) = A[\cos \omega_0 t - m_{\text{ЧМ}} \sin \Omega t \sin \omega_0 t] = \\ &= A \cos \omega_0 t - \frac{Am_{\text{ЧМ}}}{2} \cos(\omega_0 - \Omega)t + \frac{Am_{\text{ЧМ}}}{2} \cos(\omega_0 + \Omega)t. \end{aligned}$$

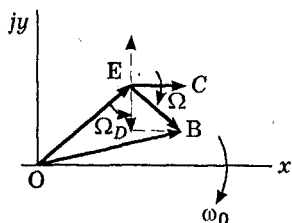


Рис.3.23. Векторная диаграмма ЧМ сигнала при малом индексе модуляции

Сравним ЧМ сигнал с АМ сигналом. Выражение для ЧМ сигнала почти совпадает с выражением для АМ-сигнала. Различие только в знаке перед составляющей с частотой $(\omega_0 - \Omega)$. Поворот фазы на 180° косинусоидальной компоненты DE на рис.3.23 приводит к "качанию" результирующего вектора OB ЧМ сигнала около центрального положения. Такого "качания" у АМ сигнала нет. Кроме того, суммарный вектор BE всегда расположен под прямым углом к вектору OE несущего колебания. Для АМ сигнала вектор BE или совпадает по направлению с вектором OE несущей, или прямо противоположен ему.

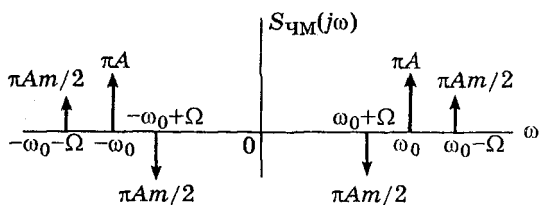


Рис.3.24. Спектральная плотность ЧМ сигнала с малым индексом модуляции

Спектральная диаграмма ЧМ сигнала с малым индексом модуляции приведена на рис.3.24. Спектральные компоненты на частотах $\pm(\omega_0 - \Omega)$ у ЧМ сигнала повернуты на 180° по отношению к подобным компонентам для АМ сигнала (см. рис.3.17, в).

3.5.2. Непрерывный ЧМ сигнал с большим индексом модуляции

В математике доказывается, что функция

$$e^{jm_{\text{ЧМ}} \sin \Omega t}$$

на отрезке $[-\pi, \pi]$ раскладывается в комплексный ряд Фурье:

$$e^{jm_{\text{ЧМ}} \sin \Omega t} = \sum_{k=-\infty}^{\infty} J_k(m_{\text{ЧМ}}) e^{jk\Omega t}, \quad (3.46)$$

где $J_k(m_{\text{ЧМ}})$ — Функции Бесселя k -го индекса от аргумента $m_{\text{ЧМ}}$.
Подставляя формулу (3.46) в (3.45) получим

$$\dot{W}(t) = A \sum_{k=-\infty}^{\infty} J_k(m_{\text{ЧМ}}) e^{j(\omega_0 + k\Omega)t}.$$

Действительная часть полученного выражения

$$u_{\text{ЧМ}}(t) = \text{Re } \dot{W}(t) = A \sum_{k=-\infty}^{\infty} J_k(m_{\text{ЧМ}}) \cos(\omega_0 + k\Omega)t. \quad (3.47)$$

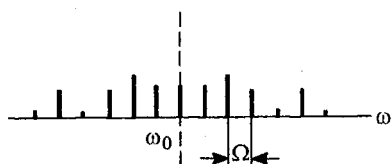


Рис.3.25. Спектр ЧМ сигнала

Из полученного выражения видно, что спектр ЧМ сигнала содержит бесконечное число компонент (рис.3.25). Их амплитуды определяются функциями Бесселя, для которых справедливо равенство:

$$J_{-k}(m_{\text{ЧМ}}) = (-1)^k J_k(m_{\text{ЧМ}}). \quad (3.48)$$

Поэтому начальные фазы боковых колебаний с частотами $\omega_0 + k\Omega$, $\omega_0 - k\Omega$ совпадают, если k — четно и отличаются на 180° , если k — нечетно.

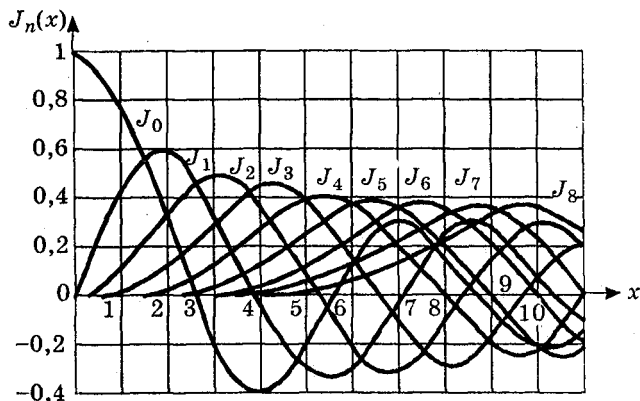


Рис.3.26. Графики функций Бесселя

Графики функций Бесселя приведены на рис.3.26.

Используя упомянутое выше свойство (3.48) функций Бесселя, выражение (3.47) запишем в виде:

$$u_{\text{ЧМ}}(t) = A \{ J_0(m_{\text{ЧМ}}) \cos \omega_0(t) - J_1(m_{\text{ЧМ}}) [\cos(\omega_0 + \Omega)t - \cos(\omega_0 - \Omega)t] + J_2(m_{\text{ЧМ}}) [\cos(\omega_0 + 2\Omega)t + \cos(\omega_0 - 2\Omega)t] + \dots$$

Обращаясь к амплитудному спектру ЧМ сигнала, отметим, что с ростом частоты амплитуды убывают. Установлено, что если число $k \leq m_{\text{ЧМ}}$, то амплитуды гармоник практически равны нулю. Поэтому *ширина спектра ЧМ сигнала*

$$\Pi = 2m_{\text{ЧМ}}\Omega = 2 \cdot \frac{\Delta\omega}{\Omega} \Omega = 2\Delta\omega,$$

где $\Delta\omega$ — девиация частоты.

Для ЧМ сигнала с малым индексом модуляции ширина полосы $\Pi = 2\Omega$, т.е. такая же, как и для АМ сигнала.

3.5.3. Однополосный непрерывный сигнал с угловой модуляцией

Сигналы с угловой модуляцией могут формироваться и с одной боковой полосой.

Покажем, что комплексный сигнал угловой модуляции, записанный в виде аналитической функции

$$\dot{W}(t) = A_0 e^{j\omega_0 t} e^{j[\theta(t) + j\theta^*(t)]},$$

является сигналом с одной боковой полосой.

Отметим, что фаза сигнала, входящая в написанное выражение, также является аналитической функцией.

При тональной частотной модуляции

$$\theta(t) = \Delta\omega \cos \Omega t, \quad \theta^*(t) = \Delta\omega \sin \Omega t.$$

$$\dot{W}(t) = A_0 e^{j\omega_0 t} e^{j\Delta\omega \int (\cos \Omega t + j \sin \Omega t) dt} = A e^{j\omega_0 t} e^{m_{\text{ЧМ}} e^{j\omega t}},$$

где $m_{\text{ЧМ}}$ — индекс ЧМ.

Используя известную формулу

$$e^x = \sum_{k=0}^{\infty} x^k / k!,$$

перепишем выражение для ЧМ сигнала

$$\dot{W}(t) = A e^{j\omega_0 t} \sum_{k=0}^{\infty} \left(\frac{1}{k!} \right) m_{\text{ЧМ}}^k e^{jk\Omega t} = A \sum_{k=0}^{\infty} \left(\frac{1}{k!} \right) m_{\text{ЧМ}}^k e^{(\omega_0 + k\Omega)t}.$$

Действительная часть полученного выражения

$$u_{\text{ОЧМ}}(t) = A \sum_{k=0}^{\infty} \left(\frac{1}{k!} \right) m_{\text{ЧМ}}^k \cos(\omega_0 + k\Omega)t. \quad (3.49)$$

Сравнивая двухполосный ЧМ сигнал с полученным, находим, что последний содержит только одну верхнюю боковую полосу, так как значения " k ", определяющие частоты гармонических составляющих, пробегает значения от 0 до ∞ .

Исключение нижней боковой полосы сопровождается некоторым расширением верхней боковой полосы. При этом ширина спектра однополосной ЧМ составляет примерно 2/3 от ширины спектра двухполосной ЧМ.

3.5.4. Методы формирования непрерывного ЧМ сигнала

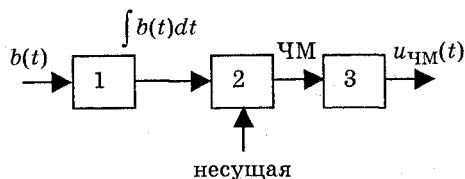


Рис.3.27. Структурная схема модулятора широкополосной ЧМ:
1 — интегратор; 2 — фазовый модулятор; 3 — умножитель частоты

На рис.3.27 приведена структурная схема модулятора широкополосной ЧМ. В этой схеме вначале формируется ФМ сигнал

$$u(t) = A \cos[\omega_0 t + 2\pi K \int b(t) dt],$$

который затем преобразуется в умножителе частоты. Пусть в умножителе производится возведение в квадрат входного напряжения, т.е. оператор нелинейного преобразователя

$$L[u(t)] = c_2 u^2(t).$$

Тогда

$$L[u(t)] = A^2 \cos^2[\omega_0 t + 2\pi K \int b(t) dt] = \frac{A^2}{2} \{1 + \cos[2\omega_0 t + 4\pi K \int b(t) dt]\}.$$

При возведении в квадрат несущая увеличилась в 2 раза. В такое же число раз увеличился и индекс частотной модуляции. Очевидно, что при многократном умножении индекс модуляции может увеличиться в несколько раз.

Другой способ получения ЧМ сигнала связан с прямым воздействием модулирующего сигнала $b(t)$ на реактивные элементы контура генератора. Допустим, что емкость контура изменяется по закону

$$c = c_0 + kb(t) = c_0 \left[1 + \frac{k}{c_0} b(t)\right],$$

где k — постоянный множитель с размерностью $[\Phi/\text{В}]$.

Частота напряжения на выходе контура

$$\omega = \frac{1}{\sqrt{LC_0} \left[1 + \frac{k}{C_0} b(t)\right]^{1/2}}. \quad (3.50)$$

Если выполнить условие $x = \frac{k}{C_0} b(t) \ll 1$, то функцию

$$f(x) = 1 / \sqrt{1+x}$$

можно разложить в ряд Лорана

$$f(x) \cong f(a) + (x-a)f'(a) + \dots + \frac{(x-a)^n}{n!} f^n(a).$$

Вычислив производную $f'(x) = -\frac{1}{2}(1+x)^{-3/2}$ и положив $a = 0$, получим

$$f(x) = 1 - x/2 = 1 - \frac{kb(t)}{2C_0}.$$

Подставляя полученное выражение в (3.50), получим

$$\omega = \frac{1}{\sqrt{LC_0}} \left[1 - \frac{kb(t)}{2C_0} \right] = \omega_0 + k_{\text{ЧМ}} b(t),$$

где

$$K_{\text{ЧМ}} = -k/2C_0.$$

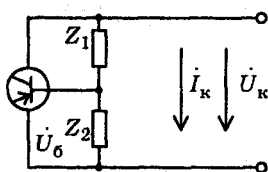


Рис.3.28. Схема реактивного каскада на транзисторе

Переменную емкость получают, например, с помощью закрытого полупроводникового диода (варикапа), емкость которого зависит от подводимого к нему напряжения. Переменную индуктивность получают с помощью насыщенной катушки индуктивности. Переменную реактивность (индуктивность или емкость) можно получить также, применяя реактивный каскад на транзисторе или на лампе.

Схема реактивного каскада на транзисторе приведена на рис.3.28. На базу транзистора подается высокочастотное колебание.

Выбором параметров z_1 , z_2 обеспечивается неравенство токов $\dot{I}_k \gg \dot{I}$. Нагрузка между коллектором и эмиттером

$$Z = \dot{U}_k / \dot{I}_k,$$

где $\dot{I}_k = s\dot{U}_6 = sIZ_2$; S — мгновенная крутизна характеристики транзистора.

Напряжение

$$\dot{U}_k = \dot{I}(Z_1 + Z_2)$$

разделим на $\dot{I}_k$. Получим

$$Z = \frac{\dot{I}}{\dot{I}_k} (Z_1 + Z_2) = (Z_1 + Z_2) / sZ_2.$$

Если выполняется условие $|Z_2| \ll |Z_1|$, то

$$Z \approx Z_1 / sZ_2.$$

Пусть $Z_1 = \frac{1}{j\omega C}$, а $Z_2 = R$.

В этом случае

$$Z = \frac{1}{j\omega C R S} = \frac{1}{j\omega C_3},$$

где

$$C_3 = sRC.$$

Реактивный каскад ведет себя как конденсатор с эквивалентной емкостью $C_3 = sRC$.

Если $Z_1 = R$, а $Z_2 = 1/j\omega C$, то

$$Z = j\omega C R / s = j\omega L_3.$$

Реактивный каскад эквивалентен катушке индуктивности величиной $L_3 = CR/s$.

Переменную C и L можно получить изменением крутизны характеристики транзистора.

3.5.5. Дискретная частотная модуляция

При дискретной ЧМ частота сигнала изменяется дискретно в соответствии с передаваемым кодовым словом. Основание кода может быть любым, но наибольшее распространение получили двоичные коды. Такие сигналы называются *двоичными*. Дискретные двоичные ЧМ сигналы являются сигналами с активной паузой. Их энергия в любом единичном интервале не равна нулю.

Два или несколько гармонических колебаний называются *когерентными (связанными)*, если их несущие частоты и начальные фазы связаны взаимнооднозначной зависимостью. Значение этой зависимости (закона) позволяет по значению частоты и фазы, например, 1-го колебания определить частоты и начальные фазы остальных колебаний.

Например, два колебания $\cos(\omega_0 t + \varphi_0)$ и $\cos(k\omega_0 t + k\varphi_0)$, где k — целое положительное число, являются когерентными.

Некогерентными называются такие гармонические колебания, несущие частоты и начальные фазы которых не связаны между собой.

Модуляторы двоичного дискретного ЧМ сигнала могут строиться с использованием как когерентных так и некогерентных гармонических сигналов. Структурная схема модулятора, построенного с применением некогерентных сигналов, приведена на рис.3.29.

Двоичный дискретный ЧМ сигнал на выходе модулятора

$$u_{\text{вых}}(t) = Aa(t)\cos\omega_1 t + A\bar{a}(t)\cos\omega_2 t, \quad (3.51)$$

где модулирующие сигналы определены следующими выражениями:

$$a(t) = \sum_{k=1}^n a_k v[t - (k-1)\tau_n]; \quad \bar{a}(t) = \sum_{k=1}^n \bar{a}_k v[t - (k-1)\tau_n];$$

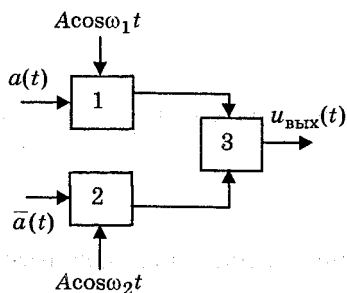


Рис.3.29. Структурная схема модулятора двоичного дискретного ЧМ сигнала:
1, 2 — умножители;
3 — сумматор

a_k — k -й элемент двоичного кодового слова, равный 1 или 0; $\bar{a}_k$ — k -й отрицательный двоичный элемент кодового слова, равный 1 или 0, причем $a_k + \bar{a}_k = 1$; + — знак сложения по модулю два; $v(t)$ — форма единичного элемента двоичного дискретного сигнала.

Как следует из выражения (3.51), двоичный ЧМ сигнал строится из двух двоичных дискретных АМ сигналов. Один из них формируется на выходе первого, другой — на выходе второго умножителя модулятора (рис.3.29). Поэтому спектральная плотность двоичного ЧМ сигнала равна сумме спектральных плотностей двух двоичных АМ сигналов (рис.3.30).

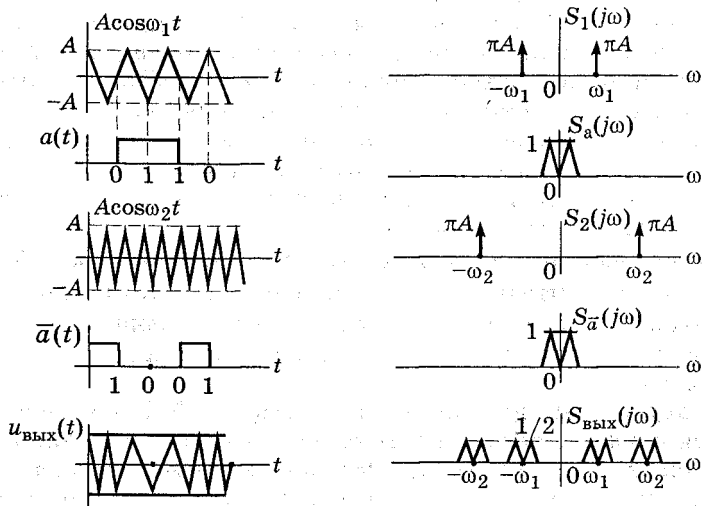


Рис.3.30. Диаграммы, поясняющие формирование двоичного дискретного ЧМ сигнала

Ширина спектра ЧМ сигнала с прямоугольной огибающей радиоимпульсов

$$\Delta\omega = \omega_2 - \omega_1 + 4\pi / \tau_{\text{и}}. \quad (3.52)$$

Если разность $\omega_2 - \omega_1 = 4\pi / \tau_{\text{и}}$, то

$$\Delta\omega = 8\pi / \tau_{\text{и}}. \quad (3.53)$$

Отношение $m = \Delta\omega / \omega_{\text{т}}$, где $\omega_{\text{т}} = 2\pi / \tau_{\text{и}}$, называется *индексом ЧМ*.

3.6. Фазовая модуляция гармонического переносчика

3.6.1. Непрерывная фазовая модуляция

Ранее в п.3.5.1. показано, что ФМ является разновидностью угловой модуляции. Для ФМ сигнала его фаза определяется формулой:

$$\theta(t) = \omega_0 t + 2\pi k_{\text{фм}} b(t) = \omega_0 t + 2\pi k_{\text{фм}} b_0 \cos \Omega t = \omega_0 t + m_{\text{ф}} \cos \Omega t,$$

где $m_{\text{ф}} = 2\pi k_{\text{фм}} b_0$ — индекс фазовой модуляции.

Мгновенная частота ФМ сигнала:

$$\omega(t) = \frac{d\theta(t)}{dt} = \omega_0 - 2\pi k_{\text{фм}} b_0 \Omega \sin \Omega t = \omega_0 - \Delta\omega \sin \Omega t,$$

где $\Delta\omega = 2\pi k_{\text{фм}} b_0 \Omega$ девиация частоты.

ФМ сигнал определяется формулой (3.47) при подстановке вместо $m_{\text{ЧМ}}$ выражения $\Delta\omega / \Omega = m_{\text{фм}} \Omega / \Omega = m_{\text{фм}}$.

Девиация частоты $\Delta\omega$ ФМ сигнала пропорциональна частоте модуляции и ширина спектра ФМ сигнала не будет постоянной, как у ЧМ сигнала, а будет меняться с изменением частоты Ω модулирующего сигнала. По этой причине ФМ сигналы в системах радиовещания не применяются. Однако ФМ сигнал получить легче, чем ЧМ. Поэтому во многих системах ЧМ сигнал получают с помощью фазовых модуляторов, на которые подают сигнал типа $\int s(t)dt$.

3.6.2. Дискретная двоичная фазовая модуляция

При *дискретной фазовой модуляции* фаза сигнала изменяется дискретно. Число возможных значений этой фазы определяет название модуляции. Если это число равно 2, то модуляция называется *двоичной ФМ*, если 3 — *троичной*, и т.д., если 2^k где k — целое положительное число, причем $k \geq 3$, то модуляция называется *многопозиционной*. При $k = 2$ модуляция также называется *двукратной фазовой модуляцией* (ДФМ).

При двоичной ФМ модулирующий сигнал

$$b(t) = \sum_{k=1}^n b_k v[t - (k-1)\tau_{\text{и}}],$$

где b_k — элементы двоичного кодового слова, равные либо 0 либо 1; $v(t)$ — форма импульса модулирующего сигнала.

Фаза сигнала в пределах некоторого k -го единичного интервала $\tau_{\text{и}}$ при $b_k=0$ $\theta(t) = \omega_0 t$, а при $b_k=1$ $\theta(t) = \omega_0 t + 2\pi k_{\text{фм}} v[t - (k-1)\tau_{\text{и}}]$.

Пусть $v(t)$ — прямоугольный импульс длительностью $\tau_{\text{и}}$ и с амплитудой равной 1, тогда

$$\theta(t) = \omega_0 t + 2\pi k_{\text{фм}}.$$

Если коэффициент $k_{\text{фм}} = 0,5$, то $2\pi k_{\text{фм}} = \pi$ и фаза $\theta(t) = \omega_0 t + \pi$.

Двоичный ФМ сигнал в пределах k -го единичного интервала

$$u_k(t) = \begin{cases} Av[t - (k-1)\tau_{\text{и}}] \cos \omega_0 t, & b_k = 0; \\ Av[t - (k-1)\tau_{\text{и}}] \cos(\omega_0 t + \pi) = -Av[t - (k-1)\tau_{\text{и}}] \cos \omega_0 t, & b_k = 1. \end{cases}$$

Этот сигнал можно записать также в виде

$$u_k(t) = Ac_k v[t - (k-1)\tau_{\text{и}}] \cos \omega_0 t,$$

где c_k — k -й элемент, полученный преобразованием k -го элемента кодового слова по правилу: 0 заменяется на 1, а 1 заменяется на -1.

Двоичный ФМ сигнал с прямоугольной огибающей

$$u_{\text{фм}}(t) = A \sum_{k=1}^n c_k v[t - (k-1)\tau_{\text{и}}] \cos \omega_0 t = Ac(t) \cos \omega_0 t, \quad (3.54)$$

где $c(t) = \sum_{k=1}^n c_k v[t - (k-1)\tau_{\text{и}}]$.

Поскольку двоичный ФМ сигнал с прямоугольной огибающей, и двоичный АМ сигнал определяются аналогичными формулами, то свойства и параметры спектральных плотностей этих сигналов совпадают. В частности, ширина спектра двоичного ФМ сигнала с прямоугольной огибающей определяется той же формулой, что и для двоичного АМ сигнала с той же огибающей:

$$\Delta\omega = 2\omega_{\text{в}},$$

где $\omega_{\text{в}} = 2\pi/\tau_{\text{и}}$, $\tau_{\text{и}}$ — длительность прямоугольного видеоимпульса.

Диаграммы, поясняющие формирование двоичной ФМ, приведены на рис.3.31.

Разновидностью двоичной ФМ является относительная фазовая модуляция (ОФМ). Математическое описание ОФМ сигнала во временной и частотной областях не отличаются от описания ФМ сигнала. Например, двоичный ОФМ сигнал с прямоугольной огибающей описывается формулой (3.54), т.е. определяется произведением модулирующей функции $c(t)$ на гармонический сигнал. Однако функция $c(t)$ при ОФМ отличается от аналогичной

функции $b(t)$ при ФМ, и определяется следующим преобразованием:

$$c_k = b_k + c_{k-1}, \quad (3.55)$$

где b_k , c_k — элементы двоичных кодовых слов, равные 0 или 1 в k -м единичном интервале; + — знак суммирования по модулю два; $k = 1, 2, \dots n$.

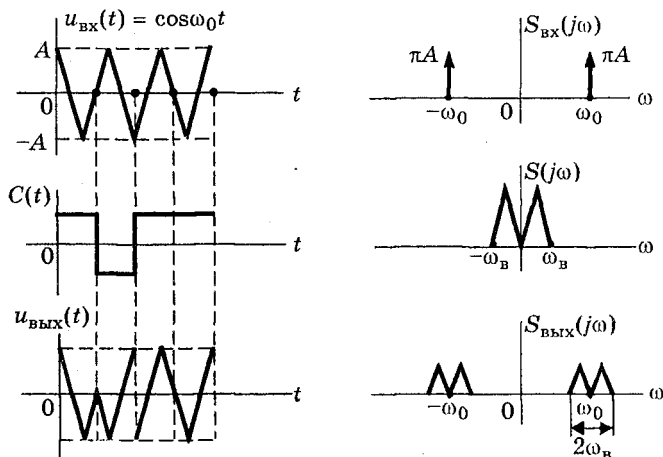


Рис.3.31. Диаграммы, поясняющие формирование двоичной ФМ

Нулю кодового слова $\{b_k\}$ соответствует нулевая разность фаз между радиоимпульсами ОФМ сигнала в k -м и $(k-1)$ -м единичных интервалах, а единице кодового слова — разность фаз, равная π . Преобразование (3.55) определяет так называемый *относительный код*. Например,

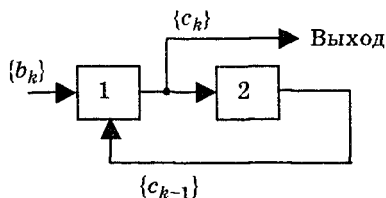


Рис.3.32. Структурная схема преобразователя: 1 — сумматор по модулю 2; 2 — линия задержки

если $c_0 = 0$ и кодовое слово $\{b_k\}$ равно 11010, то кодовое слово $\{c_k\}$ равно 10011. Структурная схема преобразователя приведена на рис.3.32.

3.6.3. Многократная фазовая модуляция

При *многократной ФМ* двоичное кодовое слово, передаваемое по каналу, разбивается на блоки одинаковой длины. Каждо-

му блоку ставится в однозначное соответствие значение фазы гармонического сигнала. Если длина двоичного блока равна " k ", то число значений фаз равно 2^k . Правило отображения блоков двоичного кодового слова в значения фазы гармонического сигнала называется *модуляционным кодом*. Кодовый алфавит состоит из 2^k букв, причем каждой букве соответствует определенное значение фазы.

Частным случаем многократной ФМ является *двукратная ФМ*. Для этой модуляции $k = 2$.

Пусть кодовое слово равно 1001101011. При двукратной ФМ длина блоков равна 2, а блоки, выбранные последовательно во времени равны 10, 01, 10, 10, 11. Если модуляционный код записан в виде табл.3.1, то двоичному кодовому слову соответствует слово модуляционного кода: Диаграммы, поясняющие эту модуляцию, приведены на рис.3.33.

Таблица 3.1

Блок	Фаза
00	$\varphi_1 = 45^\circ$
01	$\varphi_2 = 135^\circ$
10	$\varphi_3 = 225^\circ$
11	$\varphi_4 = 315^\circ$

Форма единичного элемента двукратной ФМ имеет вид прямоугольника с основанием $2\tau_{\text{и}}$, поэтому ширина его спектра равна $\Delta\omega = 4\pi/2\tau_{\text{и}} = 2\pi/\tau_{\text{и}}$, т.е. в два раза меньше, чем при двоичной ФМ.

Многократная ФМ с числом значений фаз $N = 2^k$ и с прямоугольной формой единичного элемента имеет ширину спектра

$$\Delta\omega = 4\pi/\tau_{\text{и}} \log_2 N = 4\pi/k\tau_{\text{и}}.$$

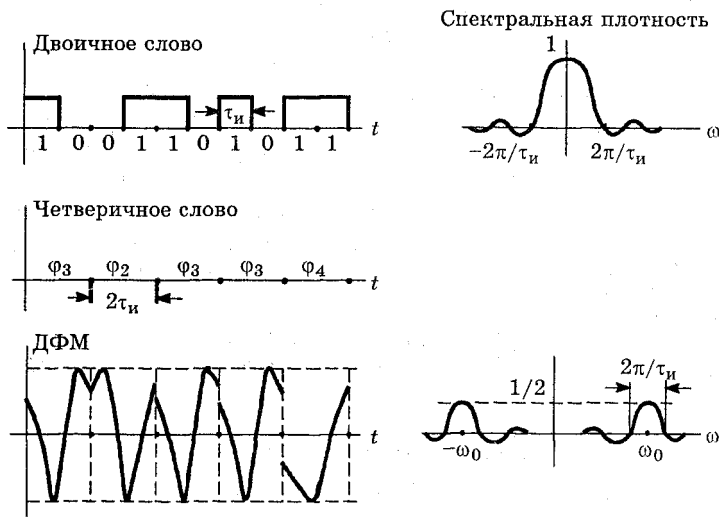


Рис.3.33. Диаграммы, поясняющие двукратную ФМ

Это позволяет при той же полосе, что и для двоичной ФМ увеличить скорость передачи информации в $k = \log_2 N$ раз.

При многократной ОФМ каждому блоку из " k " символов двоичного кодового слова ставится в однозначное соответствие значение разности фаз гармонического сигнала в " k "-м и " $(k - 1)$ "-м единичных интервалах. Правило отображения блоков двоичного кодового слова в значения разности фаз остается тем же, что и при многократной ФМ.

3.6.4. Квадратурная амплитудная модуляция

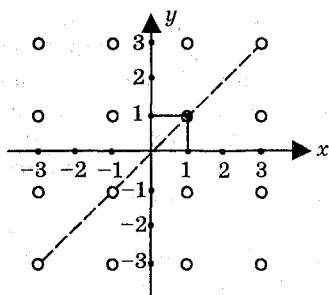


Рис.3.34. Сигналы единичных элементов квадратурной АМ

При *квадратурной амплитудной модуляции* (КАМ) каждому блоку из " k " символов двоичного кодового слова ставятся определенные значения начальной фазы и амплитуды гармонического носителя. Например, всевозможным комбинациям из четырех символов могут быть поставлены 16 гармонических сигналов одинаковой длительности (рис.3.34), отличающихся друг от друга амплитудой и фазой. Для выделенного на этом рисунке двойной линией единичного элемента начальная фаза равна 45° , а амплитуда равна $\sqrt{2}$.

Сигнал с КАМ

$$\begin{aligned} \dot{u}(t) = \sum_{k=1}^n A_k v[t - (k-1)\tau_u] \cos(\omega_0 t + \varphi_k) &= \sum_{k=1}^n A_k \cos \varphi_k v[t - (k-1)\tau_u] \cos \omega_0 t - \\ &- A_k \sin \varphi_k v[t - (k-1)\tau_u] \sin \omega_0 t. \end{aligned} \quad (3.56)$$

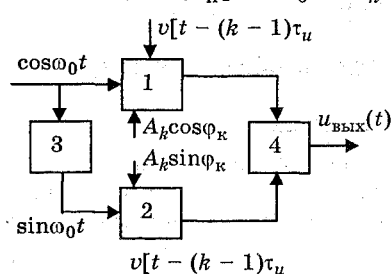


Рис.3.35. Структурная схема модулятора сигнала с КАМ: 1,2 — перемножители; 3 — фазосдвигающее на $\pi/2$ устройство; 4 — вычитающее устройство

где $v(t)$ — форма единичного элемента сигнала; A_k — амплитуда; φ_k — начальная фаза k -го единичного радиоэлемента.

Если $A_k = \text{const}$, то сигнал с КАМ превращается в сигнал с многократной ФМ (ОФМ).

Структурная схема модулятора построенного по выражению (3.56) сигнала с КАМ, показана на рис.3.35.

Использование КАМ позволяет (как и при многократной ФМ) сократить ширину спектра сигнала в $\log_2 N$ раз или при той же поло-

се, что и для двоичной ФМ, увеличить скорость передачи информации в $k = \log_2 N$ раз.

3.6.5. Дискретная фазовая модуляция с одной боковой полосой

Дискретная ФМ с ОБП мало чем отличается от рассмотренной ранее АМ-ОБП. При этом структурная схема модулятора аналогична схеме на рис.3.20. Она содержит модулятор дискретного сигнала и полосовой фильтр, настроенный на нижнюю или верхнюю боковую полосу. Рис.3.36 иллюстрирует работу модулятора на примере двоичной ФМ-ОБП.

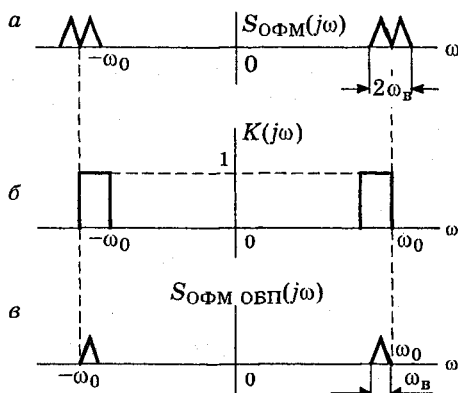


Рис.3.36. Диаграммы, поясняющие формирование двоичной дискретной ФМ-ОБП

При дискретной модуляции с ОБП сохраняются характерные для АМ-ОБП свойства преобразований во временной и частотной областях. Главными из них являются: свойство переноса спектральных компонент из области нулевой частоты в область более высоких частот с сохранением всей информации о модулирующем дискретном сигнале и сокращение в 2 раза ширины полосы.

Дискретная модуляция с ОБП применяется при двоичной ФМ и ОФМ, двукратной и многократной ФМ и КАМ.

3.7. Дискретная модуляция шумоподобного сигнала

3.7.1. Дискретная модуляция постоянного тока

Эти методы используются в системах передачи данных.

Известны две разновидности дискретной модуляции постоянного тока: *модуляция по амплитуде (АМ)* и *модуляция по знаку*.

При АМ изменяется амплитуда, а при модуляции знака — направление тока в соответствии с передаваемым кодовым словом. Во временной области сигналы определяются формулой (3.33), а в частотной — формулой (3.34). Для АМ сигнала $b_k \in \{0,1\}$, $k = 1,2,\dots,n$. Для модулированного по знаку сигнала $b_k \in \{-1,1\}$. Структурные схемы модуляторов приведены на

рис.3.37, 3.38. В структуре амплитудного модулятора содержится ключ, а в структуре знакового модулятора — переключатель. Ключ и переключатель являются управляемыми, а их состояния определяются элементами кодового слова.

3.7.2. Генерирование шумоподобных сигналов

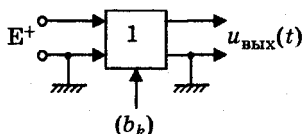


Рис.3.37. Структурная схема модулятора постоянного тока по амплитуде:

1 — управляемый ключ, $\{b_k\}$ — элемент кодового слова

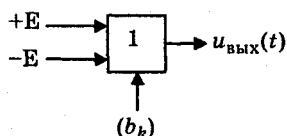


Рис.3.38. Структурная схема модулятора постоянного тока по знаку:

1 — переключатель; $\{b_k\}$ — элемент кодового слова

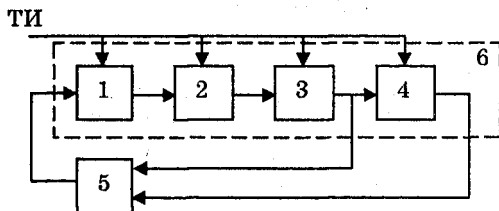


Рис.3.39. Генератор шумоподобного сигнала: 1, 2, 3, 4 — ячейки регистра сдвига; 5 — сумматор по модулю 2; 6 — регистр сдвига

Шумоподобные сигналы можно получить, используя модуляторы постоянного тока и генераторы. В случае применения модуляторов кодовые слова, определяющие закон формирования ШС, хранятся в запоминающем устройстве. Сигналы, формируемые в соответствии с этими словами, поступают на управляющий вход ключа (рис.3.37) или переключателя (рис.3.38).

Генераторы строятся по автономной схеме, структура которой определяется законом формирования ШС.

Наиболее распространенный подкласс генераторов ШС строится на базе *цифровых автоматов*. Если формируемая на его выходе двоичная последовательность получена с использованием операций задержки и логической операции суммирования по модулю два, то такой автомат называется *линейным*, а формируемая на его выходе последовательность — *линейной последовательностью*. Если кроме указанных операций используются дополнительные логические операции (например, операция И), то формируемая цифровым автоматом двоичная последовательность называется *нелинейной*.

Рассмотрим пример генерирования линейной последовательности.

Рис.3.39 приведена схема генератора ШС, кодированного M -последовательностью. Буквой M принято обозначать разновидность линейных двоичных последователь-

ностей. Генератор содержит регистр сдвига с обратными связями и сумматор по модулю два. Подключение входов сумматора к выходам определенных ячеек регистра определяет структуру M -последовательности.

Пусть до подачи тактовых импульсов (ТИ) в регистр сдвига было введено число 1000. С поступлением 1-го тактового импульса в 1-ю ячейку регистра запишется цифра "0", т.к. сумма по модулю два цифр третьего и четвертого разрядов равна 0, а во 2-ю ячейку перейдет цифра "1", ранее записанная в 1-й ячейке. Соответственно в 3-ю и 4-ю ячейки запишутся цифры, ранее содержащиеся во 2-й и 3-й ячейках и т.д. Эти состояния ячеек приведены в 1-й строке табл.3.2.

Таблица 3.2

Номер тактового импульса	Состояния ячеек регистра сдвига			
	1	2	3	4
1	0	1	0	0
2	0	0	1	0
3	1	0	0	1
4	1	1	0	0
5	0	1	1	0
6	1	0	1	1
7	0	1	0	1
8	1	0	1	0
9	1	1	0	1
10	1	1	1	0
11	1	1	1	1
12	0	1	1	1
13	0	0	1	1
14	0	0	0	1
15	1	0	0	0

С поступлением последующих тактовых импульсов процесс смены состояний ячеек продолжится. В результате на их выходах будет генерироваться ШС, кодированные циклически сдвинутыми M -последовательностями. Каждый из этих ШС состоит из единичных элементов — импульсов прямоугольной формы. Их амплитуда равна или не равна нулю в соответствии с символом $\{0,1\}$ M -последовательности.

Длительность единичных элементов $\tau_{и}$ ШС определяется периодом тактовых импульсов. Например, на выходе четвертой ячейки регистра генерируемый ШС определяется последовательностью 001001101011110. Она записана в последнем столбце табл.3.2. При непрерывном поступлении тактовых импульсов ШС периодически повторяется через время, равное $N\tau_{и} = 15\tau_{и}$, где

$N = 2^k - 1$ — период M -последовательности, определяемой числом разрядов k регистра сдвига (в нашем примере $k = 4$); $\tau_{\text{и}}$ — длительность периода тактовых импульсов. Начальную комбинацию можно задавать любым из 15 двоичных чисел, кроме числа 0000. Генерируемые ШС остаются при этом одними и теми же, но только циклически сдвигаются в зависимости от начальной кодовой комбинации.

Изменяя подключение входов сумматора по модулю два к тем или другим ячейкам регистра сдвига, можно получать разные по структуре ШС, но с одними и теми же свойствами. Число таких ШС равно $\frac{1}{k} \phi(2^k - 1)$, где $\phi(x)$ — функция Эйлера.

Например, для $k = 4$ число M -последовательностей равно 2, а для $k = 11$ — 176, т.е. с ростом длины их число резко увеличивается.

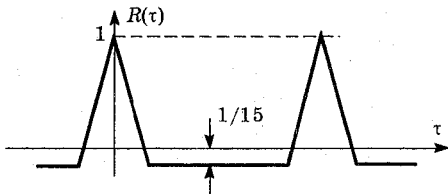


Рис.3.40. Периодическая автокорреляционная функция ШС

Периодическая нормированная автокорреляционная функция ШС, кодированного M -последовательностью длиной $N = 15$ символов, приведена на рис.3.40. Этот ШС строится из разнополярных импульсов одинаковой амплитуды и длительности по закону, заданному M -последовательностью.

Нормированная непериодическая функция автокорреляции того же ШС имеет значения не превышающие $1/\sqrt{N}$ (при сдвигах не равных нулю).

Свойства автокорреляционной функции ШС, кодированного M -последовательностью, близки к свойствам автокорреляционной функции шума, возникающего во входных цепях усилителей. По этой причине такие сигналы называются *шумоподобными*.

3.7.3. Модуляция шумоподобного сигнала по форме

При этой модуляции сообщение в виде двоичного кодового слова разбивается на блоки длиной в " k " символов. Набору 2^k $k = 1, 2, 3, \dots$ двоичных кодовых слов каждого блока ставится в однозначное соответствие набор отличающихся по форме ШС.

Частными случаями этой модуляции являются: амплитудная модуляция ШС и фазовая модуляция ШС. Для АМ и ФМ ШС длина блоков $k = 1$, а число возможных форм равно 2.

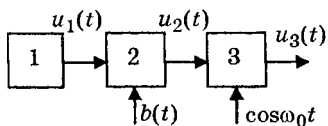


Рис.3.41. Амплитудный модулятор шумоподобного сигнала: 1 — генератор ШС; 2 — управляемый ключ (первый АМ); 3 — второй АМ

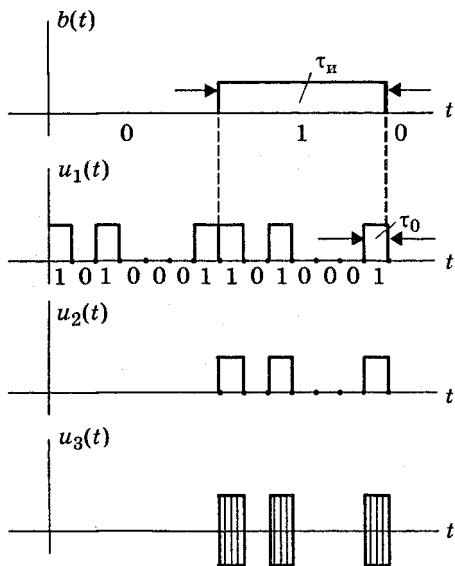


Рис.3.42. Временные диаграммы, поясняющие работу амплитудного модулятора ШС

На рис.3.41 приведена схема амплитудного модулятора ШС. Генератор 1 выдает шумоподобную последовательность $u_1(t)$, определяемую периодическим двоичным кодовым словом. Для иллюстрации работы схемы (см. рис.3.42) это слово выбрано равным коду Баркера длиной в 7 символов: 1010001. Длительность импульса шумоподобной последовательности $u_1(t)$ равна τ_0 . Она меньше длительности

импульса информационной последовательности $b(t)$ в τ_n/τ_0 раз. Положение управляемого ключа 2 определяется элементом информационного кодового слова: если этот элемент равен 1, то ключ замкнут, а если равен 0, то — разомкнут. Ширина спектра ШС $u_2(t)$ определяется шириной τ_0 импульса. Если форма этого импульса имеет прямоугольный вид, то ширина спектра ШС равна $1/\tau_0$. Она в τ_n/τ_0 раз больше ширины спектра информационной двоичной кодовой последовательности.

Произведение длительности ШС на ширину спектра ШС называется *базой сигнала*. База ШС

$$B = \tau_n \frac{1}{\tau_0} = N, \quad (3.57)$$

где N — длина кодового слова шумоподобной последовательности. В рассматриваемом нами примере $N = 7$.

В некоторых случаях ШС $u_2(t)$ непосредственно поступает в канал, а в других — преобразуется в АМ радиосигнал $u_3(t)$ при помощи 2-го АМ.

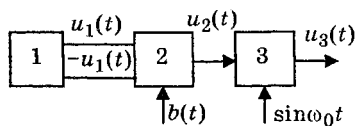


Рис.3.43. Фазовый модулятор шумоподобного сигнала:
 1 — генератор ШС;
 2 — управляемый переключатель (первый ФМ);
 3 — второй ФМ

На рис.3.43, 3.44 приведены структурная схема фазового модулятора ШС и временные диаграммы, поясняющие его работу. В зависимости от символа передаваемого кодового слова 1 или 0 меняется знак ШС на выходе переключателя 2.

Так же, как и при АМ, сигнал или поступает в канал непосредственно или преобразуется с помощью модулятора 3 в ФМ радиосигнал. База ФМ ШС равна базе АМ ШС и определяется формулой (3.57).

На рис.3.45 приведена структурная схема модулятора ШС по форме для $k \geq 1$.

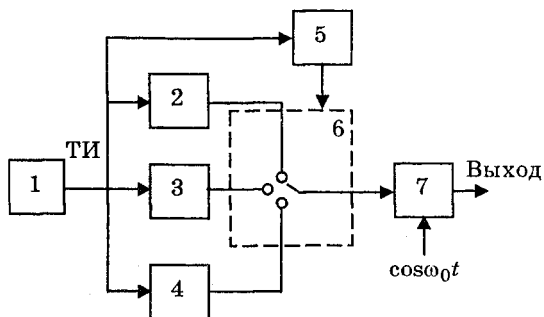


Рис.3.45. Модулятор шумоподобного сигнала по форме: 1 — тактовый генератор; 2,3,4 — 1-й, 2-й и 2^k -й генераторы шумоподобных сигналов; 5 — источник дискретных сообщений; 6 — управляемый переключатель; 7 — фазовый модулятор

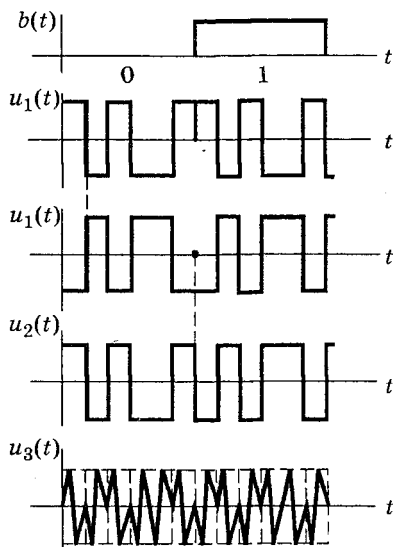


Рис.3.44. Диаграммы, поясняющие работу фазового модулятора ФМ

Поясним его работу на примере использования симплексного кода длиной в 7 символов. Код из N слов называется симплексным, если скалярное произведение любой пары слов этого кода равно $-1/(N-1)$, N — четное и равно $-1/N$, N — нечетное.

В табл.3.3 приведены кодовые слова симплексного кода над алфавитом 1, -1 для $N = 7$. Для этого кода скалярное произведе-

ние равно $-1/7$. В таблице приведены также блоки двоичных чисел, которые выделены из информационного слова дискретного источника. Число разрядов двоичных чисел равно $k = 2$, что соответствует использованию четырех форм ШС.

Таблица 3.3

Двоичный код		Кодовые слова, определяющие ШС							
0	0	-1	1	1	-1	1	-1	-1	
0	1	1	1	-1	1	-1	-1	-1	
1	0	1	-1	1	-1	-1	-1	1	
1	1	-1	1	-1	-1	-1	1	1	

Положение управляемого переключателя на рис.3.45 определяется двоичным числом блока с выхода источника дискретного сообщения. Например, если число блока равно 00, то выход управляемого переключателя соединен с выходом генератора Γ_1 , генерирующего ШС с кодовым словом первой строки табл.3.3.

Импульсы с выхода тактового генератора определяют длительность τ_0 единичного интервала ШС, а также и длительность единичного интервала $\tau_{и}$ двоичных кодовых комбинаций на выходе источника дискретных сообщений.

В модуляторе производится модуляция косинусоидальной несущей шумоподобными сигналами. В рассматриваемой схеме применена фазовая модуляция. Скорость передачи информации ШС с модуляцией по форме

$$R = 1/\tau_{и} \log_2 2^k = k/\tau_{и},$$

если любому ШС соответствует равновероятное двоичное число блока источника сообщения. С ростом числа ШС скорость передачи увеличивается.

амплитуда

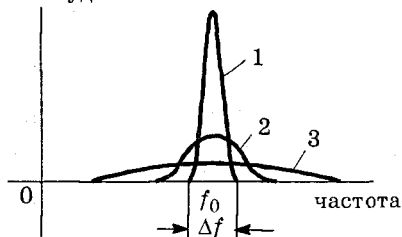


Рис.3.46. Спектры сигналов

Ширина спектра ШС на выходе управляемого переключателя определяется шириной спектра элементарного импульса. Для прямоугольного видеопульса эта ширина равна $1/\tau_0 = N/\tau_{и}$, где N — база ШС. С ростом базы ширина его спектра увеличивается. На рис.3.46 приведены спектральные плотности узкополосного сигнала (первая зависимость) и ШС (вторая и третья зависимости).

Сигналы имеют одинаковые энергии, а их базы определяют неравенством $B_1 < B_2 < B_3$. При приеме узкополосного сигнала 1 в полосе Δf шумоподобные сигналы 2, 3 создают помеху тем большей мощности, чем меньше их база. С увеличением этой базы интенсивность спектральных компонент ШС 3 в полосе Δf

уменьшается, так как его энергия остается постоянной, а спектральные компоненты ШС распределяются в большем диапазоне. При дальнейшем увеличении базы ШС его спектральные компоненты могут уменьшиться настолько, что их интенсивность будет намного меньше флуктуационного шума входных цепей приемника, осуществляющего прием сигнала 1.

Помеха, создаваемая ШС, оказывается пренебрежимо малой. Этот пример показывает возможность совмещения в одном и том же частотном диапазоне как обычных систем, так и систем с ШС. Кроме того, в системах с ШС обеспечивается скрытность передачи, если код, определяющий форму ШС, известен только своему корреспонденту, а база ШС выбрана такой величины, при которой уровень полезного сигнала меньше уровня флуктуационного шума, возникающего во входных цепях приемника.

3.8. Модуляция импульсного переносчика

3.8.1. Амплитудно-импульсная модуляция

В качестве переносчика сообщения в технике связи используется последовательность видеоимпульсов. Если в соответствии с передаваемым сообщением изменяется амплитуда импульсов, то формируемый сигнал называется *амплитудно-импульсно-модулированным (АИМ) сигналом*.

Схема модулятора и диаграммы, поясняющие его работу, приведены на рис.3.47, 3.48.

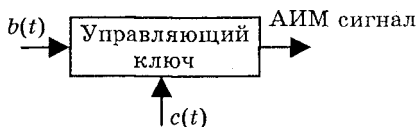


Рис.3.47. Структурная схема модулятора АИМ сигнала

На вход управляемого ключа поступает модулирующий непрерывный сигнал $b(t)$, соответствующий передаваемому сообщению. На другой вход этого ключа подаются носители сообщения в виде видеоимпульсов $C(t)$ с постоянной амплитудой C_0 и постоянной дли-

тельностью τ_n . За время действия этих импульсов выход ключа соединяется с его входом и модулирующий сигнал поступает на выход. Период T повторения импульсов обычно выбирается равным $1/2F$, где F — верхняя граничная частота спектра первичного сигнала.

Так как АИМ сигнал равен произведению

$$u_{\text{АИМ}}(t) = C(t)b(t),$$

то его спектральная плотность определяется сверткой спектральных плотностей модулирующего сигнала и видеоимпульсов.

$$S_{\text{АИМ}}(j\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(jz)S_c[j(\omega - z)dz, \quad (3.58)$$

Спектральная плотность носителя сообщения

$$S_c[j(\omega)] = \pi \sum_{k=-\infty}^{\infty} \frac{2A\tau_{\text{и}}}{T} \frac{\sin \frac{k\omega_1\tau_{\text{и}}}{2}}{k\omega_1\tau_{\text{и}}/2} \delta(k\omega_1 - \omega), \quad (3.59)$$

где

$$\omega_1 = 2\pi/T.$$

Спектральная плотность периодических видеоимпульсов вычислена в примере 2.4 и показана на рис.3.48, *д*.

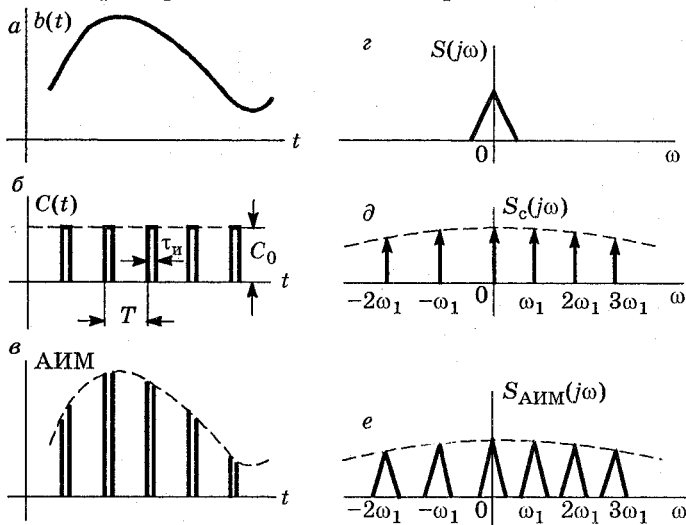


Рис.3.48. Диаграммы, поясняющие работу модулятора АИМ сигнала

Подставляя выражение (3.59) в (3.58), получим (рис.3.48, *е*):

$$S_{\text{АИМ}}(j\omega) = \frac{A\tau_{\text{и}}}{T} \sum_{k=-\infty}^{\infty} \frac{\sin \frac{k\omega_1\tau_{\text{и}}}{2}}{k\omega_1\tau_{\text{и}}/2} S[j(\omega - k\omega_1)]. \quad (3.60)$$

Анализируя полученный результат, находим, что спектральные компоненты группируются около значений частот $\omega = 0, \pm\omega_1, \pm2\omega_1, \pm3\omega_1$, и т.д. и повторяют с точностью до постоянной величины спектр $S(j\omega)$ модулирующего сигнала (рис.3.48 *г, е*). При увеличении частоты интенсивность спектральных компонент убывает по закону $\sin x/x$, причем скорость убывания определяется длительностью видеоимпульса $\tau_{\text{и}}$. Чем меньше эта длительность, тем медленнее убывают спектральные компоненты и наоборот. Длительность прямоугольного видеоимпульса $\tau_{\text{и}}$ определяет ширину спектра АИМ сигнала, которая равна $\pi = 2\pi/\tau_{\text{и}}$.

3.8.2. Широтно-импульсная модуляция. Фазово-импульсная модуляция

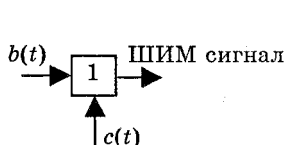


Рис.3.49. Структурная схема модулятора ШИМ сигнала: 1 — генератор импульсов переменной длительности

Сигнал, у которого изменяются длительности импульсов импульсного переносчика, называется *широтно-импульсно-модулированным (ШИМ) сигналом*. Если в соответствии с передаваемым сообщением изменяется задержка импульсов, то такой сигнал называется *фазо-импульсно-модулированным (ФИМ) сигналом*.

Структурная схема модулятора ШИМ сигнала приведена на рис.3.49, а временные диаграммы, поясняющие его работу — на рис.3.50.

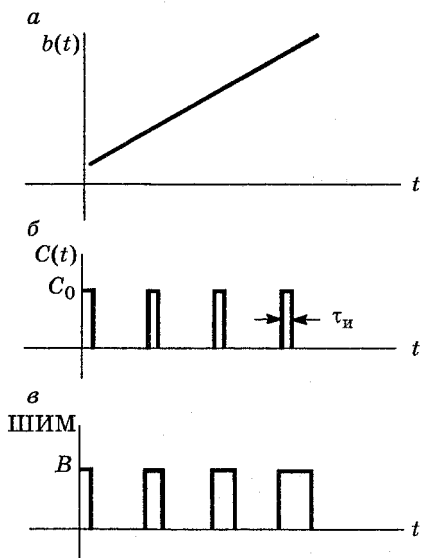


Рис.3.50. Диаграммы, поясняющие работу модулятора ШИМ сигнала

Генератор импульсов переменной длительности является импульсным устройством. Запуск генератора осуществляется видеоимпульсами $C(t)$ с амплитудой C_0 , длительностью $\tau_{и}$ и периодом $T = 1/2F_v$. С приходом каждого импульса генератор вырабатывает импульс с постоянной амплитудой B и длительностью, прямо пропорциональной значению модулирующего сигнала $b(t)$.

Формирование ФИМ сигнала поясняется рис.3.51. и рис.3.52. Схема модулятора содержит: генератор импульсов переменной длительности и формирователь выходных импульсов. В качестве генератора часто используют формирователь ШИМ сигнала.

Рис.3.52, а, б, в иллюстрируют работу генератора импульсов переменной длительности, рис.3.52, г, д поясняют работу формирователя выходных импульсов.

Основными операциями формирователя выходных импульсов являются: дифференцирование ШИМ сигнала; ограничение положительных импульсов на выходе дифференцирующей цепи;

преобразование отрицательных импульсов на выходе дифференцирующей цепи в импульсы заданной формы и длительности.

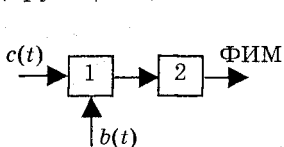


Рис.3.51. Структурная схема модулятора ФММ сигнала: 1 — генератор ШИМ сигнала; 2 — формирователь импульсов

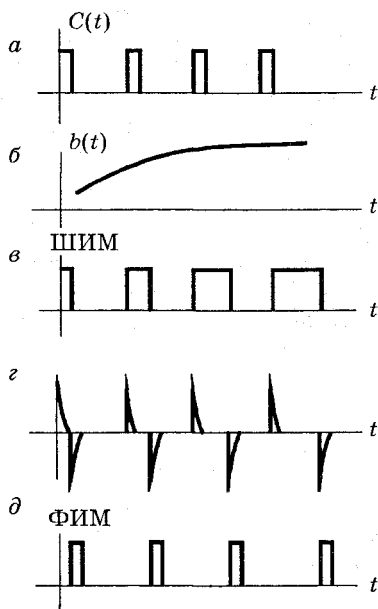


Рис.3.52. Диаграммы, поясняющие работу модулятора ФИМ сигнала

Отыскание спектральных плотностей ШИМ и ФИМ сигналов при произвольном модулирующем сигнале (как это было сделано для АИМ сигнала) затруднено. Поэтому их вычисляют для гармонического сигнала.

Методика расчета следующая. В формулу для носителя сообщения (2.12) подставляют следующие выражения для информационных параметров:

$$\tau_{и} = \tau_{и0} + k_{ш}b_0 \cos \Omega t \quad \text{для ШИМ}$$

$$t_0 = k_{ф}b_0 \cos \Omega t \quad \text{для ФИМ.}$$

Коэффициенты $k_{ш}$ и $k_{ф}$, входящие в эти выражения имеют размерность равную [с/в], и определяют крутизну модуляционных характеристик ШИМ и ФИМ сигналов.

Например, выражение для ШИМ сигнала

$$u_{ШИМ}(t) = \frac{A}{T} (\tau_{и0} + k_{ш}b_0 \cos \Omega t) + \frac{4}{T} \frac{A}{k\omega_1} \sum_{k=1}^{\infty} \sin \frac{k\omega_1}{2} (\tau_{и0} + k_{ш}b_0 \cos \Omega t) \cos k\omega_1 t.$$

Выражение для ФИМ сигнала

$$u_{ФИМ}(t) = A \frac{\tau_{и}}{T} + \frac{2A\tau_{и}}{T} \times \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_1\tau_{и}}{2}}{k\omega_1\tau_{и}/2} \cos k\omega_1(t - k_{ф}b_0 \cos \Omega t).$$

По выражениям ШИМ и ФИМ сигналов нетрудно вычислить спектральные плотности сигналов,

применяя прямое преобразование Фурье.

3.8.3. Импульсно-кодовая модуляция

Преобразование непрерывного первичного аналогового сигнала в код называется *импульсно-кодовой модуляцией (ИКМ)*.

Основание кода может быть любым. В системах электросвязи наибольшее распространение получили двоичные коды, реализуемые с меньшими аппаратными затратами.

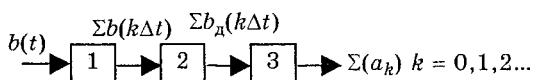


Рис.3.53. Аналого-цифровой преобразователь: 1 — дискретизатор; 2 — квантователь; 3 — кодер

Устройство, преобразующее непрерывный первичный аналоговый сигнал в двоичный код, называется *аналого-цифровым преобразователем* (АЦП).

Структурная схема АЦП и диаграммы приведены на рис.3.53, 3.54.

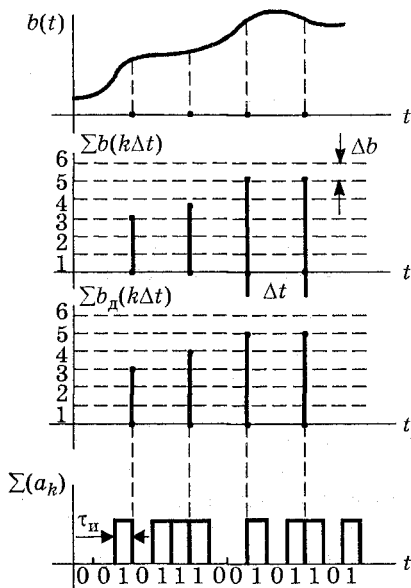


Рис.3.54. Диаграммы, поясняющие работу АЦП

Основными операциями при ИКМ являются операции дискретизации, квантования и кодирования. При дискретизации получают выборочные значения из непрерывного аналогового сигнала. Эти значения являются скалярным произведением функции $b(t)$, определяющей этот сигнал, и обобщенной функции, равной сумме бесконечного числа дельта функций:

$$\sum_{k=-\infty}^{\infty} b(k\Delta t) = \int_{-\infty}^{\infty} b(t) \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) dt,$$

$$k = 0; \pm 1; \pm 2, \dots$$

В этом нетрудно убедиться, если принять во внимание фильтрующее свойство δ -функции.

Алгоритм определения выборочных значений прост и сводится к умножению входного непрерывного сигнала на весьма короткие импульсы, возникаю-

щие в моменты времени $k\Delta t$.

Интервал Δt называется *интервалом дискретизации*. Его величина выбирается в соответствии с теоремой Котельникова. Согласно этой теореме, дискретизация не приводит к искажениям, если $\Delta t \leq 1/2F$, где F — верхняя граничная частота спектра сигнала $b(t)$. Алгоритм дискретизации аналогичен АИМ.

Полученные в результате дискретизации выборочные значения далее переводятся в двоичные числа. Вначале каждое значение $b(k\Delta t)$ сравнивается с заранее заданными постоянными уровнями, которые называются *уровнями квантования*. Затем ближайший к выборочному значению уровень $b_d(k\Delta t)$ переводится в двоичное число a_k . Процесс замены выборочного значения ближайшим к нему уровнем называется *квантованием*. При квантовании возникает погрешность

$$\Delta k = b(k\Delta t) - b_d(k\Delta t). \quad (3.61)$$

Ее величина тем меньше, чем больше уровней квантования L . Таким образом, каждое выборочное значение преобразуется в одно из L возможных значений, которое далее преобразуется в двоичное кодовое слово. Его длина $n = \log_2 L$. Длина двоичной кодовой комбинации, определяемая длиной кодового слова, не должна превышать интервал дискретизации, т.е.

$$n\tau_{\text{и}} < \Delta t.$$

Следовательно, длительность единичного интервала, равная длительности одного импульса кодовой последовательности

$$\tau_{\text{и}} \leq \frac{\Delta t}{n} = \frac{1}{2Fn} = \frac{1}{2F \log_2 L}. \quad (3.62)$$

Если кодовая последовательность содержит импульсы прямоугольной формы, то ширина спектра сигнала при ИКМ равна $1/\tau_{\text{и}} = 2F \log_2 L$.

3.9. Некогерентное детектирование

3.9.1. Некогерентное детектирование непрерывных АМ сигналов

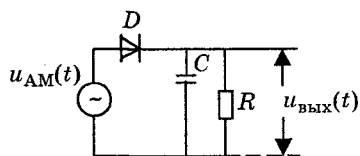


Рис.3.55. Схема некогерентного детектора

Некогерентное детектирование оказывается возможным только при приеме АМ сигнала с большим уровнем несущей и, если индекс модуляции находится в пределах $1 \geq m \geq 0$. При других видах непрерывной АМ: АМ-ПН, АМ-ОБП и АМ-БУН, когда $m > 1$ возникают искажения. Это объясняется тем, что некогерентный детектор (рис.3.55), содержащий ди-

од D и цепочку RC , выделяет огибающую, которая для АМ-ПН, АМ-ОБП и АМ-БУН при $m > 1$ не совпадает с формой модулирующего сигнала (сравните огибающие на рис.3.16, в, г).

Рассмотрим работу некогерентного детектора при АМ с $m \leq 1$. Когда входное напряжение положительно, через диод про-

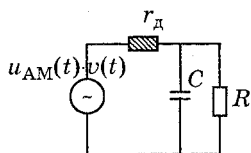


Рис.3.56.
Эквивалентная
схема
некогерентного
детектора

текает ток, а когда отрицательно, ток равен нулю. Если считать, что ограничение тока происходит на нулевом уровне, то схему на рис.3.55 можно заменить эквивалентной (рис.3.56). В этой схеме входное напряжение равно произведению двух напряжений: входного АМ сигнала и напряжения в виде последовательности прямоугольных импульсов с единичной амплитудой, длительностью импульсов $\tau_{\text{и}} = T/2$ и с периодом T повторения, равным периоду высокочастотного заполнения АМ сигнала т.е. $T = 2\pi/\omega_0$ (рис.3.57).

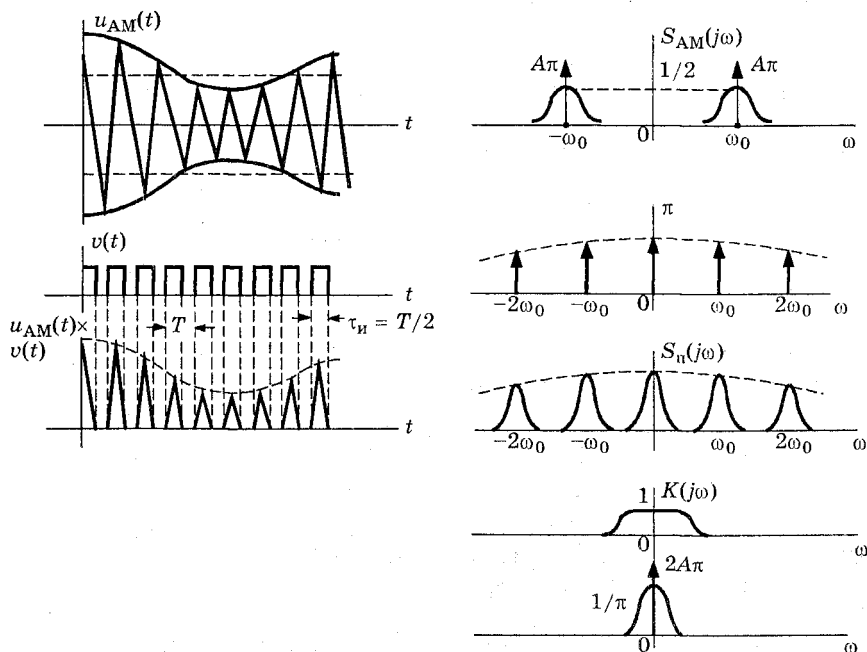


Рис.3.57. Диаграммы, поясняющие работу некогерентного детектора

Сигнал на выходе эквивалентной схемы найдем, как обратное преобразование Фурье от спектральной плотности выходного напряжения

$$u_{\text{ВЫХ}}(t) = \frac{1}{2} \int_{-\infty}^{\infty} S_{\text{ВЫХ}}(j\omega) e^{j\omega t} d\omega,$$

где спектральная плотность выходного сигнала равна спектральной плотности входного сигнала $S_{\Pi}(j\omega)$ (рис.3.57) умноженной на коэффициент передачи $K(j\omega)$ цепочки RC — фильтра нижних частот,

$$S_{\text{вых}}(j\omega) = S_{\Pi}(j\omega)K(j\omega).$$

Спектральная плотность $S_{\Pi}(j\omega)$ равна свертке спектральных плотностей АМ сигнала и последовательности прямоугольных импульсов $v(t)$.

$$S_{\Pi}(j\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{AM}}(jz) S_v[j(\omega - z)] dz, \quad (3.63)$$

где $S_{\text{AM}}(j\omega)$ — спектральная плотность АМ сигнала; $S_v(j\omega)$ спектральная плотность последовательности прямоугольных импульсов.

$$\begin{aligned} S_v(j\omega) &= \pi\delta(0) + \pi \sum_{k=1}^{\infty} \frac{\sin k\pi/2}{k\pi/2} [\delta(k\omega_0 + \omega) + \delta(k\omega_0 - \omega)] = \\ &= \pi \sum_{k=-\infty}^{\infty} \frac{\sin k\pi/2}{k\pi/2} \delta(k\omega_0 - \omega), \end{aligned} \quad (3.64)$$

где $\omega_0 = 2\pi/T$ — несущая частота АМ сигнала.

Выражение (3.64) получено прямым преобразованием Фурье функции $v(t)$ (рис.3.57). Аналогичная задача решена в примере 2.4.

Подставляя формулу (3.64) в (3.63) и учитывая фильтрующее свойство δ -функции, получим

$$S_{\Pi}(j\omega) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \frac{\sin k\pi/2}{k\pi/2} S_{\text{AM}}[j(\omega - k\omega_0)].$$

Полагая, что фильтр нижних частот пропускает только те спектральные составляющие $S_{\Pi}(j\omega)$, которые группируются около нулевой частоты, получим выражение для спектральной плотности сигнала на выходе детектора

$$\begin{aligned} S_{\text{вых}}(j\omega) &= \frac{1}{2} \left\{ \frac{2}{\pi} S_{\text{AM}}[j(\omega - \omega_0)] + \frac{2}{\pi} S_{\text{AM}}[j(\omega + \omega_0)] \right\} = \\ &= \frac{1}{\pi} \{ S_{\text{AM}}[j(\omega - \omega_0)] + S_{\text{AM}}[j(\omega + \omega_0)] \} = \\ &= \frac{1}{\pi} \left[\frac{1}{2} S(j\omega) + \frac{1}{2} S(j\omega) + 2A\pi\delta(0) \right] = \frac{1}{\pi} S(j\omega) + 2A\delta(0). \end{aligned}$$

Сигнал на выходе некогерентного детектора равен

$$u_{\text{ВЫХ}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{1}{\pi} S(j\omega) + 2A\delta(0) \right] e^{j\omega t} d\omega = \frac{1}{\pi} b(t) + \frac{1}{2\pi} 2A[e^0] = \frac{1}{\pi} [A + b(t)].$$

Выражение для сигнала на выходе некогерентного детектора содержит постоянное число A/π , которое определяет постоянное напряжение и функцию $b(t)/\pi$ описывающую переменное напряжение, совпадающее по форме с модулирующим сигналом $b(t)$.

3.9.2. Некогерентное детектирование дискретных сигналов

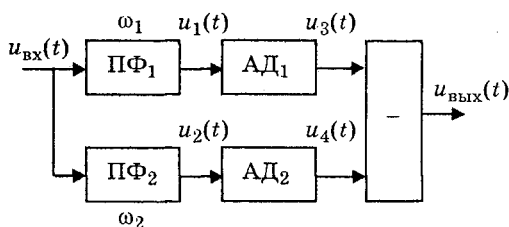


Рис.3.58. Некогерентный детектор двоичного ЧМ сигнала

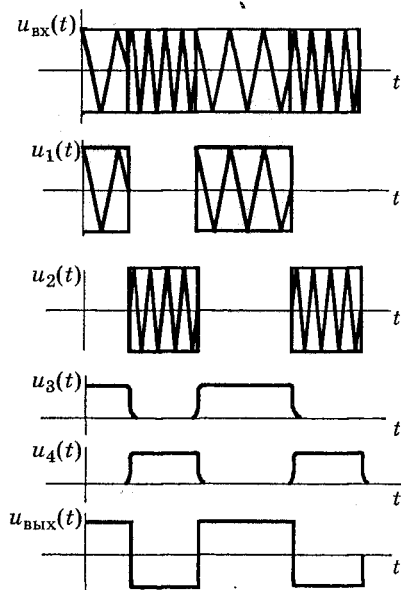


Рис.3.59. Диаграммы, поясняющие работу двоичного ЧМ детектора

Некогерентное детектирование применяется при приеме дискретных АМ и ЧМ сигналов. Оно мало отличается от некогерентного детектирования непрерывных АМ сигналов. Некоторые особенности связаны с импульсным характером сигналов. Это накладывает определенные требования к параметрам фильтра нижних частот и полупроводникового диода.

При некогерентном детектировании двоичных ЧМ сигналов используются два канала обработки. Каждый канал содержит полосовой фильтр и амплитудный некогерентный детектор.

Полосовые фильтры, настроенные на несущие частоты носителей негативных двоичных кодовых последовательностей, разделяют четные и нечетные элементы ЧМ сигнала, которые в дальнейшем детектируются амплитудными некогерентными детекторами. При вычитании выделенных на их входах негативных двоичных кодовых последовательностей образуется выходной сигнал — принимаемая кодовая последовательность (рис.3.58, 3.59).

3.9.3. Детектирование непрерывных ЧМ сигналов

Одним из самых распространенных методов детектирования является метод, основанный на преобразовании ЧМ сигнала в АМ сигнал. Диаграммы, поясняющие этот метод, приведены на рис.3.60, а вариант детектора — на рис.3.61. Частотный детектор содержит два резонансных контура, два диода и два фильтра нижних частот, выполненных на RC -цепочках. Резонансные контуры несколько расстроены относительно средней частоты ω_0 ЧМ сигнала.

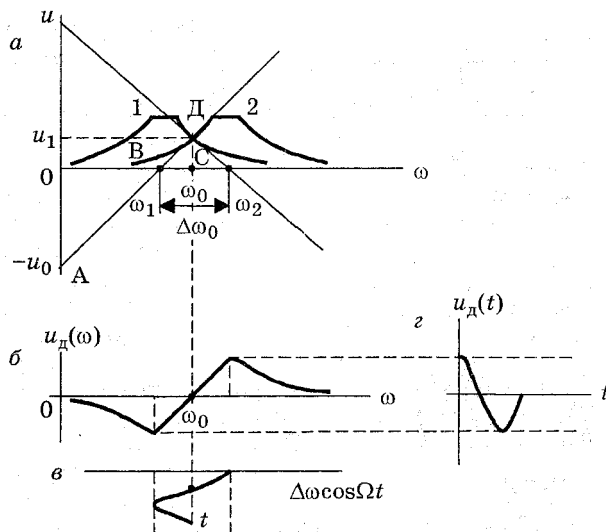


Рис.3.60. Диаграммы, поясняющие работу детектора непрерывного ЧМ сигнала

Аппроксимируем участки кривых амплитудно-частотных характеристик (АЧХ) контуров линейными функциями (рис.3.60, а):

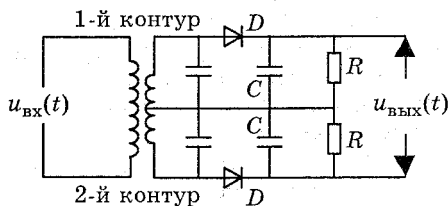


Рис.3.61. Детектор непрерывного ЧМ сигнала

$$u = -\frac{2u_1}{\Delta\omega_0} \omega + \frac{2u_1}{\Delta\omega_0} \left(\omega_0 + \frac{\Delta\omega_0}{2} \right), \quad (3.65)$$

$$u = -\frac{2u_1}{\Delta\omega_0} \omega - \frac{2u_1}{\Delta\omega_0} \left(\omega_0 - \frac{\Delta\omega_0}{2} \right), \quad (3.66)$$

где u_1 — значение напряжения на выходе контуров при $\omega = \omega_0$; $\Delta\omega_0 = \omega_2 - \omega_1$ — разность значений частот, кото-

рые совпадают с точками пересечений прямых линий с осью ω .

Каждое из приведенных уравнений найдено по исходному уравнению

$$u = a\omega + b,$$

где $a = u_0/\omega_1$, $b = -u_0$.

Из подобных треугольников OAB и BDC находим

$$\frac{u_0}{\omega_1} = \frac{2u_1}{\Delta\omega_0}, \quad u_0 = \frac{2u_1}{\Delta\omega_0} \omega_1 = \frac{2u_1}{\Delta\omega_0} (\omega_0 - \Delta\omega/2).$$

Разность левых и правых частей уравнений (3.65), (3.66) определяет линейную часть характеристики детектора

$$u_d(\omega) = \frac{4u_1}{\Delta\omega_0} \omega - \frac{4u_1\omega_0}{\Delta\omega_0} = \frac{4u_1}{\Delta\omega_0} (\omega - \omega_0) = K_d(\omega - \omega_0), \quad (3.67)$$

где $K_d = 4u_1/\Delta\omega_0$ — крутизна характеристики детектора.

Характеристика детектора $u_d(\omega)$ показана на рис.3.60, б. Она получена вычитанием значений кривой 2 из соответствующих значений кривой 1 на рис.3.60, а.

Найдем отклик частотного детектора на входной сигнал, считая, то девиация $\Delta\omega < \Delta\omega_0$.

Подставляя выражение для частоты сигнала

$$\omega = \omega_0 + \Delta\omega \cos \Omega t \quad (3.68)$$

в уравнение (3.67), получим

$$u_d(\omega) = \frac{4u_1}{\Delta\omega_0} (\omega_0 + \Delta\omega \cos \Omega t - \omega_0) = K_d \Delta\omega \cos \Omega t.$$

Отклик частотного детектора на входной ЧМ сигнал, прямо пропорционален модулирующему сигналу (рис.3.60, в). На рис.3.60, г показано напряжение на выходе детектора, которое определяется девиацией частоты $\Delta\omega$ ЧМ сигнала и крутизной K_d характеристики детектора.

Поясним работу частотного детектора. При поступлении на его вход ЧМ сигнала с постоянной амплитудой и с частотой (3.68) на выходах контуров появляются сигналы, модулированные как по частоте, так и по амплитуде.

Убедимся в появлении амплитудной модуляции, подставив в формулу (3.66) формулу (3.68) для частоты сигнала.

$$\begin{aligned} u &= \frac{2u_1}{\Delta\omega_0} (\omega_0 + \Delta\omega \cos \Omega t) - \frac{2u_1}{\Delta\omega_0} \left(\omega_0 - \frac{\Delta\omega_0}{2} \right) = \\ &= u_1 + \frac{2u_1\Delta\omega}{\Delta\omega_0} \cos \Omega t = u_1 [1 + m \cos \Omega t], \end{aligned}$$

где $m = 2\Delta\omega/\Delta\omega_0$.

Полученная огибающая выделяется амплитудным детектором.

Заметим, что ЧМ сигнал перед подачей на частотный детектор подается на двусторонний ограничитель для устранения паразитной амплитудной модуляции сигнала. Если эту паразитную АМ не устранить, то она появится на выходе частотного детектора и создаст дополнительную помеху, ухудшающую качество воспроизведения модулирующего сигнала. Причина возникновения паразитной АМ обусловлена отклонением реальной АЧХ от идеальной. Идеальная АЧХ усилителя имеет П-образный вид.

3.9.4. Детектирование сигналов с амплитудно-импульсной модуляцией

Детектором сигнала с АИМ является фильтр нижних частот (ФНЧ). Этот фильтр пропускает на выход только те спектральные компоненты, которые расположены около нулевой частоты. (см. рис.3.48, е).

Обратное преобразование Фурье от этих спектральных компонент дает сигнал на выходе ФНЧ:

$$u_{\text{вых}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\text{АИМ}}(j\omega) k(j\omega) e^{j\omega t} d\omega = \frac{A\tau_{\text{и}}}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega = \\ = \frac{A\tau_{\text{и}}}{T} b(t) = 2F\tau_{\text{и}} Ab(t),$$

где $S_{\text{АИМ}}(j\omega)$ — спектральная плотность АИМ сигнала, определяемая формулой (3.60); $k(j\omega)$ — коэффициент передачи ФНЧ; $b(t)$ — модулирующий сигнал; A , $\tau_{\text{и}}$ — соответственно амплитуда и длительность импульсов; F — верхняя граничная частота спектра сигнала $b(t)$.

Сигнал на выходе ФНЧ с точностью до постоянной величины совпадает с модулирующим сигналом $b(t)$.

3.9.5. Декодирование сигналов с импульсно-кодовой модуляцией

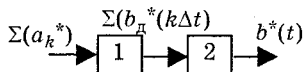


Рис.3.62. Структурная схема цифроаналогового преобразователя:

1 — преобразователь кодовых слов в выборочные значения (декодер); 2 — ФНЧ

При приеме сигнала с ИКМ двоичные кодовые слова (числа) преобразуются в значения амплитуд коротких видеоимпульсов, следующих с периодом $T = 1/2F$, где F — верхняя граничная частота первичного сигнала $b(t)$. Полученный таким способом АИМ сигнал далее декодируется с помощью ФНЧ.

Устройство, осуществляющее эти

преобразования, называется *цифроаналоговым преобразователем (ЦАП)*. Его структурная схема приведена на рис.3.62.

Теоретические основы ЦАП определены теоремой Котельникова см. разд.2.7.

3.9.6. Детектирование сигналов с широтно-импульсной и фазо-импульсной модуляцией

Детектирование ШИМ сигнала возможно с помощью ЦАП, если этот сигнал предварительно преобразуется в двоичное число. Структурная схема детектора приведена на рис.3.63.

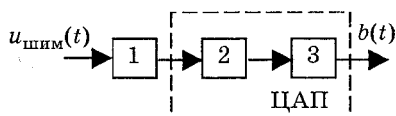


Рис.2.63. Детектор сигнала с ШИМ: 1 — преобразователь ширины импульса сигнала в двоичное число; 2 — преобразователь двоичного числа в значение амплитуды короткого видеоимпульса; 3 — ФНЧ

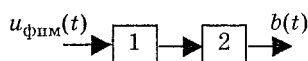


Рис.3.64. Детектор сигнала с ФИМ: 1 — преобразователь ФИМ сигнала в ШИМ сигнал; 2 — детектор ШИМ сигнала, строящийся с использованием ЦАП

В свою очередь детектор ШИМ сигнала можно использовать для детектирования ФИМ сигнала, если предварительно преобразовать его в ШИМ сигнал (рис.3.64).

3.10. Когерентное детектирование

3.10.1. Математическая модель когерентного детектора

Когерентное детектирование применяется при приеме непрерывных АМ-ПН, АМ-ПН-ОБП, ФМ и ЧМ сигналов, дискретных двоичных АМ, ЧМ, ФМ, ОФМ сигналов, а также при приеме сигналов с многократной ФМ, сигналов с КАМ и шумоподобных сигналов, модулированных по амплитуде, фазе и форме. Его можно применять и при приеме непрерывного АМ сигнала с большим уровнем несущей. Но на практике способ детектирования АМ-БУН не нашел применения из-за большей сложности когерентного детектора по сравнению с некогерентным.

Когерентное детектирование называют также *синхронным*, а когерентный детектор — *синхронным детектором*.

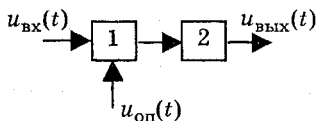


Рис.3.65. Структурная схема когерентного детектора:
1 — перемножитель;
2 — ФНЧ

Структурная схема когерентного детектора приведена на рис.3.65. Она содержит перемножитель 1 и фильтр нижних частот 2.

На первый вход перемножителя подают входной модулированный сигнал, а на второй — опорный сигнал, который генерируется (формируется) в приемнике.

Особенностью АМ-ПН и АМ-ПН-ОБП

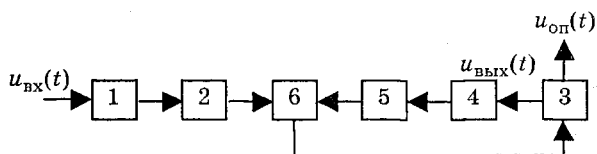


Рис.3.66. Устройство формирования опорного сигнала при АМ-ПН и АМ-ПН-ОБП:
1,4 — квадраторы; 2,5 — узкополосные фильтры, настроенные на частоту $2\omega_0$;
3 — управляемый генератор; 6 — синхронный детектор

является то, что опорный сигнал нельзя получить фильтрацией сигнала на входе детектора, так как в спектре АМ-ПН и АМ-ПН-ОБП сигналов нет спектральной компоненты на несущей частоте. Это существенно усложняет его получение.

Вариант схемы формирователя опорного сигнала $u_{оп}(t)$ показан на рис.3.66. Входной АМ-ПН или АМ-ПН-ОБП сигнал возводят в квадрат с помощью квадратора 1:

$$u_{вх}^2(t) = b^2(t) \cos^2(\omega_0 t + \varphi_0) = \frac{b^2(t)}{2} + \frac{1}{2} b^2(t) \cos(2\omega_0 t + 2\varphi_0).$$

Далее составляющую с частотой $2\omega_0$ отфильтровывают узкополосным полосовым фильтром 2. Напряжение $u_{вых}(t)$ опорного генератора 3 также возводится в квадрат, а составляющая с частотой $2\omega_0$ отфильтровывается узкополосным полосовым фильтром 5. Полученные таким образом узкополосные напряжения сравниваются по фазе с помощью фазового (синхронного) детектора 6. Выходное низкочастотное напряжение этого детектора далее используется для подстройки фазы управляемого генератора 3. На его выходе формируется опорный сигнал, необходимый для нормальной работы когерентного детектора.

В ряде случаев подстройку начальной фазы опорного сигнала ведут по специально передаваемому передатчиком гармоническому сигналу с несущей частотой ω_0 . Для него, иногда называемого пилот сигналом, выделяется отдельный узкополосный канал.

При приеме двоичного АМ сигнала опорное напряжение можно получить с помощью узкополосного полосового фильтра, настроенного на несущую частоту сигнала. Это объясняется тем, что спектральная плотность этого сигнала (рис.3.18, з) имеет спектральные составляющие на частотах $\pm\omega_0$.

Формирование опорного сигнала при двоичной дискретной ФМ производится также, как и при АМ-ПН и АМ-ПН-ОБП. Формирование опорных сигналов в схеме когерентного детектора двоичного ЧМ сигнала не отличается от формирования опорных двоичных АМ сигналов. Задача получения опорного сигнала существенно упрощается при ОФМ. Здесь в качестве опорного можно использовать сигнал на входе детектора, задержанный на время, равное одному единичному интервалу (одному такту).

Рассмотрим работу когерентного детектора, полагая, что начальная фаза опорного сигнала совпадает с начальной фазой принимаемого. Входной сигнал

$$u_{\text{вх}}(t) = b(t)\cos(\omega_0 t + \varphi),$$

где $b(t)$ — модулирующий непрерывный или дискретный сигнал; ω_0 — несущая частота; φ — начальная фаза.

Опорный сигнал

$$u_{\text{оп}}(t) = \cos(\omega_0 t + \varphi).$$

Сигнал на выходе перемножителя

$$u_{\text{п}}(t) = b(t) \cos^2(\omega t + \varphi) = b(t) \left[\frac{1}{2} + \frac{1}{2} \cos(2\omega t + 2\varphi) \right].$$

Низкочастотная часть этого сигнала

$$u_{\text{вх}}(t) = b(t)/2$$

выделяется на выходе фильтра нижних частот.

Перейдем к рассмотрению особенностей когерентного детектирования некоторых видов непрерывных и дискретных сигналов.

3.10.2. Когерентное детектирование АМ-ОБП сигнала

Пусть на вход когерентного детектора подается АМ-ОБП сигнал с нижней боковой полосой. Его спектральные компоненты определяются формулой (3.37) и зависят от спектральной плотности модулирующего сигнала (рис.3.67, а). Спектральная плотность АМ-ОБП сигнала приведена на рис.3.67, б.

В перемножителе входной сигнал умножается на когерентный опорный сигнал, спектральная плотность которого (рис.3.67, в) определяется формулой (3.29).

Спектральная плотность произведения этих двух сигналов (рис.3.67, з)

$$S_{\Pi}(j\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2} \{S_{\text{НБ}}[j(\omega_0 + z) + S_{\text{НБ}}[j(\omega_0 - z)]] \cdot \pi[\delta(\omega - z + \omega_0) + \delta(\omega - z - \omega_0)] dz = \frac{1}{4} [S_{\text{НБ}}j(2\omega_0 + \omega) + S_{\text{НБ}}(-j\omega) + S_{\text{НБ}}(j\omega) + S_{\text{НБ}}j(2\omega_0 - \omega)].$$

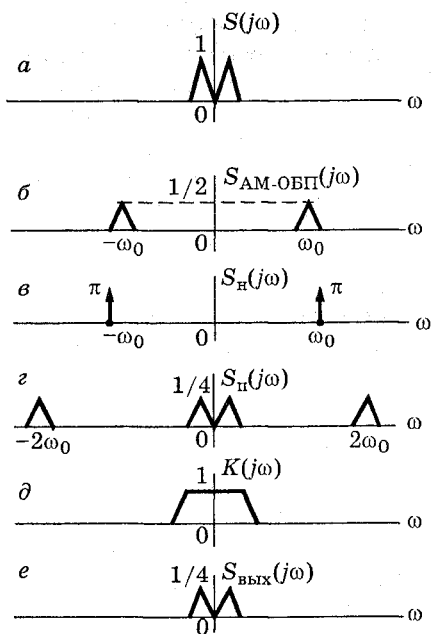


Рис.3.67. Диаграммы, поясняющие когерентное детектирование АМ-ПН-ОБП сигнала

Так как высокочастотные составляющие подавляются фильтром нижних частот, то спектральная плотность сигнала на выходе фильтра нижних частот (рис.3.67, е) определяется только теми составляющими спектральной плотности, которые группируются около нулевой частоты.

$$S_{\text{ВЫХ}}(j\omega) = S_{\Pi}(j\omega) \cdot K(j\omega) = \frac{1}{4} [S_{\text{НБ}}(-j\omega) + S_{\text{НБ}}(j\omega)] = \frac{1}{4} S(j\omega),$$

где $K(j\omega)$ — коэффициент передачи фильтра нижних частот, $S(j\omega)$ — спектральная плотность модулирующего сигнала (рис.3.67, а).

Спектральная плотность сигнала на выходе детектора по форме не отличается от спектральной плотности модулирующего сигнала. Поэтому сигнал когерентного детектора воспроизводится без искажений:

$$u_{\text{ВЫХ}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{4} S(j\omega) e^{j\omega t} d\omega = \frac{1}{4} b(t),$$

Искажения возможны при частичном подавлении боковых спектральных компонент, если условие (3.40) не выполняется и соответствующим образом не подобрана амплитудно-частотная характеристика полосового фильтра.

Найдем условия, при выполнении которых эти искажения отсутствуют. Вычислим спектральную плотность произведения АМ-ОБП сигнала и когерентного напряжения, записав спектральную плотность АМ-ОБП сигнала с частично подавленной нижней боковой полосой в виде:

$$S_{\text{AM-ОБП}}(j\omega) = \frac{1}{2} \{ S_{\text{ВБ}}[j(\omega_0 + \omega)] + S_{\text{НБ}}[j(\omega_0 + \omega)] + S_{\text{НБ}}[j(\omega_0 - \omega)] + S_{\text{ВБ}}[j(\omega_0 - \omega)] \}. \quad (3.69)$$

Составляющие спектральной плотности показаны: нижние боковые на рис.3.68, а; верхние боковые — на рис.3.68, б. Спектральная плотность АМ-ОБП сигнала, равная сумме этих составляющих, приведена на рис.3.68, в.

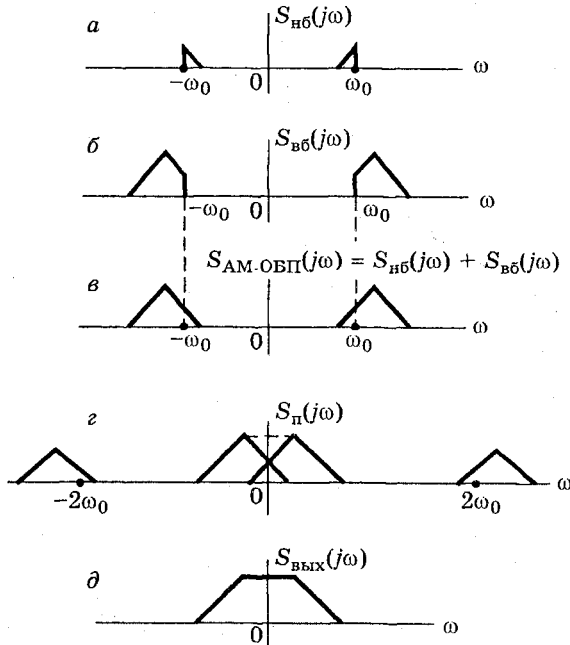


Рис.3.68. Диаграммы, поясняющие когерентное детектирование с частично подавленной боковой полосой

С учетом формулы (3.69) спектральная плотность на выходе перемножителя

$$\begin{aligned} S_{\Pi}(j\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2} \{ S_{\text{ВБ}}[j(\omega_0 + z)] + S_{\text{НБ}}[j(\omega_0 + z)] + S_{\text{НБ}}[j(\omega_0 - z)] + \\ &\quad + S_{\text{ВБ}}[j(\omega_0 - z)] \} \cdot \pi [\delta(\omega_0 + \omega - z) + \delta(-\omega_0 + \omega - z)] dz = \\ &= \frac{1}{4} \{ S_{\text{ВБ}}[j(2\omega_0 + \omega)] + S_{\text{НБ}}[j(2\omega_0 + \omega)] + S_{\text{НБ}}(-j\omega) + S_{\text{ВБ}}(-j\omega) + S_{\text{ВБ}}(j\omega) + \\ &\quad + S_{\text{НБ}}(j\omega) + S_{\text{НБ}}[j(2\omega_0 - \omega)] + S_{\text{ВБ}}[j(2\omega_0 - \omega)] \}. \end{aligned}$$

Спектральная плотность на выходе фильтра нижних частот (рис.3.68, б)

$$S_{\text{вых}}(j\omega) = S_{\Pi}(j\omega)K(j\omega) = \frac{1}{4}[S_{\text{нб}}(-j\omega) + S_{\text{вб}}(-j\omega) + S_{\text{вб}}(j\omega) + S_{\text{нб}}(j\omega)]. \quad (3.70)$$

Для неискаженного воспроизведения модулирующего сигнала необходимо, чтобы сумма спектральных плотностей, входящих в выражение (3.70), равнялась $S(j\omega)$, т.е.

$$S_{\text{вых}}(j\omega) = \frac{1}{4}[S_{\text{нб}}(-j\omega) + S_{\text{вб}}(j\omega) + S_{\text{вб}}(-j\omega) + S_{\text{нб}}(j\omega)] = \frac{1}{4}S(j\omega).$$

Отсюда условие неискаженного воспроизведения модулирующего сигнала записывается следующим образом:

$$\left. \begin{aligned} S_{\text{нб}}(-j\omega) + S_{\text{вб}}(j\omega) &= S(j\omega), \quad \omega \geq 0, \\ S_{\text{нб}}(j\omega) + S_{\text{вб}}(-j\omega) &= S(j\omega), \quad \omega \leq 0. \end{aligned} \right\} \quad (3.71)$$

Нетрудно видеть, что условия (3.40) и (3.71) идентичны и их можно обеспечить выбором АЧХ полосового фильтра передатчика.

3.10.3. Когерентное детектирование двоичных ФМ сигналов

Пусть на входе когерентного детектора напряжение

$$u_{\text{вх}}(t) = Aq(t) \cos(\omega_0 t + \varphi_1),$$

где φ_1 — начальная фаза двоичного ФМ сигнала, случайность которой обусловлена помехами, $q(t)$ — принимаемая двоичная последовательность, определяемая кодовым словом над двоичным алфавитом $-1, 1$; A — амплитуда сигнала; ω_0 — несущая частота. Напряжение опорного когерентного колебания

$$u_{\text{оп}}(t) = \cos(\omega_0 t + \varphi_2),$$

где φ_2 — начальная фаза.

Эта фаза почти повторяет фазу φ_1 входного напряжения при малых помехах, т.е. также является случайной. С ростом интенсивности помех ошибка в воспроизведении фазы опорного сигнала возрастает, а при сильных помехах эта фаза может быстро меняться на 180° .

Напряжение на выходе перемножителя когерентного детектора равно

$$Aq(t) \cos(\omega_0 t + \varphi_1) \cos(\omega_0 t + \varphi_2) = \frac{Aq(t)}{2} [\cos \Delta\varphi + \cos(2\omega_0 t + \varphi_1 + \varphi_2)],$$

где

$$\Delta\varphi = \varphi_1 - \varphi_2.$$

Оно зависит от разности фаз $\Delta\varphi$:

$$u_{\text{вых}}(t) = \frac{Aq(t)}{2} \cos \Delta\varphi.$$

Если из-за присутствия помех фаза опорного напряжения скачком меняется на 180° , то $\cos\Delta\varphi$ становится равным -1 , а

$$u_{\text{вых}}(t) = -Aq(t)/2.$$

Тогда при приеме возникает явление так называемой "обратной работы", когда вместо единиц кодового слова выделяются нули и наоборот.

3.10.4. Когерентное детектирование сигналов с ОФМ

Устранение явления "обратной работы" возможно только при переходе от ФМ к ОФМ. При двоичной ОФМ информация о символе b_i информационного кодового слова содержится в разности фаз k -го и $(k-1)$ -го элементов ОФМ сигнала. Это следует из уравнения:

$$b_k = c_k + c_{k-1},$$

которое получено из уравнения (3.55). При этом символу 0 кодового слова $\{b_k\}$, $k = 1, 2, \dots, n$ соответствует сумма одинаковых, а символу 1 — отличающихся символов, сдвинутых на один такт кодовых слов $\{c_k\}$, $\{c_{k-1}\}$. Поэтому когерентность сигнала на входе детектора достаточно сохранить в течение 2-х единичных интервалов.

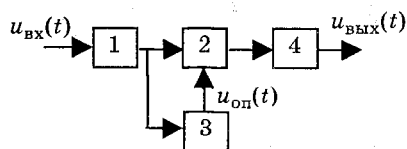


Рис.3.69. Структурная схема когерентного детектора при приеме ОФМ сигнала методом сравнения фаз: 1 — полосовой фильтр; 2 — перемножитель; 3 — линия задержки; 4 — ФНЧ

На рис.3.69 приведена схема когерентного детектора. В качестве опорного, используется входной сигнал, который задерживается на время, равное длительности единичного интервала $\tau_{\text{и}}$. Входной и опорный сигналы умножаются и результат фильтруется.

Напряжение на входе детектора в k -м единичном интервале

$$Av[t - (k-1)\tau_{\text{и}}] \cos(\omega_0 t + 2\pi k_{\text{ФМ}} c_k) = A \cos(\pi c_k) v[t - (k-1)\tau_{\text{и}}] \cos \omega_0 t,$$

где $2\pi k_{\text{ФМ}} = \pi$, $c_k \in 0,1$, а форма модулирующего сигнала $c(t)$ в k -м единичном интервале имеет прямоугольный вид.

Напряжение на выходе перемножителя в k -м единичном интервале равно

$$A^2 \cos(\pi c_k) \cdot \cos(\pi c_{k-1}) \cdot v[t - (k-1)\tau_{\text{и}}] \cdot v[t - (k-2)\tau_{\text{и}}] \times \\ \times \cos \omega_0 t \cdot \cos[\omega_0(t + \tau_{\text{и}})].$$

Напряжение на выходе ФНЧ

$$u_{\text{вых}}(t) = \frac{A^2}{2} \cos \omega_0 \tau_{\text{и}} \sum_{k=1}^n q_k q_{k-1} v[t - (k-1)\tau_{\text{и}}],$$

где $q_k = \cos \pi c_k$, причем $q_k q_{k-1} = \pm 1$.

Последняя формула получена при условии, что огибающая радиоимпульсов в каждом единичном интервале имеет также прямоугольный вид. Желательно, чтобы напряжение $u_{\text{вых}}(t)$ было наибольшим. Для этого нужно обеспечить равенство $\cos \omega_0 \tau_{\text{и}} = 1$, т.е. задержку линии выбрать равной $\tau_{\text{и}} = k/f_0$, $k = 1, 2, 3, \dots$. Из уравнения $q_k = \cos \pi c_k$ следует, что символу $c_k = 0$ соответствует символ $q_k = 1$, а символу $c_k = 1$ — символ $q_k = -1$. С учетом этого найдем, что напряжению на выходе детектора на рис.3.70, $г$ соответствует двоичное кодовое слово $\{b_k\} = 11010$.

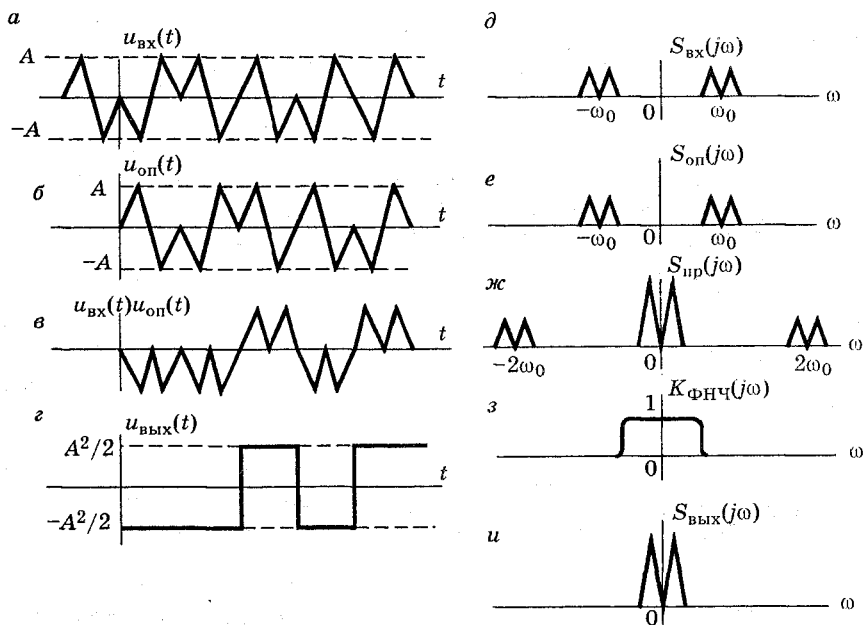


Рис.3.70. Диаграммы, поясняющие работу когерентного детектора

Для однозначного декодирования сигнала с двухкратной ОФМ (ДОФМ) необходимо, чтобы напряжение на выходе син-

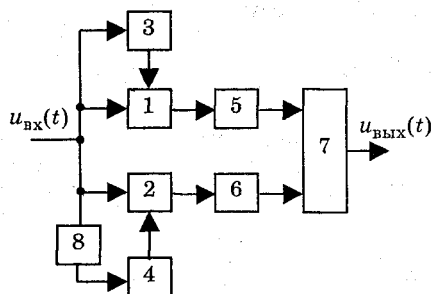


Рис.3.71. Когерентный детектор
двукратного ОФМ сигнала:

1,2 — перемножители; 3,4 — линии
задержки; 5,6 — ФНЧ;
7 — декодер; 8 — фазосдвигающая
цепь на $\pi/2$

хронного детектора соответствовало следующим значениям разности фаз: 0, $\pi/2$, π и $3\pi/2$. Детектор, схема которого приведена на рис.3.69, адекватно откликается только на разность фаз, равную 0 или π . При разности равной $\pi/2$ или $3\pi/2$ отклики одинаковы и равны нулю. В результате появляются неоднозначность декодирования слов модуляционного кода. Для исключения этой неоднозначности в схему синхронного детектора вводят второй канал. Схема синхронного детектора сигнала с ДОФМ приведена на

рис.3.71. Второй канал не отличается от первого, но при подаче входного сигнала на второй вход перемножителя 2 его фаза дополнительно сдвигается на $\pi/2$.

Если входной сигнал в k -м единичном интервале

$$u_k(t) = Av[t - (k-1)2\tau_{\text{и}}]\cos(\omega_0 t + \varphi_k),$$

то отфильтрованные ФНЧ верхнего и нижнего каналов напряжения соответственно равны

$$\frac{A^2}{2} v[t - (k-1)2\tau_{\text{и}}] \cos(\varphi_k - \varphi_{k-1}), \quad \frac{A^2}{2} v[t - (k-1)2\tau_{\text{и}}] \sin(\varphi_k - \varphi_{k-1}),$$

где $\tau_{\text{и}}$ — длительность единичного интервала двоичной информационной кодовой последовательности.

Декодирование в блоке 7 ведется по правилу, приведенному в табл.3.4.

Таблица 3.4

Разность фаз $\varphi_k - \varphi_{k-1}$	Напряжение на выходе ФНЧ		Двоичный код на выходе декодера
	верхнего канала	нижнего канала	
0	$A^2/2$	0	00
$\pi/2$	0	$A^2/2$	01
π	$-A^2/2$	0	10
$3\pi/2$	0	$A^2/2$	11

Вопросы когерентного детектирования сигналов с многократной ФМ и КАМ рассматриваются в специальных дисциплинах.

3.10.5. Детектирование шумоподобных сигналов

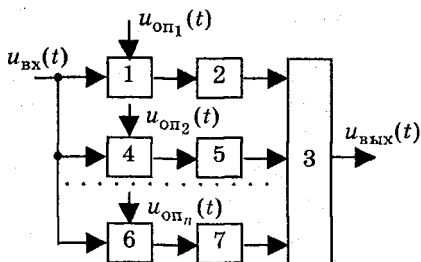


Рис.3.72. Когерентный детектор ШС с модуляцией по форме:
 1, 4, 6 — перемножители;
 2, 5, 7 — ФНЧ;
 3 — преобразователь номера канала в двоичный код

(рис.3.73) или на 2^k копий при модуляции ШС по форме (рис.3.74). Далее результат фильтруется одним (рис.3.65) или несколькими фильтрами нижних частот (рис.3.72).

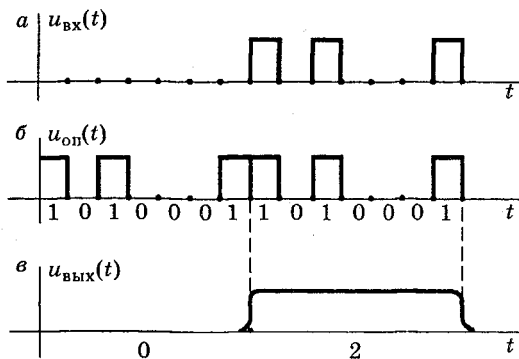


Рис.3.73. Временные диаграммы, поясняющие детектирование АМ ШС

"0", а перемножение опорного ШС на неравные нулю значения принимаемого ШС и фильтрация результата дает символ "1".

В детекторе (рис.3.72) неравное нулю напряжение на выходе одного из ФНЧ (рис.3.74, з) однозначно связано с двоичным кодовым словом передаваемого информационного блока. В передатчике от этого слова зависит форма ШС.

Цель детектирования та же, что и для простых сигналов — выделение модулирующего сигнала. Основным способом детектирования ШС является когерентный. Детектор АМ и ФМ шумоподобного сигнала не отличается от детектора, изображенного на рис.3.65, и является одноканальным. Детектор модулированного по форме ШС более сложный. Он является многоканальным (рис.3.72). При детектировании принимаемый модулированный ШС умножается или на одну копию ШС переносчика при АМ и ФМ

Напряжение на выходе ФНЧ в схеме (рис.3.65) определяет символ 0 или 1 передаваемого информационного кодового слова. Напомним, что символ "0" передается при АМ ШС отсутствием, а символ "1" — наличием ШС (рис.3.73, а). Перемножение опорного ШС (рис.3.73, б) на нулевое значение принимаемого ШС (рис.3.73, а) и фильтрация результата (рис.3.73, в) дает символ

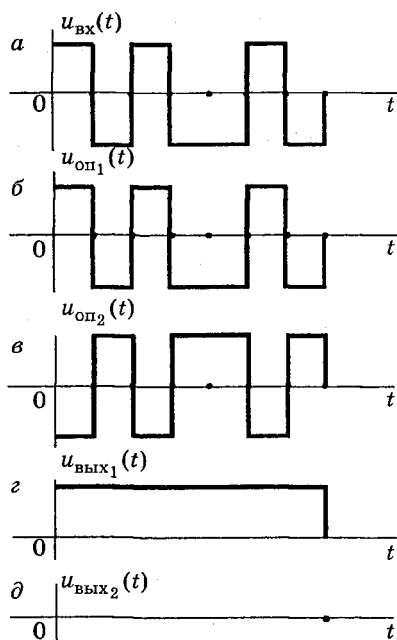


Рис.3.74. Временные диаграммы, поясняющие детектирование модулированных по форме ШС

Поясним сказанное на конкретном примере. Пусть слово 00 связано с формой ШС, показанной на рис.3.74, а, слово 01 — формой ШС, показанный на рис.3.74, б и т.д. Остальным двум словам 10 и 11 в нашем примере будут соответствовать еще две формы ШС. Таким образом в рассматриваемом примере длина передаваемого кодового блока выбрана равной 2, а число форм ШС выбрано равным 4. Когерентный детектор распознает передаваемые формы. Пусть передается форма, соответствующая кодовому слову 00 (рис.3.74, а). На перемножители всех каналов детектора одновременно с принимаемым ШС подаются отличающиеся друг от друга опорные ШС, причем опорный сигнал на первом перемножителе совпадает по форме со входным ШС. В этом случае на выходе ФНЧ первого канала будет не нулевое напряжение (рис.3.74, г), а на выходах

ФНЧ остальных каналов напряжения будут равны нулю.

При передаче ШС, соответствующего кодовому слову 01, на входе детектора будет присутствовать сигнал, совпадающий по форме с $u_{оп2}(t)$ (рис.3.74, в). Тогда не равное нулю напряжение будет на выходе второго ФНЧ, а на выходах остальных ФНЧ напряжения будут нулевые и т.д.

Преобразование номера канала, на выходе которого наблюдается не равное нулю напряжение, в двоичный код, получают кодовое слово передаваемого информационного блока.

Когерентное детектирование шумоподобных радиосигналов принципиально не отличается от рассмотренного выше детектирования на видеочастоте, но его реализация становится более сложной.

При приеме как видео-, так и радиосигналов нужно подстраивать параметры опорных ШС под параметры принимаемого сигнала. Такими параметрами являются: тактовая частота, задержка опорных сигналов, несущая частота.

Устройство, с помощью которого обеспечивается совпадение тактовых частот принимаемого и опорного сигналов, называется *устройством тактовой синхронизации*.

Устройства, обеспечивающие совпадение опорных ИС во времени и по несущей частоте называются соответственно *устройствами поиска и слежения по времени задержки ИС и по несущей частоте*. Устройства тактовой синхронизации, поиска и слежения входят в состав блока формирования опорного сигнала.

3.10.6. Особенности реализации операции умножения в устройствах формирования и преобразования сигналов

Основой для построения большинства перемножителей аналоговых сигналов служат нелинейные элементы, характеристика которых аппроксимируется полиномом (3.16). Пусть число удерживаемых членов полинома не больше трех, а входной сигнал равен сумме

$$u = a + b.$$

Тогда реакция нелинейного элемента на этот входной сигнал будет равна

$$\begin{aligned} L(a + b) &= c_0 + c_1(a + b) + c_2(a + b)^2 + \dots = \\ &= c_0 + c_1a + c_1b + c_2a^2 + 2c_2ab + c_2b^2 + \dots \end{aligned}$$

На выходе нелинейного элемента, кроме полезной составляющей $2c_2ab$, содержатся также ненужные остальные компоненты. Это создает определенные трудности при построении модуляторов и синхронных детекторов.

Устранить часть ненужных компонент можно, если применить еще один оператор:

$$L(a - b) = c_0 + c_1(a - b) + c_2(a - b)^2 = c_0 + c_1a - c_1b + c_2a^2 - 2c_2ab + c_2b^2.$$

Вычитая из $L(a + b)$ выражение $L(a - b)$, получим

$$L(a + b) - L(a - b) = 2c_1b + 4c_2ab.$$

Освободится от первого слагаемого в этой формуле можно с помощью фильтра.

Принципиальная схема модулятора, в котором реализуются приведенные выше преобразования, приведена на рис.3.75. Эту схему используют для получения АМ-ПН и АМ-ОБП сигналов.

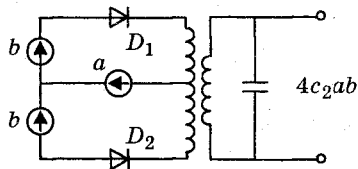


Рис.3.75. Принципиальная схема модулятора

При перемножении низко- и высокочастотных аналоговых сигналов можно использовать ключевой переключатель. Схема перемножителя приведена на рис.3.76. Когда разность потенциалов между точками

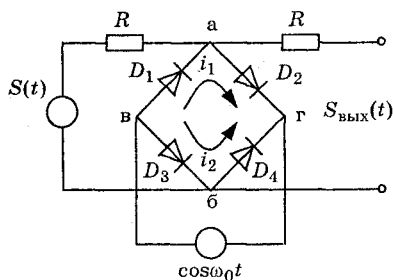


Рис.3.76. Принципиальная схема ключевого перемножителя

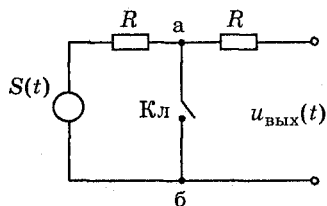


Рис.3.77. Эквивалентная схема ключевого перемножителя

в, г схемы положительная (действует положительная полуволна переменного напряжения) диоды проводят токи i_1, i_2 по направлениям, которые показаны стрелками. Сопротивление R резистора во много раз больше сопротивления r открытого диода, поэтому ответвлением тока через резисторы можно пренебречь. Если сопротивления диодов выбраны одинаковыми, то потенциал точки а равен потенциалу точки б. Это соответствует замыканию ключа в эквивалентной схеме, показанной на рис.3.77. Источник напряжения $s(t)$ нагружается на сопротивление R , а напряжение $u_{\text{вых}}(t)$ становится равным нулю.

Когда разность потенциалов между точками в, г отрицательная (действует отрицательная полуволна переменного напряжения), токи i_1, i_2 останутся равными нулю. Ключ в эквивалентной схеме размыкается, а на выходе действует входное напряжение $s(t)$.

Работа ключевого перемножителя поясняется диаграммами во временной и частотной областях (рис.3.78).

Напряжение на выходе ключа является откликом параметрической цепи на входной сигнал

$$u_{\text{вых}}(t) = s(t)K(t),$$

где $K(t)$ — коэффициент передачи параметрической цепи (ключа).

Поэтому спектральную плотность на выходе $S_{\text{вых}}(j\omega)$ можно найти по формуле свертки спектральных плотностей $S(j\omega)$, $S_k(j\omega)$. Как видно из рис.3.78, ж, спектральная плотность $S_{\text{вых}}(j\omega)$ состоит из полезной и ненужных составляющих. Полезная составляющая группируется около значений частот $\omega = \pm\omega_0$ и является спектральной плотностью произведения $s(t)\cos\omega_0 t$. Эта составляющая легко отделяется от других полосовым фильтром, настроенным на частоту ω_0 .

Ключевой перемножитель применяется в преобразователях частоты, модуляторах АМ-ПН, АМ-ОБП сигналов и в синхронных детекторах. На рис.3.79 приведена принципиальная схема модулятора, а на рис.3.80, — схема синхронного детектора АМ-ПН сигнала.

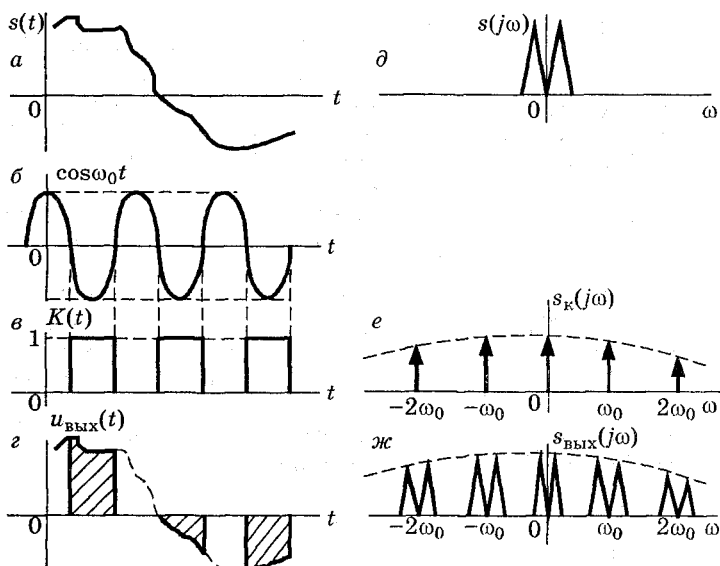


Рис.3.78. Диаграммы, поясняющие работу ключевого перемножителя

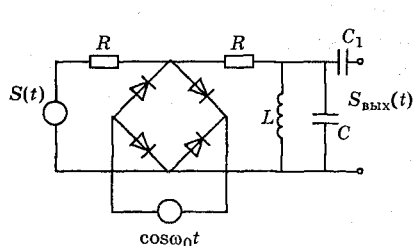


Рис.3.79. Принципиальная схема модулятора АМ-ПН сигнала

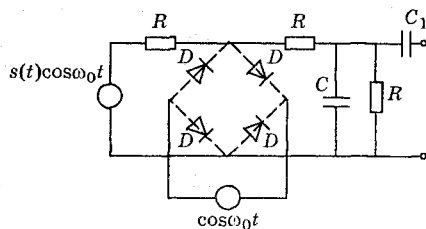


Рис.3.80. Принципиальная схема синхронного детектора АМ-ПН сигнала

3.11. Принципы цифровой обработки сигналов

3.11.1. Дискретные преобразования Фурье (прямое и обратное)

Цифровой обработкой сигналов (ЦОС) называется процесс выполнения по заданной программе вычислительных операций (суммирования, умножения, и т.д.) над числами, однозначно отображающих эти сигналы.

Преобразования ведутся комплексно, с применением аппаратных и программных средств.

При ЦОС появляется возможность выполнения операций, которые нельзя реализовать традиционными методами. Например, если требуется длительное хранение исходных или промежуточных данных или изменение алгоритмов обработки сигналов путем изменения управляющей программы и т.д.

Примерами использования ЦОС в системах связи являются: обработка звуковых сигналов, включая анализ, распознавания и синтез речи; обработка изображений с целью их улучшения; сжатие информации для ее передачи и хранения; распознавание образов.

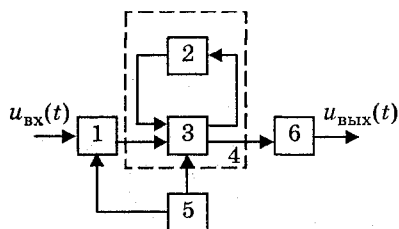


Рис.3.81. Структурная схема ЦОС: 1 — АЦП; 2 — устройство памяти; 3 — арифметическое устройство; 4 — процессор; 5 — генератор импульсов синхронизации; 6 — ЦАП

Упрощенная структурная схема ЦОС непрерывных сигналов приведена на рис.3.81. Непрерывный входной сигнал $u_{вх}(t)$ поступает на АЦП, назначением которого является преобразование этого сигнала в числа, обычно записываемые в двоичной системе счисления.

Над двоичными числами, хранящимися в устройстве памяти, и числами, поступающими с АЦП, производятся арифметические операции (умножение, сложение, сдвиг во времени) в соответствии с управляющей программой, ко-

торая также хранится в устройстве памяти.

В процессе вычислений между арифметическим устройством и устройством памяти происходит обмен данными. В соответствующие моменты времени, которые определяются генератором импульсов синхронизации, цифровой процессор выдает на выход вычисленную последовательность двоичных чисел. При необходимости, эта последовательность преобразуется в непрерывную форму с помощью цифро-аналогового преобразователя (ЦАП).

Дискретный сигнал на выходе АЦП является последовательностью двоичных чисел, выдаваемых через равные интервалы времени. Эти числа соответствуют коэффициентам обобщенного ряда (2.2) Фурье, которые, в свою очередь, определяются входным непрерывным сигналом и выбранным набором базисных функций (базисом).

В случае, когда в качестве базисных функций используются функции вида $\sin x/x$, а спектр непрерывной функции ограничен верхней граничной частотой F_v , коэффициенты ряда (2.2) совпадают с выборочными значениями непрерывной функции.

Удобным для технической реализации является также базис, состоящий из сдвинутых во времени прямоугольных импульсов (рис.3.82). Выражение для k -й базисной функции

$$\varphi_k(t) = \begin{cases} \sqrt{n/\tau_0}, (k-1)\frac{\tau_0}{n} \leq t \leq k\tau_0/n, \\ 0 \end{cases} \quad \text{для остальных значений,} \quad (3.72)$$

где n — число используемых базисных функций.

Базисные функции не перекрываются во времени, поэтому они удовлетворяют условию ортогональности. Кроме того, эти функции удовлетворяют и условию:

$$\int_{-\infty}^{\infty} \varphi_k^2(t) dt = \int_{(k-1)\tau_0/n}^{k\tau_0/n} \frac{n}{\tau_0} dt = \frac{n}{\tau_0} \cdot \frac{\tau_0}{n} = 1,$$

т.е. являются ортонормированными.

Спектральные коэффициенты обобщенного ряда Фурье, определяющие числа на выходе дискретизатора,

Рис.3.82. Базисные функции

$$a_k = \int_0^{\tau_0} u(t) \varphi_k(t) dt = \sqrt{\frac{n}{\tau_0}} \int_{(k-1)\tau_0/n}^{k\tau_0/n} u(t) dt.$$

Вычисление сводится к простой операции вычисления интеграла от непрерывного сигнала в пределах длительности k -й базисной функции. Это упрощает аппаратуру аналого-цифрового преобразования.

Запишем конечное число N выборочных значений $u(0)$, $u(1)$, ..., $u(N-1)$ и в интервале значений $(0 - N - 1)$ представим их в виде ряда Фурье

$$u(n) = \sum_{k=-\infty}^{\infty} C_k e^{j \frac{2\pi}{N} nk}, \quad (3.73)$$

где C_k — коэффициенты ряда Фурье.

Записывая функцию $u(n)$ в виде ряда Фурье, полагаем, что она периодическая с периодом N .

Ряд (3.73) получен из (2.6) заменой $1/2$ а $(jk\omega_1)$ на C_k , частоты ω_1 , на $2\pi/T = 2\pi/N\Delta$, а времени t на $n\Delta$, где Δ — интервал дискретизации.

Учитывая периодичность выборочных значений, пределы суммирования по k можно заменить от $k = 0$ до $k = N - 1$.

$$u(n) = \sum_{k=0}^{N-1} C_k e^{j \frac{2\pi}{N} nk}. \quad (3.74)$$

Умножим обе части уравнения (3.74) на $\frac{1}{N} e^{-j(2\pi/N) nq}$ и произведем суммирование по n от $n = 0$ до $N - 1$

$$\frac{1}{N} \sum_{n=0}^{N-1} u(n) e^{-j \frac{2\pi}{N} nq} = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} C_k e^{j \frac{2\pi}{N} (k-q)n}.$$

Меняем порядок суммирования в правой части последнего уравнения. В результате получим

$$\frac{1}{N} \sum_{n=0}^{N-1} u(n) e^{-j \frac{2\pi}{N} nq} = \sum_{k=0}^{N-1} C_k \left[\frac{1}{N} \sum_{n=0}^{N-1} e^{j \frac{2\pi}{N} (k-q)n} \right]. \quad (3.75)$$

Принимая во внимание, что

$$\frac{1}{N} \sum_{n=0}^{N-1} e^{j \frac{2\pi}{N} nq} = \begin{cases} 1 & \text{при } q = rN, \text{ где } r \text{ целое число,} \\ 0 & \text{при иных значениях } q. \end{cases}$$

найдем, что выражение в квадратных скобках формулы (3.75) или равно 1 при $q = k$ и равно нулю при $q \neq k$.

Следовательно,

$$C_k = \frac{1}{N} \sum_{n=0}^{N-1} u(n) e^{-j \frac{2\pi}{N} nk}. \quad (3.76)$$

Формулы (3.74), (3.76) называются *обратным и прямым дискретными преобразованиями Фурье* (ДПФ) и являются важными математическими моделями дискретных сигналов.

Пример 3.9. Дискретный сигнал задан шестью равноотстоящими отсчетами:

$$\{u(n)\} = \{1, 1, 1, 0, 0, 0\}.$$

Найти коэффициенты C_0 и C_1 ДПФ этого сигнала.

Решение. Используя формулу (3.76), вычисляем

$$\begin{aligned} C_0 &= \frac{1}{6} [u(0) + u(1) + u(2) + u(3) + u(4) + u(5)] = 3/6, \\ C_1 &= \frac{1}{6} \sum_{n=0}^5 u(n) e^{-j2\pi n/N} = \frac{1}{6} [u(0) + u(1)e^{-j2\pi/6} + u(2)e^{-j2\pi \cdot 2/6}] = \\ &= \frac{1}{6} [1 + e^{-j\pi/3} + e^{-j2\pi/3}]. \end{aligned}$$

Как видно из формул (3.74), (3.76), для вычисления ДПФ или ОДПФ последовательности из N элементов, требуется выполнить N^2 операций с комплексными числами. Если N велико (порядка тысячи и более), то вычисления в реальном масштабе времени становятся затруднительными из-за ограниченного быстродействия вычислительных средств. Решить возникшую проблему можно, используя *алгоритм быстрого преобразования Фурье* (БПФ), который позволяет существенно сократить число вычислительных операций.

3.11.2. Квантование

Продолжим анализ аналого-цифрового преобразования.

Полученные в результате дискретизации численные значения коэффициентов a_k ряда (2.2), переводятся в двоичные числа либо с их округлением либо с усечением. Поэтому при квантовании возникает ошибка

$$\Delta_k = a_k^* - a_k,$$

где a_k — число до квантования, a_k^* — число после квантования.

Поясним возникновение ошибки при округлении.

Число a_k , записанное в двоичной системе счисления,

$$a_k = \sum_{n=0}^{\infty} \alpha_n 2^{-n}, \quad (3.77),$$

где $\alpha_n = \{0,1\}$.

Запись числа в форме (3.77) означает, что в устройствах ЦОС используется нормирование обрабатываемых сигналов так, что все математические операции выполняются с числами, по абсолютному значению не превышающие единицу. Таким образом, числа отображаются правильными десятичными дробями. Далее эти дроби переводятся в двоичные числа.

Процесс квантования с помощью округления поясняется рис.3.83.

Амплитудная характеристика квантователя представлена в виде сплошной ступенчатой линии. Высота ступеньки определяется значением младшего разряда в двоичном представлении a_k^* . Это шаг квантования, равный 2^{-l} . Пунктирной линией показана характеристика идеального квантователя, когда число разрядов двоичного числа выбрано равным ∞ . Ошибка, возникающая при округлении, удовлетворяет неравенству

$$-2^{-l-1} \leq \Delta_k \leq 2^{-l-1}.$$

Если значение a_k находится между точками б, в, то погрешность округления будет отрицательной, но не более 2^{-l-1} , а если — между точками а, б, то — положительной, но не более 2^{-l-1} . Из рис.3.83 следует, что с увеличением числа разрядов двоичного числа погрешность округления уменьшается. Для уменьшения ошибки квантования следует использовать большее число разрядов. Однако это приводит к усложнению устройств ЦОС и снижению их быстродействия.

После преобразования непрерывного сигнала в совокупность двоичных чисел над ними ведутся дальнейшие преобразования. Рассмотрим некоторые из них.

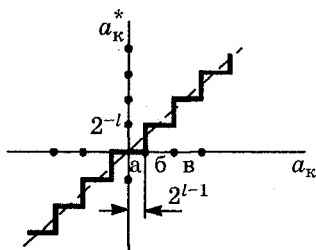


Рис.3.83. К пояснению процесса квантования

3.11.3. Прямое и обратное z-преобразование

Линейная цифровая фильтрация заимствует основные положения теории линейных систем.

Линейный цифровой фильтр (ЦФ) преобразует последовательность числовых отсчетов входного сигнала $\{a_k\}$ в последовательность числовых отсчетов $\{b_k\}$ выходного сигнала

$$(a_0, a_1, \dots, a_k \dots) \rightarrow (b_0, b_1, \dots, b_k \dots)$$

по формуле свертки, аналогичной (3.7)

$$b_m = \sum_{k=0}^m a_k h_{m-k}, \quad (3.78)$$

где $\{h_k\}$ — дискретный сигнал, являющийся реакцией цифрового сигнала на единичный импульс $(1, 0, 0, \dots)$.

Смысл формулы прост и нагляден: значение b_m выходного сигнала равно взвешенной сумме a_m и всех предыдущих значений входного сигнала, причем роль весовых коэффициентов выполняют отсчеты импульсной характеристики ЦФ. ЦФ, также как и линейный, обладает "памятью" по отношению к прошлым воздействиям.

Для физически реализуемых систем реакция ЦФ равна нулю до подачи входного импульса, поэтому коэффициенты $h_{-1}, h_{-2}, \dots$ импульсной характеристики равны нулю и суммирование в (3.78) можно распространить на все положительные значения индекса k .

$$b_m = \sum_{k=0}^{\infty} a_k h_{m-k}. \quad (3.79)$$

В частотной области ЦФ характеризуется коэффициентом передачи

$$K(j\omega) = \sum_{n=0}^{\infty} e^{-j\omega n\Delta} h_n, \quad (3.80)$$

где h_n — выборочные значения импульсной характеристики ЦФ.

Применяя к (3.80) обратное преобразование Фурье, получаем импульсную характеристику ЦФ в виде следующей последовательности

$$h(t) = h_0 \delta(t) + h_1 \delta(t - \Delta) + \dots$$

Значение для $K(j\omega)$ найдем по аналогии с реакцией линейной цепи на гармонический сигнал.

При дискретизации этого сигнала получим

$$\{a_k\} = \{A \exp[j(\omega k\Delta + \varphi)]\}.$$

Выходной сигнал ЦФ найдем по формуле свертки (3.78).

$$\{b_m\} = \sum_{k=-\infty}^m a_k h_{m-k} = A e^{j\varphi} \sum_{k=-\infty}^m e^{j\omega k\Delta} h_{m-k} = A e^{j(\omega m\Delta + \varphi)} \sum_{k=-\infty}^m e^{j\omega(k-m)\Delta} h_{m-k}.$$

Заменим индекс $m - k$ на n . В результате получим

$$\{b_m\} = Ae^{j(\omega m\Delta + \varphi)} \sum_{n=-\infty}^m e^{j\omega n\Delta} h_n = Ae^{j(\omega m\Delta + \varphi)} K(j\omega),$$

где $K(j\omega)$ — коэффициент передачи, определяемый формулой (3.80).

Расчет коэффициента передачи ЦФ ведут также с использованием z -преобразований.

Каждой числовой последовательности $\{u_k\} = \{u_0, u_1, \dots, u_m\}$, соответствующей выборочным значениям сигнала, можно поставить в однозначное соответствие сумму ряда по отрицательным степеням комплексной переменной:

$$u(z) = u_0 + u_1/z + u_2/z^2 + \dots = \sum_{k=1}^{\infty} u_k z^{-k}. \quad (3.81)$$

Сумма (3.81) называется z -преобразованием числовой последовательности $\{u_k\}$.

Умножим левую и правую части уравнения (3.81) на множитель z^{m-1} :

$$z^{m-1}u(z) = u_0 z^{m-1} + u_1 z^{m-2} + \dots + u_m z^{-1} + \dots$$

Полученный результат проинтегрируем по замкнутой кривой, которая лежит в области аналитичности функции $u(z)$ и охватывает все ее полюсы. Примем во внимание известное из математики выражение

$$\oint z^n dz = \begin{cases} 2\pi j, & \text{если } n = -1, \\ 0, & \text{если } n \neq -1. \end{cases}$$

В результате получим

$$u_m = \frac{1}{2\pi j} \oint z^{m-1} U(z) dz. \quad (3.82)$$

Выражение (3.82) является обратным z -преобразованием.

Для числовой последовательности $\{v_k\}$ на выходе ЦФ аналогично находится соответствующее z -преобразование $V(z)$, причем

$$V(z) = U(z)H(z),$$

где $H(z)$ — системная функция.

Системная функция равна отношению z -преобразования выходного сигнала к z -преобразованию входного

$$H(z) = V(z)/U(z) = \sum_{k=0}^{\infty} h_k z^{-k}, \quad (3.83)$$

где h_k — числовая последовательность, соответствующая импульсной характеристике.

Формула (3.83) определяет системную функцию как z -преобразование от числовой последовательности $\{h_k\}$. Сравнивая (3.80) и (3.83), определим правило нахождения коэффициента передачи $K(j\omega)$ по системной функции $H(z)$: для этого в формуле (3.83) нужно сделать подстановку $z = e^{j\omega\Delta}$.

Пример 3.10. Импульсная характеристика ЦФ определяется последовательностью, содержащую два ненулевых отсчета: $\{h_k\} = (1, -1, 0, \dots)$.

Вычислить частотный коэффициент передачи $K(j\omega)$ ЦФ.

Решение. Системная функция имеет вид

$$H(z) = 1 - z^{-1}.$$

Подставляя в полученную формулу $z = e^{-j\omega\Delta}$ получим

$$K(j\omega) = 1 - e^{-j\omega\Delta} = 1 - \cos \omega\Delta + j \sin \omega\Delta.$$

Рассмотрим примеры реализации ЦФ.

3.11.4. Трансверсальные цифровые фильтры

Под трансверсальными фильтрами понимают такие ЦФ, которые работают в соответствии с алгоритмом

$$v_i = a_0 u_i + a_1 u_{i-1} + a_2 u_{i-2} + \dots + a_m u_{i-m}, \quad (3.84)$$

где $a_0, a_1, a_2, \dots, a_m$ — последовательность постоянных коэффициентов; m — число, определяющее порядок фильтра.

Трансверсальный фильтр производит взвешенное суммирование предшествующих отсчетов входного сигнала.

Применив z -преобразование к обеим частям уравнения (3.84), получим

$$V(z) = (a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_m z^{-m})U(z).$$

Отсюда системная функция

$$H(z) = a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_m z^{-m} = \frac{a_0 z^m + a_1 z^{m-1} + \dots + a_m}{z^m}. \quad (3.85)$$

Частотную характеристику фильтра получаем по уравнению (3.85) заменой $z = e^{j\omega\Delta}$.

$$K(j\omega) = a_0 + a_1 e^{-j\omega\Delta} + a_2 e^{-j2\omega\Delta} + \dots + a_m e^{-jm\omega\Delta}.$$

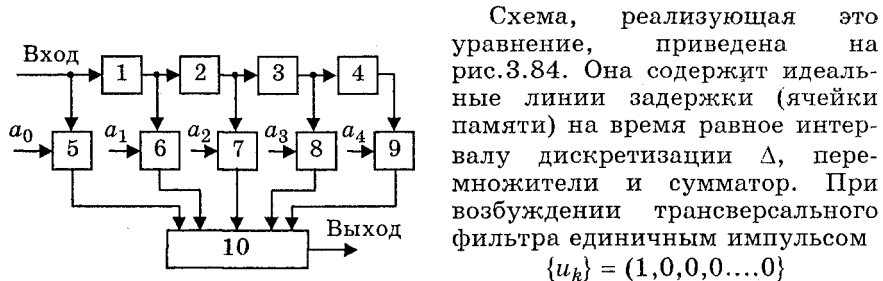


Рис.3.84. Схема трансверсального фильтра: 1, 2, 3, 4 — устройства задержки на 1 такт; 5, 6, 7, 8, 9 — перемножители; 10 — сумматор

Схема, реализующая это уравнение, приведена на рис.3.84. Она содержит идеальные линии задержки (ячейки памяти) на время равное интервалу дискретизации Δ , перемножители и сумматор. При возбуждении трансверсального фильтра единичным импульсом $\{u_k\} = (1, 0, 0, 0, \dots, 0)$

и перемещения содержимого ячеек памяти путем сдвига через время дискретизации Δ на его выходе формируются выборочные значения импульсной

характеристики ЦФ. При подаче на его вход совокупности выборок $\{u_k\}$, $k = 0, \dots, m$ и перемещении содержимого ячеек памяти на его выходе реализуется совокупность дискретных сигналов, определяемая формулой свертки (3.79). Элементы задержки трансверсального фильтра должны обеспечивать задержку l -разрядных двоичных чисел, а перемножители и сумматоры — соответствующие математические операции с этими числами.

3.11.5. Рекурсивный цифровой фильтр

Этот фильтр работает по следующему алгоритму:

$$v_i = a_0 u_i + a_1 u_{i-1} + \dots + a_m u_{i-m} + b_1 v_{i-1} + b_2 v_{i-2} + \dots + b_n v_{i-n}, \quad (3.86)$$

где $a_0, a_1, \dots, a_m$; $b_1, b_2, \dots, b_n$ — постоянные коэффициенты, не равные нулю.

В уравнении (3.86) участвуют выборки как входного, так и выходного сигналов.

Системная функция рекурсивного фильтра определяется выражением

$$H(z) = \frac{V(z)}{U(z)} = \frac{a_0 z^n + a_1 z^{n-1} + \dots + a_m z^{n-m}}{z^n - b_1 z^{n-1} + \dots - b_n}.$$

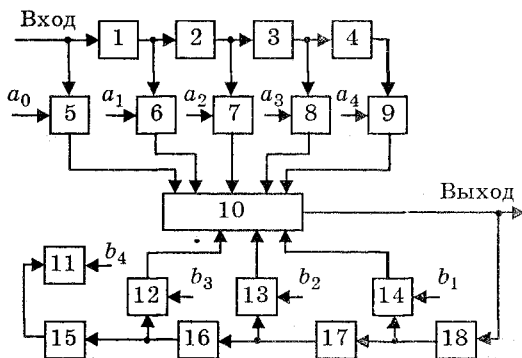


Рис.3.85. Схема рекурсивного фильтра: 1, 2, 3, 4, 15, 16, 17, 18 — устройства задержки на 1 такт; 5, 6, 7, 8, 9, 11, 12, 13, 14 — перемножители; 10 — сумматор

Схема рекурсивного ЦФ приведена на рис.3.85. Она содержит элементы задержки (памяти), умножители и сумматор, обеспечивающий математические операции l -разрядными двоичными числами.

При подаче на его вход совокупности выборок $\{u_k\}$, $k = 0, \dots, m-1$ и перемещения содержимого ячеек памяти через время дискретизации Δ на его выходе

формируются выборки выходного сигнала в соответствии с формулой (3.86).

3.11.6. Пример синтеза линейных цифровых фильтров

Под синтезом ЦФ понимают совокупность математических преобразований для определения структуры ЦФ. В результате

синтеза получают математическое выражение, из которого следует структура требуемого ЦФ.

В общем случае синтез структуры ЦФ осуществляется путем применения z -преобразования к совокупности выборочных значений импульсной характеристики

$$\{h_k\} = \{h_0, h_1, \dots, h_m\}.$$

Предполагается, что импульсная характеристика обращается в нуль при $t < 0$, что означает, что ЦФ является физически реализуемым фильтром. Если число выборочных значений конечно, синтезируемый ЦФ будет трансверсальным. При неограниченно протяженной импульсной характеристике получают рекурсивный фильтр.

В результате преобразований находят системную функцию $H(z)$. Затем заменой z на $e^{j\omega\Delta}$ находят коэффициент передачи ЦФ. Коэффициент передачи является искомым математическим выражением, которое определяет структуру ЦФ.

Пример 3.11. Синтезировать трансверсальный ЦФ, подобный интегрирующей RC -цепи, с импульсной характеристикой

$$h(t) = \begin{cases} 0, & t < 0, \\ \exp(-t/\tau), & t > 0. \end{cases}$$

Решение. Пусть импульсная характеристика трансверсального ЦФ содержит три равноотстоящих отсчета:

$$\{h_k\} = \{1, e^{-\Delta/\tau}, e^{-2\Delta/\tau}\}.$$

Применим z -преобразование к приведенной совокупности отсчетов. Получим системную функцию ЦФ

$$H(z) = 1 + e^{-\Delta/\tau} z^{-1} + e^{-2\Delta/\tau} z^{-2}.$$

Коэффициент передачи $K(j\omega)$, определяющий структуру трансверсального ЦФ:

$$K(j\omega) = 1 + e^{-\Delta/\tau} e^{-j\omega\Delta} + e^{-2\Delta/\tau} e^{-j2\omega\Delta}.$$

3.11.7. Цифровые модуляторы и детекторы

Принцип работы цифровых модуляторов и детекторов основан на взаимнооднозначном соответствии между модулированными сигналами, и их спектральными коэффициентами.

Изменение спектральных коэффициентов, адекватное модулированному сигналу (АМ, ЧМ, ФМ или ОФМ), можно рассчитывать по заранее составленной программе. Эти операции производятся в блоке расчета коэффициентов (рис.3.86).

Рис.3.87 иллюстрирует пример расчета спектральных коэффициентов для синусоидального сигнала в базисе ортонормированных функций типа (3.72).

Очевидно, что при изменении амплитуды A гармонического колебания изменяются и амплитуды спектральных компонент $\{a_k\}$.

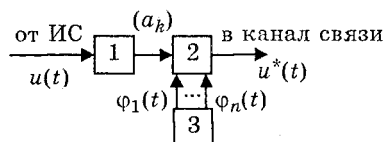


Рис.3.86. Структурная схема модулятора: 1 — блок расчета, 2 — синтезатор сигнала; 3 — генератор базисных функций

С помощью рис.3.87 можно убедиться и в том, что сдвиг гармонического колебания во времени (ФМ), или изменение периода колебания (ЧМ), также приведет к изменению численных значений спектральных компонент $\{a_k\}$.

Из процедуры вычисления $\{a_k\}$ становится очевидным недостаток рассматриваемого принципа, связанный с необходимостью преобразований быстропротекающих процессов. Для этого требуются высокопроизводительные аппаратные и программные средства.

В синтезаторе сигнала производится создание модулированного сигнала на основе рассчитанных коэффициентов и набора базисных функций. С этой целью эти функции умножаются на соответствующие коэффициенты и результаты умножений складываются (см. формулу (2.2) для обобщенного ряда Фурье).

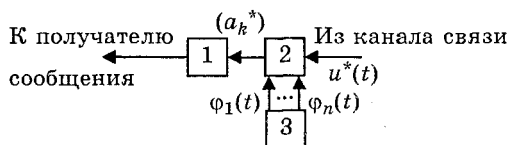


Рис.3.88. Структурная схема демодулятора: 1 — блок обработки; 2 — анализатор спектра; 3 — генератор базисных функций

На выходе синтезатора формируется сигнал $u^*(t)$.

В цифровом демодуляторе осуществляются обратные операции. В анализаторе спектра (рис.3.88) вычисляются спектральные коэффициенты, по которым в блоке обработки сигнала восстанавливается первичный сигнал.

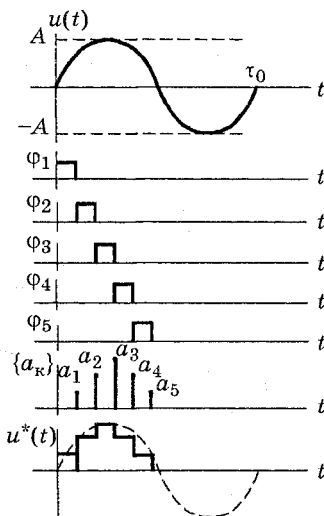


Рис.3.87. К расчету спектральных коэффициентов для синусоидального сигнала

В качестве синтезатора сигнала может быть использован вариант рассмотренного ранее трансверсального фильтра, выполненный на регистре сдвига. На его вход подается базисная функция вида ϕ_i , а на входы перемножителей — значения спектральных коэффициентов $\{a_k\}$. На

Контрольные вопросы

1. Что Вы понимаете под математической моделью блока или системы?
2. Каким условиям удовлетворяют операторы: линейной, параметрической и нелинейной систем?
3. Дайте определение линейной, параметрической и нелинейной системы.
4. Как формулируется и решается задача преобразования дискретного спектра линейной, параметрической и нелинейной системами?
5. Как формулируется и решается задача преобразования спектральной плотности сигнала линейной, параметрической и нелинейной системами?
6. Какие способы аппроксимации вольт-амперных характеристик нелинейных элементов (диодов, транзисторов, ламп) Вы знаете?
7. Что Вы понимаете под преобразованием частоты сигнала?
8. Какие виды амплитудной непрерывной модуляции Вы знаете?
9. Что Вы понимаете под индексом амплитудной модуляции?
10. Чем отличаются спектры АМ, АМ-ПН, АМ-ПН-ОБП сигналов?
11. Какие методы детектирования АМ непрерывных сигналов Вам известны?
12. В чем заключается сходство и различие между ФМ и ЧМ сигналами?
13. Как связаны между собой частота модуляции, девиация частоты и индекс модуляции?
14. В чем сходство и различие спектра ЧМ сигнала с малым индексом модуляции от спектра АМ сигнала?
15. Какие методы формирования ЧМ сигналов Вы знаете?
16. Перечислите виды модуляций импульсного переносчика. В чем сходство и отличие спектров этих сигналов?
17. Какие виды дискретной модуляций непрерывного гармонического сигнала Вы знаете?
18. Какие преимущества имеют сигнал с многократной фазовой модуляцией по сравнению с сигналом дискретной двоичной ФМ?
19. В чем сходство и отличие между сигналами: многократной ФМ и КАМ сигналом?
20. В чем сходство и отличие трансверсального и рекурсивного ЦФ?
21. На каком принципе строятся цифровые модуляторы и детекторы систем связи?

Глава 4. СЛУЧАЙНЫЕ СИГНАЛЫ И ИХ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

4.1 Математические модели непрерывных случайных процессов

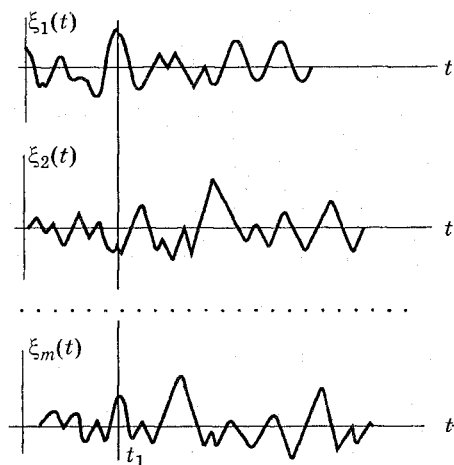


Рис.4.1. Ансамбль реализаций случайного процесса

Случайной называется такая функция, которая может принять тот или иной конкретный вид, неизвестно заранее, какой именно. Этот вид называется *реализацией*. Случайные функции зависят от времени и других параметров. Примеры реализаций случайных функций, зависящих от времени, приведены на рис.4.1 и рис.4.2. На рис.4.1. реализации более хаотичны, чем на рис.4.2. Совокупность реализаций называется *ансамблем*.

Проведем сечение ансамбля при некотором значении времени t . В результате получим несколько случайных

выборочных значений. Их число равно числу реализаций.

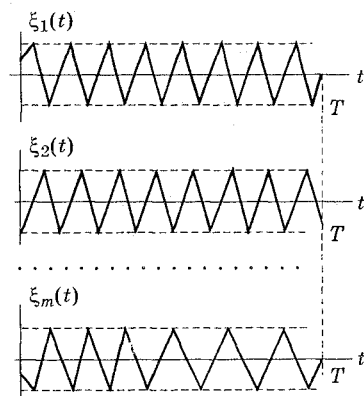


Рис.4.2. Ансамбль реализаций сигнала со случайной начальной фазой

Случайная функция $\xi(t)$ считается заданной, если для последовательных моментов времени $t_1, t_2, \dots, t_n$ и заранее выбранных значений неслучайных аргументов $x_1, x_2, \dots, x_n$ можно вычислить вероятность того, что случайные выборочные значения $\xi(t_1), \xi(t_2), \dots, \xi(t_n)$ соответственно равны или меньше $x_1, x_2, \dots, x_n$.

Эта вероятность является n -мерной интегральной функцией распределения:

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) =$$

$$P\{\xi(t_1) \leq x_1; \xi(t_2) \leq x_2; \dots; \xi(t_n) \leq x_n\}. \quad (4.1)$$

Вычисление функции (4.1) предполагает переход от непрерывных во времени реализаций $\xi(t)$ к их выборочным значениям $\xi(t_1), \xi(t_2), \dots, \xi(t_n)$, взятых в моменты времени $t_1, t_2, \dots, t_n$. Очевидно, что чем большее число реализаций и выборочных значений задано, тем более точно в статистическом смысле описывается случайный процесс.

Если существуют частные производные функции распределения по всем x_n , то можно найти n -мерную плотность распределения вероятности

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \partial^n F \frac{(x_1, x_2, \dots, x_n; t_1, \dots, t_n)}{\partial x_1, \partial x_2, \dots, \partial x_n}. \quad (4.2)$$

Зная n -мерную плотность, всегда можно получить m -мерную плотность для меньшей группы случайных величин ($m < n$) интегрированием этой n -мерной плотности по остальным ($n - m$) переменным. Например, двумерная плотность находится по n -мерной интегрированием по $n - 2$ переменным:

$$w(x_1, x_2; t_1, t_2) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_3 \dots dx_n. \quad (4.3)$$

Важным понятием является *условная плотность распределения вероятности*.

Примером условной плотности, выраженной через двумерную и одномерную плотности вероятности, является плотность:

$$w(x_1, t_1 / x_2, t_2) = \frac{w(x_1, x_2; t_1, t_2)}{w(x_2, t_2)}. \quad (4.4)$$

Формула (4.4) может быть обобщена на большее число переменных и означает, что знание о случайной величине $\xi(t_2)$ Дает и информацию о величине $\xi(t_1)$. Из формулы (4.4) также следует, что двумерная плотность распределения вероятности

$$w(x_1, x_2; t_1, t_2) = w(x_2, t_2)w(x_1, t_1 / x_2, t_2) = w(x_1, t_1)w(x_2, t_2 / x_1, t_1). \quad (4.5)$$

Если выборочные значения случайной функции независимы, то n -мерная плотность распределения вероятности равна произведению одномерных плотностей

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = w_1(x_1, t_1)w_2(x_2, t_2) \dots w_n(x_n, t_n). \quad (4.6)$$

В этом случае исчерпывающей характеристикой случайной функции является ее *одномерная плотность*.

В ряде задач вместо n -мерной плотности вероятности достаточно ограничиться такими числовыми характеристиками, как *моменты распределения*.

По одномерной плотности распределения вероятности можно вычислить начальные моменты распределения, в том числе *моменты 1-го и 2-го порядка*.

Момент 1-го порядка является средним значением случайной выборки при некотором значении времени t :

$$M\{\xi(t)\} = \int_{-\infty}^{\infty} xw_1(x, t)dx = \overline{\xi(t)} = a(t). \quad (4.7)$$

Момент 2-го порядка равен среднему от квадрата случайной выборки:

$$M\{\xi^2(t)\} = \int_{-\infty}^{\infty} x^2 w(x, t)dx. \quad (4.8)$$

Разность между случайным процессом и его математическим ожиданием называется *центрированным процессом* и обозначается

$$\overset{\circ}{\xi}(t) = \xi(t) - \overline{\xi(t)} = \xi(t) - a(t).$$

Математическое ожидание квадрата центрированной случайной выборки называется *дисперсией*:

$$D\{\xi(t)\} = M\{[\xi(t) - a(t)]^2\} = \int_{-\infty}^{\infty} [x - a(t)]^2 w(x, t)dx = \sigma^2(t). \quad (4.9)$$

Дисперсия является мерой разброса выборочных значений около их среднего значения.

Ряд задач решается с использованием двумерной функции распределения. Знание двумерной плотности распределения вероятности позволяет вычислить *смешанный второй начальный момент* двумерного закона распределения или *функцию корреляции*:

$$B(t_2, t_1) = M\{\xi(t_1), \xi(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (4.10)$$

Для центрированного случайного процесса функция корреляции

$$R(t_2, t_1) = M\{\overset{\circ}{\xi}(t_1), \overset{\circ}{\xi}(t_2)\} = \int \int [x_1 - a(t_1)][x_2 - a(t_2)] \times \\ \times w(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (4.11)$$

Статистическая связь между двумя случайными процессами $\xi(t)$ и $\gamma(t)$ в моменты времени t_1, t_2 определяется *функцией взаимной корреляции*

$$M\{\xi(t_1), \gamma(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyw(x, y; t_1, t_2) dx dy. \quad (4.12)$$

Если случайные функции статистически независимы, то двумерную плотность можно записать в виде произведения одномерных плотностей:

$$w(x, y; t_1, t_2) = w(x, t_1)w(y, t_2).$$

При этом формула (4.12) упрощается:

$$M\{\xi(t_1), \gamma(t_2)\} = \int_{-\infty}^{\infty} xw(x, t_1)dx \int_{-\infty}^{\infty} yw(y, t_2)dy. \quad (4.13)$$

Для центрированных случайных процессов функция взаимной корреляции

$$M\{\xi(t_1), \gamma(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - a(t_1)][y - a(t_2)] \cdot w(x, y; t_1, t_2) dx dy. \quad (4.14)$$

Для центрированных статистически независимых случайных процессов функция взаимной корреляции

$$M\{\xi(t_1), \gamma(t_2)\} = \int_{-\infty}^{\infty} [x - a(t_1)]w(x, t_1)dx \int_{-\infty}^{\infty} [y - a(t_2)]w(y, t_2)dy = 0. \quad (4.15)$$

Под *стационарным* понимают такой случайный процесс, n -мерная плотность распределения вероятности которого не меняется при любом сдвиге всей группы точек $t_1, t_2, \dots, t_n$ вдоль оси времени. Это условие записывается в виде:

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = w(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau), \quad (4.16)$$

где τ — временной сдвиг.

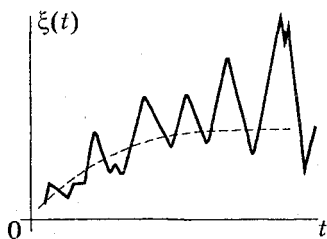


Рис.4.3. Реализация нестационарного случайного процесса

Случайные процессы, удовлетворяющие этому условию, называются *стационарными в узком смысле*. Случайный процесс, у которого математическое ожидание и дисперсия не зависят от времени, а функция корреляции зависит только от разности $\tau = t_2 - t_1$, называется *стационарным в широком смысле*. Случайные процессы, не удовлетворяющие условию (4.16), называются *нестационарными*. Реализация нестационарного процесса приведена на рис.4.3. У этой

реализации меняется не только среднее значение, но и дисперсия.

Стационарный случайный процесс называется *эргодическим*, если любая его реализация имеет одни и те же статистические свойства (рис.4.1). Поэтому для эргодических случайных процессов любая статистическая характеристика, полученная усреднением по множеству реализаций, может быть получена также усреднением за достаточно большой промежуток времени по одной реализации случайного процесса.

Например:

$$M\{\xi(t)\} = \int_{-\infty}^{\infty} xw(x, t)dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t)dt = \overline{\xi(t)}; \quad (4.17)$$

$$M\{\xi^2(t)\} = \int_{-\infty}^{\infty} x^2w(x, t)dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi^2(t)dt = \overline{\xi^2(t)};$$

$$\begin{aligned} M\{[\xi(t) - a(t)]^2\} &= \int_{-\infty}^{\infty} [x - a(t)]^2 w(x, t)dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\xi(t) - a(t)]^2 dt = D\{\xi(t)\}; \end{aligned} \quad (4.18)$$

$$\begin{aligned} B(t_1, t_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w(x_1, x_2; t_1, t_2) dx_1 dx_2 = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int \xi(t) \xi(t + \tau) dt = B(\tau), \end{aligned} \quad (4.19)$$

где $\tau = t_2 - t_1$ — разность моментов времени;

$$\begin{aligned} B_{x,y}(t_1, t_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyw(x, y; t_1, t_2) dx dy = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int \xi(t) \gamma(t + \tau) dt = B_{xy}(\tau). \end{aligned} \quad (4.20)$$

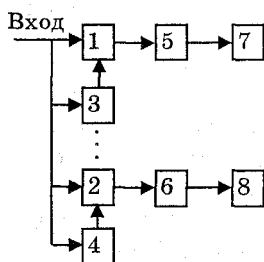


Рис.4.4. Структурная схема коррелометра:
1,2 — перемножители;
3,4 — линии задержки;
5,6 — интеграторы;
7,8 — вольтметры

Эргодические процессы чаще других используются в качестве математических моделей сообщений сигналов и помех. Это объясняется тем, что свойство эргодичности позволяет рассматривать не множество, а всего одну реализацию. Например, для вычисления функции автокорреляции шума на выходе усилителя достаточно подать его на многоканальный коррелометр (рис.4.4) и измерить напряжения на выходах его каналов. Каждый канал содержит перемножитель, линию задержки, интегратор и вольтметр. Линии задержки задерживают входную реализацию шума на заданные величины, перемножители умножают входную

реализацию шума на задержанные реализации, а интеграторы осуществляют интегрирование перемноженных реализаций. Вольтметрами измеряются значения автокорреляционной функции.

4.2. Математические модели случайных дискретных процессов

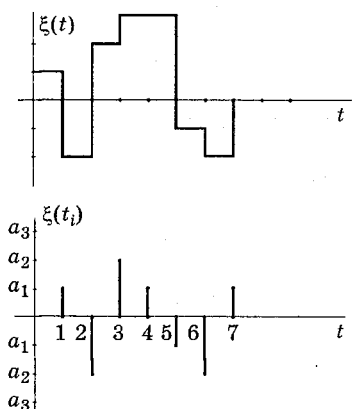


Рис.4.5. Примеры дискретных процессов

Случайные процессы включают в себя, кроме рассмотренных выше непрерывных случайных процессов непрерывного времени, также дискретные процессы. Примерами дискретных процессов являются:

а) дискретные процессы непрерывного времени (рис.4.5, а);

б) дискретные процессы дискретного времени (рис.4.5, б).

Математическими моделями дискретных процессов являются случайные последовательности. Рассмотрим математическую модель дискретного процесса дискретного времени (рис.4.5, б). Процесс математически задан, если для счетного числа моментов времени t_i и счетного числа дискретных значений a_{r_i} процесса задана совместная вероятность

$$P(\xi_1 = a_{r_1}, \xi_2 = a_{r_2}, \dots, \xi_N = a_{r_N}) = P(\xi_1 = a_{r_1}) \cdot P(\xi_2 = a_{r_2} / \xi_1 = a_{r_1}) \dots \\ \dots P(\xi_N = a_{r_N} / \xi_{N-1} = a_{r_{N-1}}, \xi_{N-2} = a_{r_{N-2}}, \dots, \xi_1 = a_{r_1}), \quad (4.21)$$

где $i = 1, 2, \dots, N$; $r = -m, -m+1, \dots, 0, 1, 2, \dots, m$; a_{r_i} — дискретное значение процесса в момент времени t_i ; $P(\xi_k = a_{rk} / \xi_{k-1} = a_{rk-1}, \xi_{k-2} = a_{rk-2}, \dots, \xi_1 = a_{r_1})$ — вероятность появления в момент времени t_k , $N \geq k \geq 1$ дискретного значения a_{rk} при условии, что в предыдущие моменты времени $t_{k-1}, t_{k-2}, \dots, t_1$ дискретные значения были $a_{rk-1}, a_{rk-2}, \dots, a_{r_1}$.

Дискретный случайный процесс, определяемый вероятностью (4.21), иногда называют дискретным процессом с памятью, имея ввиду статистическую зависимость между значениями этого процесса в моменты времени $t_1, t_2, \dots, t_N$.

Дискретный случайный процесс является *простым марковским* процессом, если условная вероятность зависит только от соседних моментов времени:

$$P(\xi_k = a_{rk} / \xi_{k-1} = a_{rk-1}, \xi_{k-2} = a_{rk-2}, \dots, \xi_1 = a_{r_1}) = \\ = P(\xi_k = a_{rk} / \xi_{k-1} = a_{rk-1}). \quad (4.22)$$

Для простого марковского дискретного случайного процесса совместная вероятность

$$P(\xi_1 = a_{r_1}, \xi_2 = a_{r_2}, \dots, \xi_N = a_{r_N}) = P(\xi_1 = a_{r_1}) \cdot P(\xi_2 = a_{r_2} / \xi_1 = a_{r_1}) \dots \\ \dots P(\xi_N = a_{r_N} / \xi_{N-1} = a_{r_{N-1}})$$

определяется вероятностью того, что в момент времени t_1 значение дискретного процесса равно a_{r_1} , т.е. $P(\xi_1 = a_{r_1})$, и условными вероятностями $P(\xi_k = a_{r_k} / \xi_{k-1} = a_{r_{k-1}})$, которые называются вероятностями перехода.

Простой марковский процесс также является дискретным случайным процессом с памятью, но статистическая зависимость имеется только между значениями процесса в соседние моменты времени.

Наиболее простой моделью является *дискретная случайная последовательность* с независимыми дискретными значениями (*последовательность Бернулли*). Такая дискретная последовательность определяется совместной вероятностью, равной произведению вероятностей:

$$P(\xi_1 = a_{r_1}, \xi_2 = a_{r_2}, \dots, \xi_N = a_{r_N}) = \prod_{i=1}^N P(\xi_i = a_{r_i}). \quad (4.23)$$

Эта математическая модель описывает сообщение *дискретного источника без памяти*.

Дискретный источник называется *стационарным*, если его статистическое описание не зависит от начала отсчета времени.

4.3. Функция корреляции и энергетический спектр

Функция корреляции отражает степень статистической связи между выборочными значениями случайного процесса. При усреднении по ансамблю функция корреляции нецентрированного случайного процесса определяется формулой (4.10), причем выборочные значения случайных реализаций берут в моменты времени t_1 и t_2 .

Если процесс эргодический, то усреднение по ансамблю заменяется усреднением по времени и функция корреляции определяется формулой (4.19). Для центрированных случайных процессов функция корреляции определяется формулой (4.11), а если процесс эргодический, то формулой

$$B_0(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} [(\xi(t) - a(t))(\xi(t + \tau) - a(t + \tau))] dt. \quad (4.24)$$

Таким образом, функция корреляции эргодического случайного процесса определяет степень статистической связи между реализацией случайного процесса и сдвинутой на время τ копией этой реализации.

Функции (4.19), (4.24), принято называть автокорреляционными функциями. Функция (4.20) называется *взаимной корреляционной функцией*. При $\xi(t) = \gamma(t)$ функция взаимной корреляции переходит в функцию автокорреляции.

Рассмотрим основные свойства автокорреляционной функции эргодического случайного процесса.

1. Автокорреляционная функция является убывающей функцией аргумента τ . Значение ее при $\tau = 0$ максимально и равно средней мощности случайного процесса:

$$B(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi^2(t) dt = P_{\text{ср}}.$$

Значение автокорреляционной функции централизованного случайного процесса равно дисперсии случайного процесса:

$$B_0(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\xi(t) - a(t)]^2 dt = P_{0\text{ср}} = D\{\xi(t)\}.$$

Если среднее значение случайного процесса является постоянной величиной $a(t) = a$, то значение автокорреляционной функции при $\tau = \infty$

$$B(\infty) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\xi(t) + a][\xi(\infty) + a] dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a^2 dt = a^2,$$

так как

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(\infty) dt = B_0(\infty) = 0.$$

Последняя формула означает, что чем дальше отстоят друг от друга выборочные значения случайного процесса, тем слабее статистическая зависимость между ними. При $\tau = \infty$ эта зависимость отсутствует.

2. Автокорреляционная функция является четной функцией от τ :

$$B(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t + \tau) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t - \tau) dt = B(-\tau).$$

3. В теории вероятности доказывается, что значение автокорреляционной функции максимально при сдвиге $\tau = 0$, т.е.

$$|B(0)| \geq B(\tau).$$

При описании случайных процессов используют понятия: *нормированная автокорреляционная функция и интервал корреляции*. Нормированная автокорреляционная функция случайного процесса определяет степень статистической связи между двумя

его сечениями (значениями случайной функции в моменты времени t_2 и t_1):

$$b(t_2, t_1) = B(t_2, t_1) / \sqrt{D\{\xi(t_1)\}D\{\xi(t_2)\}}.$$

Для стационарного процесса

$$b(\tau) = B(\tau) / D\{\xi(t)\}.$$

Интервал корреляции равен временному сдвигу между сечениями случайного процесса, при котором статистическая связь еще существует. Его определяют как основание прямоугольника с высотой, равной $B(0)$. Площадь прямоугольника равна площади, ограниченной кривой $B(\tau)$ и осями координат:

$$\tau_k = \int_0^{\infty} B(\tau) d\tau / B(0). \quad (4.25)$$

Понятием автокорреляционная функция сигнала пользуются не только применительно к случайным, но и к детерминированным функциям. Автокорреляционная функция детерминированного сигнала рассмотрена в разд. 2.4.

Во второй главе введено понятие спектральной плотности детерминированного сигнала, найдено выражение (2.20) для спектральной плотности мощности, а также получена формула (2.25), которая определяет автокорреляционную функцию $B(\tau)$ через спектральную плотность мощности. Указанные формулы справедливы и для стационарного случайного процесса. Для оценки ширины спектральной плотности мощности принято использовать величину

$$\Delta f_3 = \frac{1}{G(0)} \int_0^{\infty} G(\omega) \frac{d\omega}{2\pi} = \frac{1}{G(0)} \int_0^{\infty} G(f) df. \quad (4.26)$$

Принимая во внимание уравнение (2.26), найдем, что

$$G(0) = \int_{-\infty}^{\infty} B(\tau) d\tau = 2 \int_0^{\infty} B(\tau) d\tau.$$

Из последнего уравнения получаем

$$\int_0^{\infty} B(\tau) d\tau = \frac{1}{2} G(0). \quad (4.27)$$

Подставляя (4.27) в (4.25), получим

$$\tau_k = \frac{1}{2} \frac{G(0)}{B(0)}. \quad (4.28)$$

Принимая во внимание уравнение (2.25) найдем, что

$$B(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega) d\omega = \frac{1}{\pi} \int_0^{\infty} G(\omega) d\omega = 2 \int_0^{\infty} G(f) df = 2G(0)\Delta f_3.$$

Подставляя последнее равенство в (4.28), получим

$$\tau_k \Delta f_3 = 1/4.$$

Время корреляции тем меньше, чем больше ширина спектральной плотности и наоборот.

4.4 Нормальный случайный процесс

Нормальным или гауссовским случайным процессом называется такой процесс, n -мерная плотность распределения вероятности которого определяется формулой:

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = K_n \exp \left\{ - \sum_{i=1}^n \sum_{j=1}^n C_{ij} [(x_i - a_i)(x_j - a_j)] \right\}, \quad (4.29)$$

где K_n , C_{ij} , a_i , a_j — постоянные, связанные с моментами времени $t_1, t_2, \dots, t_n$ и между собой определенными зависимостями.

При $n = 1$ из (4.29) получаем выражение для одномерной плотности распределения вероятности:

$$w(x, t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-a)^2/2\sigma^2}, \quad (4.30)$$

где a — среднее значение, а σ^2 — дисперсия нормального процесса для момента времени t .

Для одномерной плотности коэффициент $K_1 = 1/\sqrt{2\pi\sigma}$, а коэффициент $C_{11} = 1/2\sigma^2$.

Нормальный случайный процесс занимает особое место при анализе систем связи. Он чаще, чем другие законы, встречается на практике, а его использование существенно упрощает расчеты.

Совместная плотность вероятности n независимых случайных величин определяется формулой (4.6).

Если одномерные плотности вероятности для каждого момента времени t_i , $i = 1, 2, \dots, n$ одинаковы и распределены по нормальному закону, причем $\sigma_1 = \sigma_2 = \dots = \sigma_n = \sigma$; $a_1 = a_2 = \dots = a_n = a$, то

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - a)^2}. \quad (4.31)$$

Используя формулу (4.31), вычислим n -мерную плотность нормального стационарного эргодического центрированного процесса, который имеет равномерную спектральную плотность в полосе частот от 0 до F_B .

Пусть некоторая реализация задана непрерывной функцией $x(t)$ в интервале времени 0- T , а ее спектр ограничен верхней граничной частотой F_B . Эту функцию с некоторой погрешностью

можно представить рядом (2.46), который в интервале $0-T$ совпадает с рядом Котельникова.

Энергия реализации

$$E = \int_0^T x^2(t) dt = \int_0^T \left[\sum_{k=1}^n x_k \frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)} \right]^2 dt =$$

$$= \sum_{k=1}^n x_k^2 \int_0^T \left[\frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)} \right]^2 dt + \sum_{k \neq l} \sum_{l=1}^n x_k x_l \int_0^T \frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)} \times$$

$$\times \frac{\sin 2\pi F_B(t - l\Delta t)}{2\pi F_B(t - l\Delta t)} dt = \frac{1}{2F_B} \sum_{k=1}^n x_k^2, \quad (4.32)$$

где $n = T/\Delta t = 2F_B T$ — число выборочных значений x_k интервале $0-T$, отсчитанных в дискретные моменты времени t_k , $k = 1, 2, \dots, n$ через интервал $\Delta t = 1/2F_B$.

При выводе формулы (4.32) принято во внимание, что базисные функции (2.44), из которых строится ряд (2.43), являются ортонормированными и удовлетворяют условию:

$$\int_{-\infty}^{\infty} \frac{\sin 2\pi F_B(t - k\Delta t)}{2\pi F_B(t - k\Delta t)} \cdot \frac{\sin 2\pi F_B(t - l\Delta t)}{2\pi F_B(t - l\Delta t)} dt = \begin{cases} 0, & k \neq l, \\ 1/2F_B, & k = l. \end{cases}$$

Полагая, что переменные x_k , входящие в выражение (4.31), являются выборочными значениями функции $x(t)$, а среднее значение стационарного, нормального, эргодического процесса $a = 0$, перепишем выражение (4.31) с учетом (4.32) в виде:

$$w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = w(\vec{x}) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{F_B}{\sigma^2} \int_0^T x^2(t) dt} =$$

$$= \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{N_0} \int_0^T x^2(t) dt}, \quad (4.33)$$

где $N_0 = \sigma^2 / F_B$ — односторонняя спектральная плотность эргодического нормального случайного процесса, равная отношению мощности этого процесса σ^2 к верхней граничной частоте F_B .

Формула (4.33) позволяет найти:

а) вероятность реализации $\xi(t)$ стационарного эргодического нормального случайного процесса:

$P[x_1 < \xi_1 \leq x_1 + dx_1, x_2 < \xi_2 \leq x_2 + dx_2, \dots, x_n < \xi_n \leq x_n + dx_n] = w(\vec{x}) d\vec{x}$,
где ξ_k , $k = 1, 2, \dots, n$ — выборки этой реализации; $d\vec{x} = dx_1 dx_2 \dots dx_n$

б) значение плотности распределения вероятности по заданной реализации $\xi(t)$:

$$w(\tilde{\xi}) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{N_0} \int_0^T \xi^2(t) dt}. \quad (4.34)$$

в) значение вероятности реализации $\xi(t)$:

$$w(\tilde{\xi}) d\tilde{\xi} = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{N_0} \int_0^T \xi^2(t) dt} d\tilde{\xi}.$$

Вероятность реализации тем больше, чем меньше энергия этой реализации

$$\int_0^T \xi^2(t) dt.$$

Найдем корреляционную функцию нормального стационарного эргодического центрированного процесса, полагая, что спектральная плотность мощности в полосе $-\omega_B - \omega_B$ постоянна и равна $N_0/2$.

$$B(\tau) = \frac{N_0}{2} \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} e^{j\omega\tau} d\omega = N_0 F_B \frac{\sin \omega_B \tau}{\omega_B \tau}. \quad (4.35)$$

При значении $\omega_B \tau = \pi$ функция $B(\tau) = 0$. Это соответствует разности времени $\tau = 1/2F_B$ между двумя соседними выборками. Следовательно, сечения процесса, разделенные этим интервалом, не коррелированы между собой. Это доказывает правомерность использования формулы (4.31) при выводе формул (4.33), (4.34).

Найдем предел, приняв во внимание формулу (2.17):

$$\lim_{\omega_B \rightarrow \infty} B(\tau) = \lim_{\omega_B \rightarrow \infty} \frac{N_0}{2} \cdot \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} e^{j\omega\tau} d\omega = \frac{N_0}{2} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} d\omega = \frac{N_0}{2} \delta(\tau). \quad (4.36)$$

При $\omega_B \rightarrow \infty$ корреляционная функция $B(\tau)$ стремится к δ -функции, для которой любые два несовпадающих сечения не коррелированы между собой.

Спектральная плотность такого процесса является постоянной величиной, равной

$$\int_{-\infty}^{\infty} B(\tau) e^{-j\omega\tau} d\tau = \frac{N_0}{2} \int_{-\infty}^{\infty} \delta(\tau) e^{-j\omega\tau} d\tau = \frac{N_0}{2}. \quad (4.37)$$

Эргодический центрированный стационарный нормальный процесс, корреляционная функция которого равна δ -функции, а спектральная плотность постоянна при всех частотах, называется *белым шумом*.

Белый шум является математической моделью реального эргодического нормального процесса, применимой в тех случаях, когда его время корреляции много меньше постоянной времени четырехполосника, на который действует этот процесс, или когда в пределах амплитудно-частотной характеристики четырехполосника спектральную плотность процесса можно приближенно считать постоянной.

Дисперсия белого шума

$$B(0) = \frac{N_0}{2} \delta(0) = \infty.$$

4.5. Узкополосные случайные процессы

Случайный процесс называется *узкополосным*, если ширина его спектра относительно мала по сравнению со средней частотой этого спектра. Подобные процессы возникают, например, на выходах линейных частотно-избирательных многокаскадных усилителей при действии на их входах теплового шума.

Аналитически узкополосный случайный процесс можно записать по аналогии с (2.32) в виде:

$$\begin{aligned} \xi(t) &= \eta(t) \cos[\omega_0 t + \Delta(t)] = \eta(t) \cos \Delta(t) \cos \omega_0 t - \eta(t) \sin \Delta(t) \sin \omega_0 t = \\ &= A(t) \cos \omega_0 t - B(t) \sin \omega_0 t = a(t) + b(t), \end{aligned}$$

где $a(t)$ — синфазная, $b(t)$ — квадратурная составляющие узкополосного случайного процесса; $\eta(t)$ — медленно меняющаяся амплитуда; $\Delta(t)$ — медленно меняющаяся фаза; ω_0 — средняя частота спектра; $A(t) = \eta(t) \cos \Delta(t)$; $B(t) = \eta(t) \sin \Delta(t)$ — случайные медленно меняющиеся функции времени.

Огибающая

$$\eta(t) = \sqrt{A^2(t) + B^2(t)}.$$

Фаза

$$\Delta(t) = -\arctg(B(t) / A(t)).$$

Синфазная $a(t)$ и квадратурная $b(t)$ составляющие имеют одинаковые мощности, их средние значения равны нулю, а функция взаимной корреляции

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a(t)b(t)dt = 0.$$

Если узкополосный случайный процесс подчиняется нормальному закону, то синфазная и квадратурные составляющие также будут подчинены нормальному закону.

Найдем плотность распределения вероятности огибающей $w(\eta)$ и фазы $\omega(\varphi)$ узкополосного случайного процесса $\xi(t)$. В пря-

моугольных координатах вероятность того, что случайная величина η будет находиться в пределах прямоугольника, ограниченного сторонами $(X, X + dX)$ и $(Y, Y + dY)$, равна $w(X, Y) dX dY$.

Вероятность этого же события с использованием полярных координат ϑ, φ

$$w(X, Y) dX, dY = w(\vartheta, \varphi) d\vartheta d\varphi, \quad (4.38)$$

где $w(\vartheta, \varphi) d\vartheta d\varphi$ — вероятность того, что случайная величина η будет находиться в пределах площади, равной $\vartheta d\vartheta d\varphi$ (рис. 4.6).

Учитывая, что синфазная и квадратурные составляющие независимы, найдем, что

$$\begin{aligned} w(X, Y) dX, dY &= \\ &= w(X) dX w(Y) dY. \end{aligned} \quad (4.39)$$

Кроме того, случайные величины $a(t)$, $b(t)$ распределены по нормальным законам с нулевыми средними значениями и одинаковыми дисперсиями, т.е.

$$\left. \begin{aligned} w(X) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-X^2/2\sigma^2}, \\ w(Y) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-Y^2/2\sigma^2}, \end{aligned} \right\} \quad (4.40)$$

где σ^2 — дисперсия синфазной и квадратурной составляющих.

Покажем, что дисперсии случайных величин $a(t)$ и $b(t)$ одинаковы и равны половине дисперсии узкополосного случайного процесса.

Вычислим дисперсию узкополосного случайного процесса $\xi(t)$:

$$\begin{aligned} \sigma_1^2 &= \overline{\xi^2(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi^2(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [A(t) \cos \omega_0 t - \\ &- B(t) \sin \omega_0 t]^2 dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A^2(t) \cos^2 \omega_0 t dt - \\ &- \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T A(t) B(t) \cos \omega_0 t \sin \omega_0 t dt + \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T B^2(t) \sin^2 \omega_0 t dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{A^2(t)}{2} dt + \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{B^2(t)}{2} dt = \frac{\sigma_1^2}{2} + \frac{\sigma_1^2}{2} = 2\sigma. \end{aligned}$$

При вычислении принято во внимание, что

$$\int_0^T \cos 2\omega_0 t dt \approx 0, \quad \int_0^T \sin 2\omega_0 t dt = 0.$$

Подставляя выражение (4.40) в (4.39), получим

$$w(X, Y) dX dY = \frac{1}{2\pi\sigma^2} e^{-\frac{X^2+Y^2}{2\sigma^2}} dX dY = \frac{\vartheta}{2\pi\sigma^2} e^{-\frac{\vartheta^2}{2\sigma^2}} d\vartheta d\varphi, \quad (4.41)$$

так как $\vartheta^2 = X^2 + Y^2$, а площади $dX dY$, $\vartheta d\vartheta d\varphi$ равны.

С учетом (4.38) из (4.41) находим совместную плотность распределения случайных величин η и Δ :

$$w(\vartheta, \varphi) = \frac{\vartheta}{2\pi\sigma^2} e^{-\vartheta^2/2\sigma^2}. \quad (4.42)$$

Плотность вероятности огибающей найдем, интегрируя (4.42) по φ :

$$w(\vartheta) = \int_0^{2\pi} w(\vartheta, \varphi) d\varphi = \int_0^{2\pi} \frac{\vartheta}{2\pi\sigma^2} e^{-\vartheta^2/2\sigma^2} d\varphi = \frac{\vartheta}{\sigma^2} e^{-\vartheta^2/2\sigma^2}. \quad (4.43)$$

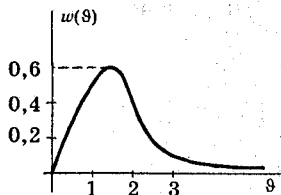


Рис.4.7. Плотность вероятности огибающей узкополосного случайного процесса

Эта плотность называется распределением Рэлея. График плотности показан на рис.4.7.

Плотность вероятности фазы найдем, интегрируя (4.42) по ϑ :

$$w(\varphi) = \int_0^{\infty} \frac{\vartheta}{2\pi\sigma^2} e^{-\vartheta^2/2\sigma^2} d\vartheta = \frac{1}{2\pi} \int_0^{\infty} w(\vartheta) d\vartheta = \frac{1}{2\pi},$$

так как $w(\vartheta) = \frac{\vartheta}{\sigma^2} e^{-\vartheta^2/2\sigma^2}$, а $\int_0^{\infty} w(\vartheta) d\vartheta = 1$.

График функций $w(\varphi)$ приведен на рис.4.8.

Аналогично вычисляются плотности вероятности для огибающей и фазы смеси узкополосного случайного шума и сигнала:

$$Z(t) = u(t) + \xi(t) = \eta_0 \cos \omega_0 t + \eta(t) \cos[\omega_0 t + \varphi(t)],$$

где η_0 — амплитуда сигнала. Плотность вероятности огибающей суммы сигнала и шума:

$$w(\vartheta) = \frac{\vartheta}{\sigma^2} e^{-\frac{\vartheta^2 + \eta_0^2}{2\sigma^2}} I_0\left(\frac{\vartheta\eta_0}{\sigma^2}\right), \quad (4.44)$$

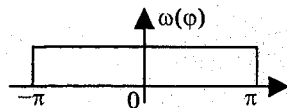


Рис.4.8. Плотность распределения вероятности фазы узкополосного случайного процесса

где $I_0(x)$ — модифицированная функция Бесселя нулевого порядка.

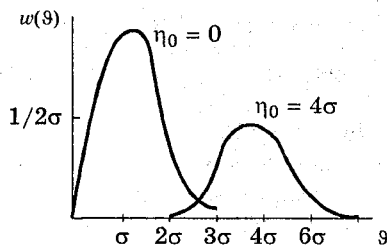


Рис.4.9. Плотность распределения вероятности огибающей суммы гармонического сигнала и узкополосного шума

При малых отношениях сигнал/шум плотность близка к релеевской, а при больших — к нормальной (рис.4.9). Выражение для плотности вероятности фазы

$$w(\varphi) = \frac{1}{\pi} e^{-\frac{\eta_0^2}{2\sigma^2}} + \frac{1}{2} \frac{\eta_0 \cos \varphi}{\sqrt{2\pi}\sigma^2} \times \\ \times \left[1 + \Phi\left(\frac{\eta_0 \cos \varphi}{2\sigma^2}\right) \right] e^{-\frac{\eta_0^2 \sin^2 \varphi}{2\sigma^2}}, \quad (4.45)$$

где $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$ — функция Крампа.

4.6. Геометрические представления случайных сигналов и помех

Геометрические представления детерминированных сигналов, рассмотренные в разд.2.6, используются и при анализе случайных сигналов и помех. Так совокупность выборочных значений реализации случайного процесса ξ_i ; $i = 1, 2, \dots, n$ может быть представлена n -мерным вектором $\bar{\xi}_i = \{\xi_1, \xi_2, \dots, \xi_n\}$, ансамблю поставлено в однозначное соответствие множество n -мерных векторов, которые будут занимать некоторую область в n -мерном пространстве.

Если задана n -мерная плотность распределения вероятности, то вероятность того, что множество n -мерных векторов находится в области V

$$P = \int_V w(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n,$$

где интегрирование ведется по этой области.

Можно найти также вероятность того, что вектор $\bar{\xi}$, определяющий заданную реализацию $\xi(t)$, попадет в бесконечно малую часть области dV :

$$dP = w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2 \dots dx_n. \quad (4.46)$$

Эта вероятность называется вероятностью появления заданной реализации.

Длина случайного n -мерного вектора

$$d = \|\bar{\xi}\| = \sqrt{\sum_{i=1}^n \xi_i^2}.$$

Квадрат длины случайного вектора с учетом (4.32)

$$d^2 = \sum_{i=1}^n \xi_i^2 = 2F_B E = 2F_B T P, \quad (4.47)$$

где P — мощность случайной реализации.

Рассмотрим две случайные реализации, с ограниченным спектром. Пусть значения верхних частот одинаковы и равны F_B . Вычислим расстояние между соответствующими этим реализациям n -мерными векторами:

$$d = \|\bar{\xi}_k - \bar{\xi}_m\| = \sqrt{\sum_{i=1}^n (\xi_{ik} - \xi_{im})^2}.$$

Квадрат расстояния

$$\begin{aligned} d^2 &= \sum_{i=1}^n (\xi_{ik} - \xi_{im})^2 = \sum_{i=1}^n \xi_{ik}^2 + \sum_{i=1}^n \xi_{im}^2 - 2 \sum_{i=1}^n \xi_{ik} \xi_{im} = \\ &= 2F_B T P_k + 2F_B T P_m - 2\bar{\xi}_k \bar{\xi}_m = 2F_B T (P_k + P_m) - 2\bar{\xi}_k \bar{\xi}_m, \end{aligned} \quad (4.48)$$

где $\bar{\xi}_k \bar{\xi}_m = \sum_{i=1}^n \xi_{ik} \xi_{im}$ — скалярное произведение векторов.

Расстояние между реализациями случайного процесса увеличивается с ростом F_B, T и мощности P случайных реализаций. Расстояние также зависит от скалярного произведения векторов. Если реализации $\xi_k(t)$, $\xi_m(t)$ равны и противоположны, то расстояние достигает максимальной величины.

Если реализации ортогональны, то $\bar{\xi}_k \bar{\xi}_m = 0$ и

$$d^2 = 2F_B T (P_k + P_m).$$

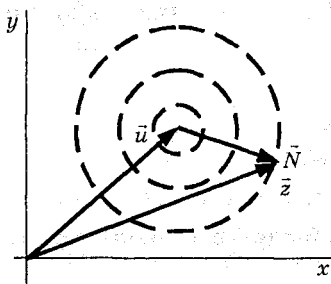


Рис.4.10. Векторная диаграмма суммы сигнала и помехи

Аддитивной смеси сигнала и шума $z(t) = u(t) + N(t)$, не содержащей спектральных компонент с частотой большей F_B , можно сопоставить сумму n -мерных векторов $\bar{u}$ и $\bar{N}$. При переходе от одной реализации $z(t)$ к другой вектор шума $\bar{N}$ занимает новое положение относительно вектора сигнала $\bar{u}$ (рис.4.10). При этом около конца вектора $\bar{u}$ можно указать область, за пределами которой вероятность нахождения вектора $\bar{N}$ будет

весьма малой. Для этой области радиус

$$d = \sqrt{2TF_{\text{ш}}P_{\text{ш max}}},$$

где $P_{\text{ш max}}$ — значение максимальной мощности реализации шума.

Для функционального пространства норма случайной реализации

$$\|\xi(t)\| = \sqrt{\int_{-T/2}^{T/2} \xi^2(t) dt},$$

где T — интервал времени, в пределах которого случайный процесс реализуется.

Расстояние между двумя реализациями $\xi_1(t)$ и $\xi_2(t)$ определяется нормой разности

$$\|\xi_1(t) - \xi_2(t)\| = \sqrt{\int_{-T/2}^{T/2} [\xi_1(t) - \xi_2(t)]^2 dt}.$$

Квадрат расстояния

$$\begin{aligned} d^2 &= \int_{-T/2}^{T/2} [\xi_1(t) - \xi_2(t)]^2 dt = \int_{-T/2}^{T/2} \xi_1^2(t) dt - 2 \int_{-T/2}^{T/2} \xi_1(t)\xi_2(t) dt + \\ &+ \int_{-T/2}^{T/2} \xi_2^2(t) dt = E_1 + E_2 - 2 \int_{-T/2}^{T/2} \xi_1(t)\xi_2(t) dt, \end{aligned}$$

где E_1, E_2 — энергия реализаций; $\int_{-T/2}^{T/2} \xi_1(t)\xi_2(t) dt$ — взаимнокорреляционная функция.

Если взаимнокорреляционная функция отрицательна, то квадрат расстояния достигает максимума, а если положительна, то — минимума. В случае, если взаимнокорреляционная функция равна нулю, то

$$d^2 = E_1 + E_2 = P_1 T + P_2 T,$$

где P_1, P_2 — средние мощности реализаций.

4.7. Математические модели сообщений

Экспериментально установлено, что основная энергия разговорного телефонного сигнала сосредоточена в полосе частот от 100 Гц до 5 кГц и зависит от характеристик человеческого голоса и системы передачи. С уменьшением полосы системы передачи экономические показатели улучшаются, но падает разборчивость речи.

Оптимальная полоса частот, при которой обеспечиваются приемлемые экономические показатели и достаточная разборчивость речи, лежит в пределах от 300 Гц до 3400 Гц. Динамический диапазон речи диктора составляет 25-35 дБ.

Для передачи спектра телевизионного сигнала требуется во много раз большая полоса. Она определяется общим числом различных элементов в одном кадре изображения, которое равно $625 \cdot 833 = 520000$, так как число различных элементов вдоль строки равно $4/3 \cdot 625 = 833$, где $4/3$ — отношение сторон кадра. Поскольку частота кадров выбирается равной 25 Гц, на передачу всех различных элементов требуется $1/25$ с, а на передачу одной пары элементов — $1/260000 \cdot 25$. Последняя цифра определяет верхнюю частоту телевизионного сигнала, которая равна 6,5 МГц. Динамический диапазон телевизионного сигнала составляет 40 дБ. Первичные сигналы, в том числе речевые и видеосигналы, являются нестационарными случайными процессами с изменяющимися во времени временными и спектральными характеристиками. Поэтому точное математическое описание таких сигналов, как правило, связано с затруднениями. Для облегчения задачи анализа обычно используют различные приближения.

Например, речевой сигнал представляют в виде стационарного нормального случайного процесса с корреляционной функцией

$$B(\tau) = e^{-\alpha|\tau|} \cos 2\pi f_0 \tau,$$

где $\alpha \approx 10^3$ Гц, $f_0 = 400$ Гц.

Сигналы черно-белого изображения также описываются стационарным нормальным случайным процессом $\xi(t_1, t_2)$ с корреляционной функцией

$$B(\tau_1, \tau_2) = \exp(-\alpha|\tau_1| - \beta|\tau_2|),$$

где τ_1, τ_2 — пространственные координаты изображения; α, β — постоянные коэффициенты.

В факсимильной связи первичный сигнал формируется методом развертки, как и в телевидении. Поэтому между сигналами в телевидении и факсимильной связи много общих черт. В то же время между ними имеются и существенные различия. Ширина спектра факсимильного сигнала зависит от скорости развертки изображения и размеров светового пятна. Она во много раз меньше ширины спектра телевизионного сигнала и составляет при передаче газет примерно 180 кГц. Изображения бывают штриховые, содержащие всего две градации яркости, и полутоновые. Для полутонового изображения динамический диапазон составляет примерно 25 дБ. При передаче изображения на выходе фотоэлектрического преобразователя получается сигнал сложной фор-

мы, энергетический спектр которого содержит все частоты от нуля до максимальной.

Поскольку передача изображения ведется построчно, то сигналы каждой строки отличаются друг от друга и являются реализациями некоторого случайного процесса.

Простейшей математической моделью первичного сигнала двухградационного изображения является случайная последовательность независимых друг от друга единиц и нулей.

Математическая модель оказывается полностью заданной, если известны вероятности $P(0)$ и $P(1)$. Эти вероятности могут быть найдены экспериментально путём подсчета числа белых и черных элементов изображения и деления полученных чисел на общее число элементов в строке. Реальные изображения отличаются статистической зависимостью между элементами изображения. В этом случае более адекватной математической моделью является марковская случайная последовательность. Как было указано ранее (разд.4.2), для полного описания марковской случайной последовательности необходимо задать априорные вероятности единиц и нулей и переходные вероятности.

Первичные сигналы телеграфии при передаче данных обычно имеют вид последовательностей импульсов прямоугольной формы: 0 или $v(t)$; $-v(t)$ или $+v(t)$. Как показано в третьей главе, ширина спектра первичного сигнала определяется шириной спектра импульса $v(t)$ и для прямоугольного импульса равна $1/\tau_{\text{и}}$ (Гц), где $\tau_{\text{и}}$ — длительность прямоугольного импульса.

Для стандартной скорости телеграфирования 50 Бод ширина спектра телеграфного сигнала равна 50 Гц. При скорости 2400 Бод ширина спектра равна соответственно 2400 Гц. При обмене данных между ЭВМ ширина спектра первичного сигнала определяется типом используемой сети. Для низкоскоростной сети ширина спектра равна 200 Гц, для среднескоростной от 600 до 9600 Гц и для высокоскоростной более 9600 Гц.

Источниками первичных сигналов в системах передачи данных могут быть оконечные устройства, управляемые непосредственно человеком: телегайды, графические устройства. Информация в эти устройства вводится в виде линий или графиков, с помощью, например, светового пера. Более сложными источниками являются электромеханические устройства считывания с магнитных лент и дисков. Наконец, в качестве источников первичных сигналов могут использоваться датчики системы комплексной автоматизации.

Часто сообщения от недалеко удаленных друг от друга оконечных устройств объединяются в пакеты.

Рассмотрим математические модели при передаче сообщений в виде пакетов. Сообщения, передаваемые в системах переда-

чи данных, необходимо трактовать как неизвестные величины. Простейшей вероятностной моделью сообщений, передаваемых оконечным устройством, является нестационарный случайный процесс со сложными вероятностными связями между пакетами и паузами.

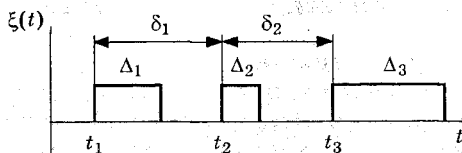


Рис.4.11. Реализация процесса, используемого в качестве модели потока сообщений

Типичная реализация $\xi(t)$ процесса, описывающего поток сообщений, имеет вид последовательности пакетов (рис.4.11). Последовательность задается длинами Δ_i и моментами появления отдельных пакетов t_i .

При математическом описании реализации $\xi(t)$ накладываются следующие условия: 1) процесс, описывающий моменты появления пакетов, стационарный; 2) распределение длины данного пакета не зависит от длины других пакетов; 3) последовательность моментов появления пакетов и последовательность длин пакетов независимы.

Исследования свойств моментов возникновения пакетов показали, что длина пакета имеет следующее распределение [13]:

$$P(n = k) = \mu(1 - \mu)^{k-1},$$

где μ — вероятность того, что появившейся пакет закончится в данном единичном интервале, причем среднее значение $\bar{n} = 1/\mu$; $k = 1, 2, \dots$

Длина пакета $n = \Delta/\tau_0$, где τ_0 — продолжительность единичного элемента. Минимальный интервал времени τ_0 , который требуется для передачи одного элементарного сообщения (при передаче двоичных цифр это двоичные числа 0 и 1) называется единичным интервалом.

Элемент сигнала, имеющий длительность, равную единичному интервалу, называется единичным. Если длительность единичного элемента $\tau_0 \rightarrow 0$, распределение длины пакета непрерывно и подчинено экспоненциальному закону:

$$P(\Delta) = \begin{cases} \mu e^{-\mu\Delta}, & \Delta \geq 0; \\ 0, & \Delta < 0, \end{cases}$$

а число пакетов, появляющихся в промежутке времени $[t, t + T]$ подчиняется распределению Пуассона:

$$P(n_x = k) = \frac{(\lambda T)^k}{k!} e^{-\lambda T},$$

где n_k — число пакетов, λ — вероятность возникновения пакета в указанном промежутке времени.

Распределение промежутка между началами соседних пакетов подчиняется экспоненциальному закону:

$$P(\delta) = \begin{cases} \lambda e^{-\lambda\delta} & \text{при } \delta \geq 0, \\ 0 & \text{при } \delta < 0. \end{cases}$$

Контрольные вопросы

1. В чем отличие описаний случайного и детерминированного процессов?
2. Какие характеристики случайных процессов Вам известны? Как связаны между собой функция распределения и плотность распределения вероятности? Перечислите свойства плотности распределения вероятности.
3. При выполнении какого условия исчерпывающей характеристикой случайной функции является ее одномерная плотность вероятности?
4. Какие числовые характеристики распределения вероятности Вам известны?
5. Что такое центрированный случайный процесс?
6. Как определяются математическое ожидание, дисперсия и корреляционная функция?
7. Какие случайные процессы называются стационарными, (нестационарными); эргодическими?
8. Как определяется математическое ожидание, дисперсия и корреляционная функция эргодического случайного процесса?
9. Какие модели дискретных сообщений Вам известны?
10. Какой случайный процесс называется марковским?
11. Какими основными свойствами обладает корреляционная функция эргодического случайного процесса?
12. Какая связь существует между шириной спектральной плотности мощности и интервалом корреляции случайного процесса?
13. Какой случайный процесс называется нормальным?
14. Как аналитически выражается n -мерная плотность распределения нормального белого шума?
15. По каким законам распределены огибающая и фаза узкополосного нормального случайного процесса?

Глава 5. КАНАЛЫ СВЯЗИ

5.1. Общие сведения о каналах передачи информации

5.1.1. Классификация каналов

Классификацию каналов ведут по различным показателям: назначению (*телеграфные, телефонные, телевизионные* и т.д.), виду физической среды распространения (*проводные, волноводные, радио* и др.), диапазону используемых частот, способу распространения электромагнитных колебаний (с *открытым* или *закрытым* распространением), характеру информационных параметров сигналов на входе и выходе канала (*непрерывные, дискретные, дискретно-непрерывные* и т.д.).

В каналах с закрытым распространением электромагнитная волна распространяется по направляющим линиям: *кабельным, волноводным, оптоволоконным*.

В каналах с открытым распространением электромагнитная волна распространяется в *космическом, околоземном* или *водном* пространстве.

К радиодиапазону относят частоты в пределах от 30 до $30 \cdot 10^{12}$ Гц. Данные о радиодиапазонах приведены в табл.5.1.

Таблица 5.1

Диапазон частот	Диапазон волн	Название частот	Название волн
30..300 Гц	1000..10000 км	Сверхнизкие (СНЧ)	Мириаметровые Километровые Гектометровые Декаметровые Метровые Дециметровые Сантиметровые Миллиметровые Децимиллиметровые
300..3000 Гц	100..1000 км	Инфранизкие (ИНЧ)	
3..30 кГц	10..100 км	Очень низкие (ОНЧ)	
30..300 кГц	1..10 км	Низкие (НЧ)	
300..3000 кГц	100..1000 м	Средние (СЧ)	
3..30 МГц	10..100 м	Высокие (ВЧ)	
30..300 МГц	1..10 м	Очень высокие (ОВЧ)	
300..3000 МГц	10..100 см	Ультравысокие (УВЧ)	
3..30 ГГц	1..10 см	Сверхвысокие (СВЧ)	
30..300 ГГц	1..10 мм	Крайне высокие (КВЧ)	
300..3000 ГГц	0,1..1 мм	Гипервысокие (ГВЧ)	

К оптическому относят диапазон частот от 10^{13} до 10^{15} Гц.

5.1.2. Краткие сведения о искажениях сигналов

При передаче сигналов по каналу связи их параметры изменяются. Различают *детерминированные* и *случайные* изменения этих параметров. Детерминированные сводятся к изменению масштаба сигнала (усилению или ослаблению) и к задержке его

во времени. Изменение масштаба можно скомпенсировать усилением (ослаблением) сигнала. Задержку скомпенсировать нельзя, что сказывается на своевременности доставки сообщения (особенное сильно в протяженных каналах).

При передаче сигнала изменяется и его форма. Искажения формы могут быть детерминированными и случайными, линейными и нелинейными. Они связаны с нарушением соотношений между амплитудами, а также фазами гармонических составляющих исходного сигнала при его передаче. Для того, чтобы детерминированные линейные искажения отсутствовали, нужно, чтобы модуль коэффициента передачи линейного четырехполюсника (АЧХ) равнялся постоянной величине K , а время запаздывания было бы постоянным для всех гармонических составляющих сигнала.

Неравномерность АЧХ, ФЧХ и ограниченность ширины полосы канала приводят к искажению формы передаваемых сигналов. Например, единичные импульсы кодовой последовательности "расплываются" во времени. При этом на выходе канала одновременно присутствуют отклики от нескольких единичных импульсов: от принимаемого в данном единичном интервале и от принятых в предыдущие единичные интервалы. Подобное наложение существенно ограничивает скорость передачи дискретной информации.

Случайные линейные искажения сигнала возникают при случайном изменении параметров среды распространения: свободного пространства, кабеля или оптического волновода. Рассмотрим искажения, которые возникают в системах радиосвязи. Они во многом определяются характером распространения радиоволн. В околосемном пространстве излученный антенной передатчика радиосигнал, отражаясь от ионизированных слоев, приходит к антенне приемника разными путями. Каждому пути соответствует свой сигнал с определенными амплитудой и задержкой. Если бы ионизированные слои были стационарными, то и параметры этих сигналов (амплитуда и задержка) были бы постоянными. Однако, из-за нестационарности ионизированных слоев, эти параметры флуктуируют.

Пусть излучаемый передающей антенной сигнал $u(t) = \text{Re} e^{j\omega t}$. На выходе радиоканала суммарный сигнал

$$u_{\text{вых}}(t) = \text{Re} \sum_{i=1}^L K_i e^{j\omega(t-\tau_i)} = \text{Re} \dot{K} e^{j\omega t}, \quad (5.1)$$

где L — число путей, K_i — коэффициент ослабления амплитуды сигнала; τ_i — в i -м пути распространения;

$$\dot{K} = \sum_{i=1}^L K_i e^{-j\omega\tau_i}. \quad (5.2)$$

Передаточная функция канала (5.2), равная отношению выходного сигнала к входному, является случайной величиной из-за случайного изменения параметров K_i и τ_i . Кроме того, она зависит от частоты ω .

Перепишем (5.2) в виде

$$\dot{K} = \sum_{i=1}^L K_i e^{-j\omega\tau_i} = \sum_{i=1}^L K_i \cos \omega\tau_i - j \sum_{i=1}^L K_i \sin \omega\tau_i = X(\omega) + jY(\omega), \quad (5.3)$$

где $X(\omega) = \sum_{i=1}^L K_i \cos \omega\tau_i$, $Y(\omega) = -\sum_{i=1}^L K_i \sin \omega\tau_i$.

Для большого числа путей L $X(\omega)$ и $Y(\omega)$ являются нормально распределенными случайными величинами. Они имеют одинаковые дисперсии σ^2 , а их математические ожидания $m_x = m_y = 0$. Можно показать, что одномерное распределение модуля $|\dot{K}|$ подчинено *рэлеевскому* закону:

$$\left. \begin{aligned} w(k) &= \frac{k}{\sigma^2} e^{-k^2/2\sigma^2}, & k \geq 0, \\ w(k) &= 0, & k < 0. \end{aligned} \right\} \quad (5.5)$$

Фаза результирующего сигнала (5.1) распределена на интервале $(0 - 2\pi)$ по равномерному закону. Изменение его амплитуды на входе приемника (на выходе линии связи), обусловленное условиями распространения радиоволн, называется *замиранием сигнала*. Замирание, подчиняющееся закону (5.5), называется *рэлеевским*.

Если в месте приема присутствуют диффузионно отраженные подлучи и регулярный (незамирающий) луч с коэффициентом k_i , который значительно больше других, то коэффициент передачи канала подчиняется *обобщенному рэлеевскому закону*:

$$w(k) = \frac{k}{\sigma^2} \exp\left(-\frac{k^2}{2\sigma^2} - q^2\right) I_0\left(\frac{k\sqrt{2}q}{\sigma}\right), \quad k \geq 0, \quad (5.6)$$

где q^2 — отношение мощностей регулярной и флуктуирующих составляющих, $I_0(x)$ — модифицированная функция Бесселя нулевого порядка.

Пусть по каналу с замираниями сигнала передается узкополосный сигнал, причем среднеквадратичное отклонение запаздывания в каждом отдельном луче $\Delta\tau$ удовлетворяет условию

$$\Delta\tau \ll 1/F,$$

где F — ширина спектра сигнала.

Если изменения начальных фаз спектральных компонент сигнала на различных частотах почти одинаковы, то все состав-

ляющие спектра замирают "дружно", что свидетельствует об одинаковом изменении их амплитуд и фаз. Такие замирания называются *общими или гладкими*.

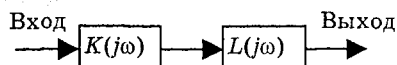


Рис.5.1. Корректор линейных детерминированных искажений

Если условие $\Delta t \ll 1/F$ не выполняется, то замирания разных спектральных компонент не одинаковы. Они называются *селективными*.

Детерминированные линейные искажения можно скомпенсировать. Одним из способов компенсации является включение в канал корректирующего четырехполосника (рис.5.1). Четырехполосник подбирается так, чтобы выполнялось равенство

$$K(j\omega)L(j\omega) = Ke^{-j\omega t_3}, \quad (5.7)$$

где K — постоянное число, t_3 —задержка сигнала; $L(j\omega)$ — коэффициент передачи корректирующего четырехполосника.

На практике полностью выполнить условие (5.7) не удастся. Поэтому коррекцию делают так, чтобы искажения не выходили за установленные пределы. Могут указываться допустимые отклонения отклика от прямоугольного импульсного воздействия или допустимые отклонения АЧХ и ФЧХ реальных каналов от заданных. Случайные линейные искажения скомпенсировать нельзя.

Нелинейные искажения возникают в выходных цепях передатчиков, входных цепях приемников, усилителях промежуточной частоты и других устройствах, входящих в аппаратуру канала связи, из-за нелинейности их амплитудных характеристик.

Анализ нелинейных искажений ведут, используя представление нелинейного четырехполосника в виде безынерционной и инерционной частей. При нелинейном безынерционном преобразовании выходной спектр сигнала обогащается новыми спектральными составляющими. Часть из них вместе с полезными спектральными составляющими попадает в полосу пропускания линейной части нелинейного четырехполосника и появляется на его выходе.

Для оценки нелинейных искажений используется коэффициент нелинейных искажений, определяемый как отношение корня квадратного из суммы квадратов амплитуд всех паразитных продуктов преобразования на выходе нелинейного четырехполосника к амплитуде полезной составляющей. Частным случаем коэффициента нелинейных искажений является *коэффициент гармоник*. Он представляет отношение корня квадратного из суммы квадратов амплитуд всех гармоник сигнала, за исключением

полезной гармоникой, к амплитуде полезной составляющей. Если полезной является первая гармоника, коэффициент

$$K_r = \frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + \dots}}{A_1}. \quad (5.8)$$

5.1.3. Помехи в каналах связи

Помехи, как указывалось выше, разрушают форму сигнала. Это приводит к необратимым потерям передаваемой информации. В зависимости от места нахождения источника различают *внутренние* и *внешние* помехи.

К внутренним помехам, возникающим в самом канале, относятся: *флуктуационные* шумы входных каскадов приемника, линий передачи сигналов от передатчика к приемнику и т.д.; электрические сигналы, попадающие в приемник из-за плохой экранировки между каскадами и блоками аппаратуры. Флуктуационные шумы содержат тепловую и дробовую составляющие. *Тепловой шум* возникает из-за теплового движения носителей заряда. *Дробовой шум* обусловлен дискретной природой носителей заряда и возникает на выходе электровакуумных и полупроводниковых приборов. Тепловые и дробовые шумы распределяются по нормальному закону с нулевым средним значением.

В диапазоне оптических частот к тепловым и дробовым шумам добавляется "*квантовый*", вызванный дискретной природой преобразования оптического сигнала в электрический ток.

Внешние помехи возникают вне канала связи. К ним относятся: *атмосферные, космические, промышленные, помехи от посторонних средств связи*. Атмосферные помехи возникают в результате физических процессов, протекающих в земной атмосфере. Мощным их источником являются грозовые разряды, излучение которых сказывается практически во всем радиодиапазоне.

Космические помехи оказывают влияние на системы радиосвязи, работающие на частотах более 30 МГц. Источниками этих помех являются Солнце, радиозвезды, межзвездные газы.

Промышленные помехи создаются различными промышленными установками, работой электрического транспорта, линиями электропередач. Помехи от посторонних средств связи, радиолокационных и навигационных систем могут превышать другие помехи. Увеличению их уровня способствует нарушение регламента распределения рабочих частот, недостаточность стабильности генераторов, плохая фильтрация побочных гармоник излучаемых сигналов.

Помехи делят на *аддитивные* и *мультипликативные*. Воздействие помех на сигнал можно представить в виде:

$$u^*(t) = u(t)\xi_m(t) + \xi_a(t),$$

где $u(t)$ — неискаженный помехами сигнал, $\xi_m(t)$ — мультипликативная помеха, $\xi_a(t)$ — аддитивная помеха $u^*(t)$ — искаженный помехами сигнал.

Причиной возникновения мультипликативной помехи является непрерывное или дискретное изменение параметров тракта передачи.

В зависимости от спектральных и временных характеристик аддитивную помеху делят на сосредоточенную по спектру, импульсную и флуктуационную. К сосредоточенным по спектру помехам относят сигналы посторонних радиостанций, а также излучения генераторов высокой частоты, используемых в промышленности, на транспорте, в медицинских учреждениях. Ширина спектральной плотности этих помех, как правило, намного меньше ширины полосы пропускания системы связи, но интенсивность может быть весьма большой. В диапазоне коротких волн сосредоточенные помехи являются основными. К импульсным помехам относят одиночные импульсы или импульсы, следующие с большим периодом повторения. Источником импульсных помех являются: грозовые разряды, электротранспорт, электрические промышленные установки.

5.2. Математические модели каналов связи

5.2.1. Модель непрерывного канала без шума

Математическая модель канала связи, необходимая для проведения исследований, считается заданной, если известны оператор преобразования, а также условия и ограничения, накладываемые как на канал, так и на входные и выходные сигналы. Математическая модель реального канала связи является достаточно сложной. Это объясняется следующими причинами.

1. Оператор L , кроме линейных, содержит также нелинейные и параметрические преобразования.
2. В канале присутствуют помехи.
3. Входной сигнал может быть случайным.

Часто сложная математическая модель не позволяет найти решение поставленной задачи. Поэтому пользуются упрощенными моделями. В них используют представление канала в виде последовательно соединенных четырехполюсников (линейных, нелинейных, параметрических). Полезным является также выделение из канала его дискретной, непрерывной и дискретно-непрерывных частей.

Результаты анализа во многом зависят от условий и ограничений, которые накладываются на сигналы и помехи. Это касается законов их распределения и физических характеристик сигналов, таких, как длительность, ширина спектра, мощность.

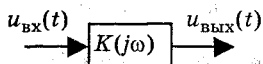


Рис.5.2. Простая модель непрерывного канала

Например, при анализе кабельных каналов используют линейный четырехполюсник с постоянным коэффициентом передачи (рис.5.2). Сигналы на его входе и выходе являются непрерывными по уровню и времени. Спектр входного сигнала лежит в заданной полосе частот. Мощность сигнала ограничена. Помеха на входе канала отсутствует. Модель имеет существенный недостаток, так как не учитывает присутствие помех, которые всегда имеют место в реальных каналах.

5.2.2. Модель непрерывного канала с аддитивным гауссовым шумом

Математическая модель, более точно описывающая реальные кабельные каналы, учитывает присутствие в них помех. Она включает: представление канала в виде линейного четырехполюсника с постоянными параметрами и заданной шириной полосы; описание сигналов на входе и выходе канала; описание помех.

Если входной сигнал является случайным, то задают не только его физические характеристики, но и n -мерную плотность распределения вероятности $w(\vec{x})$. То же относится и к помехе на выходе канала. Для нее задают физические характеристики (например, среднюю мощность) и n -мерную плотность распределения вероятности $w(\vec{r})$. Наиболее часто встречающаяся помеха аппроксимируется эргодическим нормальным случайным процессом с нулевым средним значением и n -мерной плотностью вида (4.34). Как правило эта помеха является аддитивной. Поэтому смесь сигнала и помехи (шума) записывают в виде

$$z(t) = ku(t - \tau) + \xi(t),$$

где k — постоянный коэффициент, определяющий масштаб (усиление или ослабление) сигнала; τ — постоянная задержка сигнала.

Смесь определяется условной n -мерной плотностью распределения вероятности

$$w(\vec{y} / \vec{u}) = w(y_1, y_2, \dots, y_n / u_1, u_2, \dots, u_n) = \frac{\partial^n F(\vec{y} / \vec{u})}{\partial \vec{y}},$$

где $F(\bar{y}/\bar{u}) = P[z_1 \leq y_1, z_2 \leq y_2, \dots, z_n \leq y_n / u_1, u_2, \dots, u_n]$ — вероятность того, что для заданной реализации сигнала $u(t)$ (следовательно, и заданном векторе $\bar{u}$) выборочные значения смеси $z(t)$ меньше или равны значениям неслучайных аргументов y_k , $k = 1, 2, \dots, n$. Учитывая, что $y_k = x_k + r_k$ и полагая, что $u_k = x_k$, найдем, что

$$\begin{aligned} P[z_1 \leq x_1 + r_1, z_2 \leq x_2 + r_2, \dots, z_n \leq x_n + r_n] = \\ = P[z_1 - u_1 \leq r_1, z_2 - u_2 \leq r_2, \dots, z_n - u_n \leq r_n]. \end{aligned}$$

Следовательно, $w(\bar{y}/\bar{u}) = w(\bar{r})$, причем эту условную плотность можно найти по формуле (4.34) с помощью подстановки:

$$\xi(t) = z(t) - ku(t - \tau).$$

Условная плотность

$$w(\bar{y}/\bar{u}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{N_0} \int_0^T [z(t) - ku(t - \tau)]^2 dt} \quad (5.10)$$

где N_0 — мощность шума, приходящаяся на единицу полосы, равная отношению этой мощности σ^2 к верхней граничной частоте F_B .

Полученное выражение называется *функцией правдоподобия*. Она позволяет определить вероятность реализации смеси сигнала и шума на выходе канала при условии присутствия на входе заданной реализации сигнала.

5.2.3. Канал с неопределенной фазой сигнала

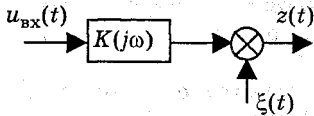


Рис.5.3. Модель непрерывного канала с аддитивным гауссовым шумом

Данная математическая модель используется для описания каналов связи, с флуктуациями сред, по которым распространяются сигналы. Это радиоканалы для связи в пределах прямой видимости, проводные каналы и др. Флуктуации среды случайным образом изменяют задержку сигнала, что приводит к случайному изменению его фазы.

Структура канала определяется линейным четырехполюсником (рис.5.3).

Входной сигнал

$$u_{\text{вх}}(t) = \eta \cos \omega_0 t. \quad (5.11)$$

Его амплитуда η меняется случайным образом в соответствии с сообщением.

Для дискретного источника значения амплитуды характеризуются заданными вероятностями, а для непрерывного — плотностью распределения вероятности.

Сигнал на выходе канала

$$u_{\text{вых}}(t) = K\eta \cos[\omega_0(t - \tau) + \varphi_0] = A \cos(\omega_0 t + \varphi), \quad (5.12)$$

где амплитуда A сигнала равна ослабленной (усиленной) в k раз амплитуде η входного сигнала; φ — начальная фаза сигнала.

При прохождении по каналу связи сигнал приобретает случайный набег фазы из-за флуктуации длины пути распространения и нестабильностей частот опорных генераторов, которые входят в состав аппаратуры канала связи и обеспечивают ее работоспособность.

Смесь сигнала и шума на выходе канала

$$z(t) = A \cos(\omega_0 t + \varphi) + \xi(t), \quad (5.13)$$

где амплитуда A повторяет изменения амплитуды η входного сигнала; фаза φ — случайна и имеет равномерный закон распределения в интервале $0-2\pi$ (см. рис.4.8); шум $\xi(t)$ является белым нормальным шумом с равномерной спектральной плотностью N_0 и нулевым средним значением.

Функция правдоподобия (условная плотность распределения вероятности) определяется выражением:

$$w(\bar{y} / A, y) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{N_0} \int_0^T [z(t) - A \cos(\omega_0 t + \varphi)]^2 dt} \quad (5.14)$$

Выражение (5.14), в отличие от (5.10), зависит от дополнительных случайных параметров: фазы φ и амплитуды A сигнала. Поскольку в фазе в данном случае не содержится передаваемой информации (она содержится в амплитуде), ее относят к паразитным случайным параметрам.

Математическая модель гауссовского канала с общими замираниями описывается аналогично, но множитель k , входящий в (5.12), как и фаза, считается паразитным случайным параметром.

Как показано в разд.5.1 одномерное распределение огибающей может быть рэлеевским (5.5). В этом случае канал называется каналом с рэлеевскими замираниями.

5.2.4. Канал с межсимвольной интерференцией и аддитивным шумом

Межсимвольная интерференция возникает при передаче двоичных кодовых последовательностей, если скорость передачи больше ширины полосы пропускания канала.

В радиоканалах причиной межсимвольной интерференции является многолучевое распространение радиоволн. При этом к аддитивному шуму прибавляется специфическая помеха, равная сумме откликов канала на импульсы кодовой последовательности, передававшиеся ранее. Математическая модель канала описывает входной сигнал, структуру канала, определяемую коэффициентом передачи $K(j\omega)$, выходной сигнал, специфическую помеху и флуктуационный шум.

Более простой является модель, для которой параметры коэффициента передачи не изменяются во времени, а сигналы на входе и выходе связаны детерминированными соотношениями.

5.2.5. Математические модели дискретных каналов связи

Сигналы на входе и выходе дискретного канала являются дискретными.

Математическая модель этого канала определяется:

а) алфавитом кодовых символов на входе $\{a_i\}$, $i = 1, 2, \dots, m$ и выходе $\{b_j\}$, $j = 1, 2, \dots, n$;

б) априорными вероятностями появления символов на входе канала;

в) вероятностями перехода $P(b_j/a_i)$, которые определяют вероятность того, что при передаче символа a_i на выходе канала появится символ b_j .

Если вероятность $P(b_j/a_i)$ не зависит от времени, то такой канал называется *однородным*.

В симметричном однородном канале без памяти алфавит кодовых символов на входе совпадает с алфавитом на выходе, а вероятности перехода определяются равенствами

$$P(b_j/a_i) = \begin{cases} p/m - 1 & \text{при } i \neq j, \\ 1 - p & \text{при } i = j. \end{cases} \quad (5.15)$$

Любой символ a_i , ($i = 1, 2, \dots, m$) может перейти в другой символ b_j ($j = 1, 2, \dots, m$) с равной вероятностью $p/m - 1$. Эти переходы определяют вероятность ошибки, равную p . Кроме того, любой символ a_i может с вероятностью $(1 - p)$ перейти в символ b_i , т.е. принят правильно.

Для двоичного симметричного канала без памяти $m = 2$ и

$$P(b_j/a_i) = \begin{cases} p & \text{при } i \neq j, \\ 1 - p & \text{при } i = j. \end{cases} \quad (5.16)$$

Вероятности перехода (5.16) схематично показаны на рис. 5.4. Вероятность перехода нуля в ноль равна $1 - p$, а нуля в единицу равна p . Соответственно, вероятность перехода 1 в 1 равна $1 - p$, а вероятность перехода 1 в ноль равна p .

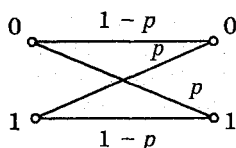


Рис.5.4.
Переходные
вероятности
в двоичном
симметричном
канале

Ошибка называется одно-, двух- и т.д., q -кратной, если из n передаваемых символов q переданы с ошибкой. Вектором ошибки в двоичном канале $\{\bar{e}\}$ является последовательность нулей и единиц, каждый элемент которой определен суммой по модулю 2 элементов входной и выходной последовательностей, находящихся на одних и тех же позициях, т.е.

$$\{e_i\} = \{a_i + b_i\}, \quad i = 1, 2, \dots, n.$$

где $+$ — знак суммирования по модулю 2.

Очевидно, что q -кратной ошибке в двоичном симметричном канале будет соответствовать такой вектор ошибки, у которого на любых позициях число единиц равно q . На остальных $n - q$ позициях этого вектора элементы равны нулю. В симметричном канале без памяти статистическая зависимость между передаваемыми символами отсутствует. Поэтому вероятность того, что произойдет q ошибок, расположенных как угодно на протяжении последовательности длиной n , равна:

$$P(q) = C_n^q \left(\frac{p}{m-1} \right)^q (1-p)^{n-q}, \quad (5.17)$$

где $C_n^q = \frac{n!}{q!(n-q)!}$ — биномиальный коэффициент, равный числу различных сочетаний q ошибок в последовательности длиной n .

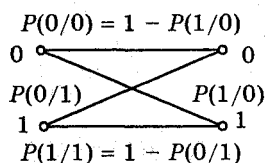


Рис.5.5. **Переходные**
вероятности
в двоичном
несимметричном
канале

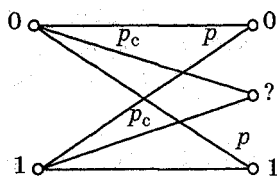


Рис.5.6. **Переходные**
вероятности
в симметричном
однородном канале без
памяти и со стиранием

Несимметричный однородный канал без памяти отличается от рассмотренного тем, что вероятности перехода зависят от того, какой символ передается. Поэтому $P(1/0) \neq P(0/1)$ (рис.5.5).

Симметричный однородный канал без памяти со стиранием содержит дополнительный выходной символ, обозначенный на рис.5.6 знаком "?". Вероятность правильной передачи нуля и единицы в двоичном симметричном канале определяется не только условными вероятностями перехода $P(1/0) = P(0/1) = p$, но и вероятностью стирания p_c . Состояние

стирания символа в канале связи возникает всякий раз, когда решающая схема демодулятора, принимающая решение о том, какой из символов "0" или "1" передается, не может надежно опознать этот символ.

5.2.6. Математическая модель дискретно-непрерывного канала связи

В этом канале входные сигналы дискретны, а выходные — непрерывны.

Математические модели дискретно-непрерывного канала с аддитивным гауссовым белым шумом почти совпадает с моделями непрерывного канала, за исключением того, что на входе вместо непрерывного задают дискретный сигнал, который определяется количеством дискретных значений (букв алфавита) и соответствующими априорными вероятностями. Функции правдоподобия находятся аналогично. Например, для канала с аддитивным гауссовским шумом эта функция

$$w(\bar{y} / b_i) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{N_0} \int_0^T [z(t) - u_i(t)]^2 dt}, \quad (5.18)$$

где b_i — i -е дискретное значение сигнала; $u_i(t)$ — форма сигнала, соответствующая букве b_i .

Функция (5.18) задается при условии, что все параметры сигнала $u_i(t)$ (за исключением информационного) известны в месте приема. В случае, если сигнал содержит паразитные параметры, функция правдоподобия зависит и от.

Если плотность распределения вероятности (5.18) не зависит от времени и от того, какие символы передавались ранее, дискретно-непрерывный канал называется *однородным (стационарным) и без памяти*. Если эта плотность зависит от времени, канал называется *нестационарным*. Канал имеет память, если плотность распределения вероятности зависит от предыдущих символов.

5.2.7. Особенности оптических каналов связи

Оптические каналы связи могут быть *ненаправленными и направленными*.

В ненаправленном оптическом канале электромагнитное поле распространяется в области пространства, которое определяется диаграммой направленности передающей антенны. Средой распространения при этом является свободное космическое, воздушное или водное пространство.

В направленных каналах для распространения электромагнитного поля используются специальные волноводы — оптические кабели.

В простейшем типе ненаправленного канала — космическом пространстве искажения сигнала отсутствуют. Изменяется только амплитуда и происходит его задержка.

Если в среде распространения появляются физические неоднородности и посторонние частицы, то характер принимаемого сигнала будет определяться также и эффектами, связанными с рассеянием электромагнитного поля этими неоднородностями. При рассеивании, как правило, возникает многолучевое распространение, что приводит, как и в радиоканале, к замираниям сигнала.

В направленном канале оптический луч распространяется в оптическом кабеле, основным элементом которого является оптический волновод — круглый стержень из оптически прозрачного диэлектрика. Оптические волноводы называют также волоконными световодами.

Основными параметрами волоконных световодов являются: *затухание сигнала, дисперсия и полоса пропускания.*

Затухание обусловлено в основном поглощением энергии в диэлектрике и ее рассеянием из-за неоднородностей, которые возникают при изготовлении оптического волновода. Дисперсия обусловлена множеством типов колебаний (модами) распространяющегося поля. Эти колебания в направленном канале накладываются друг на друга, причем каждая мода имеет свой путь и свою скорость распространения. Поэтому при наложении множества задержанных во времени колебаний возникает эффект, напоминающий многолучевое распространение электромагнитных волн.

При передаче импульсов вследствие этого эффекта возникает искажение их формы — импульсы удлиняются. Это ведет к ограничению скорости передаваемой информации.

Особенностью оптических систем связи является случайный характер сигнала на выходе преобразователя (фотодетектора) принятого поля в электрическое напряжение или ток, вне зависимости от того, является оптическое поле стохастическим из-за его рассеяния, или детерминированным. Процесс выделения поля фотодетектором сопровождается помехами различных типов. В ненаправленном канале наиболее мощной помехой является фоновое излучение. Его можно существенно уменьшить, применяя оптические волноводы. Среди других помех наибольшее влияние оказывают флуктуационные шумы, возникающие в фотодетекторе, во входной цепи усилителя оптического приемника, а также в других усилителях.

5.3. Воздействие сигналов и помех на линейные и нелинейные звенья каналов связи

5.3.1. Постановка задачи

При прохождении сигналов и помех через каналы связи происходит изменение их реализаций.

Для оценки качества передачи информации необходимо знать эти изменения, уметь рассчитывать соответствующие вероятности.

Расчеты законов распределения случайных процессов и их статистических характеристик для большинства реальных каналов связи являются достаточно сложными.

Упрощение расчетов может быть достигнуто: а) разбиением канала связи на отдельные функционально определенные блоки; б) представлением каждого отдельного блока в виде последовательно соединенных линейных и нелинейных четырехполюсников; в) разбиением нелинейных четырехполюсников на две части: линейную (инерционную) и нелинейную (безынерционную).

Применение подобной идеализации позволяет свести сложную задачу расчета к двум частным: а) расчету закона распределения на выходе линейного четырехполюсника; б) расчету закона распределения на выходе нелинейного четырехполюсника.

Каждая из этих задач формулируется следующим образом. Для заданной n -мерной плотности распределения вероятности входного случайного процесса $\xi_1(t)$ найти n -мерную плотность распределения вероятностей выходного случайного процесса $\xi_2(t)$.

Наиболее сложной является задача определения n -мерной плотности на выходе линейного (инерционного) четырехполюсника, которая легко решается только для одного частного случая, когда на входе действует нормальный шум. Менее сложной (но трудно решаемой задачей) является задача определения плотности распределения вероятности случайного процесса на выходе безынерционного нелинейного звена.

Как показано в четвертой главе, по известной плотности распределения можно вычислить моменты распределения, в том числе среднее значение и корреляционную функцию. В ряде случаев можно вычислить эти моменты, не прибегая к более сложному вычислению плотности распределения вероятности.

5.3.2. Прохождение случайных сигналов и помех через линейные системы

Пусть четырехполюсник является линейным и на его вход подан стационарный случайный процесс с нулевым средним значением. Спектральная плотность этого процесса на входе равна

$G_1(\omega)$, а корреляционная функция — $B_1(\tau)$. Коэффициент передачи четырехполосника равен $K(j\omega)$.

Требуется найти корреляционную функцию $B_2(\tau)$ и соответствующую ей спектральную плотность мощности случайного процесса $G_2(\omega)$ на выходе четырехполосника.

Если на вход подан стационарный случайный процесс с нулевым средним значением, то на выходе установившийся процесс также будет стационарным.

Можно показать, что спектральная плотность мощности входного стационарного случайного процесса определяется также, как и спектральная плотность мощности детерминированного сигнала, т.е.

$$G_1(\omega) = \lim_{T \rightarrow \infty} \frac{|S_1(j\omega)|^2}{T}, \quad (5.19)$$

где $S_1(j\omega) = \int_0^T \xi_1(t) e^{-j\omega t} dt$ — спектральная плотность отрезка

входного стационарного случайного процесса $\xi_1(t)$ в интервале $0-T$

Спектральная плотность мощности выходного случайного стационарного процесса

$$G_2(\omega) = \lim_{T \rightarrow \infty} \frac{|S_1(j\omega)K(j\omega)|^2}{T} = G_1(\omega) \cdot |K(j\omega)|^2, \quad (5.20)$$

так как спектральная плотность отрезка выходной стационарной случайной функции равна

$$S_2(j\omega) = S_1(j\omega)K(j\omega).$$

Корреляционная функция случайного процесса на выходе линейного четырехполосника

$$B_2(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_2(\omega) e^{j\omega\tau} d\omega. \quad (5.21)$$

Если линейный четырехполосник задан импульсной характеристикой $h(t)$, то его реакция на входной случайный процесс $\xi_1(t)$ может быть вычислена с помощью интеграла Дюамеля

$$\xi_2(t) = \int_0^t h(\tau) \xi_1(t - \tau) d\tau. \quad (5.22)$$

Полагаем, что входной случайный процесс является стационарным. В этом случае при вычислении среднего значения и корреляционной функции можно воспользоваться формулами (4.17), (4.19). Среднее значение

$$M\{\xi_2(t)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi_2(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \int_0^t h(\tau) \xi_1(t - \tau) d\tau dt =$$

$$= \int_0^t h(\tau) \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi_1(t - \tau) dt d\tau = \int_0^t h(\tau) M\{\xi_1(t)\} d\tau = M\{\xi_1(t)\} \int_0^t h(\tau) d\tau. \quad (5.23)$$

Корреляционная функция

$$\begin{aligned} B_2(t_1, t_2) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi_2(t_1) \xi_2(t_2) dt = \\ &= \int_0^{t_1} h(\tau_1) \int_0^{t_2} h(\tau_2) \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi_1(t_1 - \tau_1) \xi_1(t_2 - \tau_2) dt \right] d\tau_1 d\tau_2 = \\ &= \int_0^{t_1} h(\tau_1) \left[\int_0^{t_2} h(\tau_2) B_1(t_2 - t_1 - \tau_2 + \tau_1) d\tau_2 \right] d\tau_1. \end{aligned} \quad (5.24)$$

Как следует из (5.23), (5.24), при преобразовании линейным четырехполосником случайного стационарного процесса он становится нестационарным.

Признаками нестационарности являются: а) зависимость среднего значения от времени; б) зависимость корреляционной функции не от разности $t_2 - t_1$, а от моментов времени t_2 и t_1 .

Отметим следующие особенности, связанные с прохождением случайных сигналов и помех через линейные системы.

1. Если входной случайный процесс $\xi_1(t)$ подчиняется нормальному закону, распределения, то и выходной случайный процесс $\xi_2(t)$ также подчиняется нормальному закону.

2. Случайный процесс $\xi_1(t)$, подчиняющийся любому закону распределения, при прохождении через линейный четырехполосник нормализуется, причем закон распределения случайного процесса $\xi_2(t)$, тем ближе к нормальному закону, чем уже полоса пропускания этого четырехполосника.

Доказательства этих утверждений можно найти в [28].

Пример 5.1. На входе линейной цепи (рис. 5.7) действует напряжение нормального белого шума с нулевым средним значением и спектральной плотностью N_0 [В²/Гц]. Найти спектральную плотность мощности, корреляционную функцию $B_2(\tau)$ и дисперсию σ_2^2 на выходе линейной цепи.

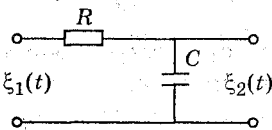


Рис. 5.7. Линейная цепь

Решение. Коэффициент передачи линейной цепи

$$K(j\omega) = \frac{\xi_2(t)}{\xi_1(t)} = \frac{1}{1 + j\omega RC}.$$

Спектральную плотность на выходе найдем по формуле (5.20):

$$G_2(\omega) = N_0 |K(j\omega)|^2 = \frac{N_0}{1 + (\omega RC)^2}.$$

Корреляционная функция случайного процесса на выходе линейной цепи определяется формулой (5.21):

$$B_2(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_1(\omega) |K(j\omega)|^2 e^{j\omega\tau} d\omega = \frac{N_0}{2\pi} \int_{-\infty}^{\infty} \frac{1}{1 + (\omega RC)^2} e^{j\omega\tau} d\omega = \frac{N_0}{2RC} e^{-\frac{|\tau|}{RC}}.$$

Дисперсия

$$\sigma_2^2 = B_0(0) = N_0 / 2RC.$$

5.3.3. Прохождение случайных сигналов и помех через нелинейные системы

Задача формулируется следующим образом. Для данной функции распределения $w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ случайной величины $\xi_1(t)$ на входе и заданном операторе нелинейного преобразования $\bar{y} = f(\bar{x})$, найти функцию распределения случайной величины $\xi_2(t)$ на выходе нелинейного звена системы.

Если функция распределения $w(y_1, y_1, \dots, y_n; t_1, \dots, t_n)$ будет найдена, то по ней нетрудно найти всевозможные статистические характеристики (моменты n -го порядка).

Задача более наглядна и понятна для одномерного случая, когда нелинейный оператор представлен только одним уравнением $y = f(x)$, а плотность распределения вероятности входного случайного процесса $\xi_1(t)$ равна $w(x)$.

Пусть существует обратная функция $x = \varphi(y)$. В этом случае, если случайная величина ξ_1 находится в пределах $x_0 < \xi_1 \leq x_0 + dx$, то случайная величина ξ_2 будет находиться в пределах

$y_0 < \xi_2 \leq y_0 + dy$ (рис.5.8). Вероятность этих событий равны. Поэтому будут равны и заштрихованные площади:

$$w(x)dx = w(y)dy.$$

Из полученного равенства находим

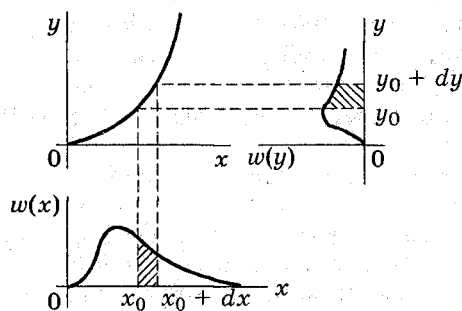


Рис.5.8. Преобразование плотности распределения вероятности безинерционной нелинейной цепью

$$w(y) = w(x) \frac{1}{|dy/dx|} = w[\varphi(y)] \frac{1}{|dy/dx|}. \quad (5.26)$$

Если известны оператор $y = f(x)$ и двумерная плотность распределения вероятности входного случайного процесса $w(x_1, x_2; t_1, t_2)$, то статистические характеристики случайного процесса (среднее значение и корреляционная функция) определяются по формулам, приведенным ниже.

Среднее значение

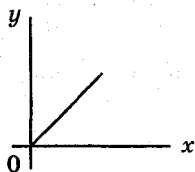
$$M\{\xi_2(t)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) w(x_1, x_2; t_1, t_2) dx_1 dx_2 = \int_{-\infty}^{\infty} f(x_1) w(x_1, t) dx_1, \quad (5.27)$$

Корреляционная функция

$$B(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) f(x_2) w(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (5.28)$$

Пример 5.2. Оператор нелинейного преобразования задан уравнениями (рис.5.9)

$$\left. \begin{aligned} y &= x \\ y &= 0 \end{aligned} \right\} \begin{aligned} x &> 0, \\ x &\leq 0. \end{aligned}$$



Плотность распределения случайной величины $\xi(t)$ на входе

$$w(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}.$$

Рис.5.9.
Оператор
нелинейного
преобразования

Требуется найти $w(y)$ плотность распределения случайной величины на выходе нелинейной цепи (рис.5.10).

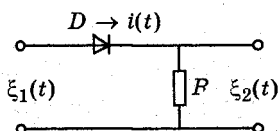


Рис.5.10. Нелинейная
цепь

Решение. Задача связана с прохождением случайного сигнала через диод с линейной характеристикой. Следует обратить внимание на то, что, какие бы ни были по величине отрицательные значения случайного процесса $\xi_1(t)$, значения случайного процесса $\xi_2(t)$ равны нулю. (рис.5.11). Следовательно, вероятность того, что случайная величина на выходе цепи в некоторый момент времени t равна нулю определяется выражением:

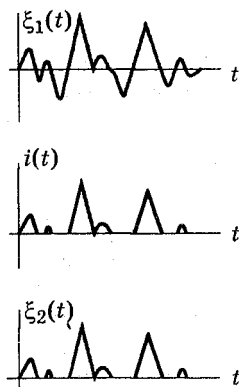


Рис.5.11.
Временные
диаграммы на входе
и выходе
нелинейной цепи

$$P(\xi_2(t) = 0) = \int_{-\infty}^0 \omega(x) dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^0 e^{-\frac{x^2}{2\sigma^2}} dx = 0,5.$$

Для положительных значений случайной величины

$$w(y) = w(x) \frac{1}{|dy/dx|} = w[x(y)] \frac{1}{|dy/dx|} = w[x(y)],$$

так как $dy/dx = 1$.

Окончательное выражение для спектральной плотности

$$w(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-y^2/2\sigma^2}, \quad y > 0$$

$$w(y) = 0,5, \quad y = 0.$$

Контрольные вопросы

1. По каким признакам можно классифицировать каналы связи?
2. Что понимают под линейными и нелинейными искажениями сигналов при их передаче по каналу связи?
3. Что понимают под аддитивной помехой? В чем ее отличие от мультипликативной помехи?
4. Как математически описываются дискретный, дискретно-непрерывный и непрерывный каналы?
5. Какие основные задачи решаются при рассмотрении прохождения случайных воздействий через канал связи и его звенья?

Глава 6. ТЕОРИЯ ПЕРЕДАЧИ И КОДИРОВАНИЯ СООБЩЕНИЙ

6.1. Общие сведения

6.1.1. Количество информации в дискретных сообщениях.

Понятие энтропии

При анализе систем связи используют *количественную меру*, позволяющую оценивать объем информации, содержащейся в сообщении, и объем передаваемой информации.

Для определения этой количественной меры обратимся к модели дискретного канала связи (разд.5.2). Рассмотрим случай, когда кодовые символы $\{a_i\}$ ($i = 1, 2, \dots, m$) независимы, а источник полностью описывается априорными вероятностями $P(a_i)$. При отсутствии помех значения принятых символов $\{b_j\}$ не отличаются от переданных $\{a_i\}$. При наличии помех между символами $\{a_i\}$, $\{b_j\}$ нет однозначного соответствия и любой символ a_i может перейти в любой другой символ b_j с некоторой вероятностью перехода $P(b_j/a_i)$. Зная априорные вероятности $P(a_i)$ и вероятности перехода $P(b_j/a_i)$, можно вычислить вероятность появления символа a_i при условии, что на выходе наблюдался символ b_j . Эта условная вероятность $P(a_i/b_j)$ называется *апостериорной (послеопытной) вероятностью*.

Апостериорная вероятность находится из известного равенства [28], определяющего совместную вероятность появления событий a_i , b_j .

$$P(a_i, b_j) = P(a_i)P(b_j / a_i) = P(b_j)P(a_i / b_j), \quad (6.1)$$

Апостериорная вероятность того, что в тех случаях, когда был принят символ b_j , передавался символ a_i

$$P(a_i / b_j) = \frac{P(a_i)}{P(b_j)} P(b_j / a_i), \quad (6.2)$$

Вероятность $P(a_i)$ определяется по результатам наблюдения за работой дискретного источника в прошлом и называется *априорной вероятностью*. Выражение $P(b_j/a_i)$ определяет степень правдоподобия того, что если был послан какой-то символ a_i , то будет принят символ b_j . Эта вероятность называется функцией правдоподобия и обозначается буквой $L(a_i)$. Вероятность $P(b_j)$ является известной величиной и определяется из условия нормировки:

$$\sum_{i=1}^m P(a_i / b_j) = \sum_{i=1}^m K P(a_i) P(b_j / a_i) = 1,$$

где k — постоянная величина.

С учетом приведенных замечаний апостериорная вероятность

$$P(a_i / b_j) = KP(a_i)L(a_i). \quad (6.3)$$

Рассмотрим процесс приема символа a_i . До его поступления в канал связи неопределенность относительно посылаемого символа определялась априорной вероятностью $P(a_i)$. Приняв символ b_j приемник располагает определенными данными относительно действительно переданного символа a_i . Его неопределенность теперь представляется новым значением вероятности $P(a_i/b_j)$. Если это новое значение равно значению априорной вероятности $P(a_i)$ то в результате приема каких-либо новых сведений о передаваемом символе не поступает. Если $P(a_i/b_j) > P(a_i)$, то шанс приема конкретного символа a_i , возрастает, и при сравнении указанных вероятностей может быть извлечена информация о символе a_i . За меру количества извлекаемой информации принимают величину

$$I(a_i, b_j) = \log \frac{P(a_i / b_j)}{P(a_i)}. \quad (6.4)$$

Основание логарифма может быть любым. Принято использовать основание 2, а количество информации измерять в двоичных единицах или битах.

Пусть $P(a_i/b_j) = P(a_i)$. Тогда

$$I(a_i, b_j) = \log \frac{P(a_i)}{P(a_i)} = \log 1 = 0.$$

Как отмечалось ранее, информация, доставляемая символом b_j о символе a_i , в этом случае равна нулю.

Рассмотрим случай приема, когда принимаемый символ b_j однозначно определяет символ a_i . В этом случае $P(a_i/b_j) = 1$ и количество извлекаемой информации на выходе канала о символе a_i

$$I(a_i, b_j) = I(a_i) = \log \frac{1}{P(a_i)} = -\log P(a_i). \quad (6.5)$$

Величина $I(a_i)$ называется *собственной информацией* о событии a_i . Это количество информации, которое может быть передано по каналу связи.

Собственная информация является положительной величиной. Она тем больше, чем меньше вероятность $P(a_i)$.

Перепишем выражение (6.4)

$$\begin{aligned} I(a_i, b_j) &= \log \frac{P(a_i / b_j)}{P(a_i)} = \log \frac{P(a_i / b_j)P(b_j)}{P(a_i)P(b_j)} = \\ &= \log \frac{P(a_i, b_j)}{P(a_i)P(b_j)} = \log \frac{P(b_j / a_i)}{P(b_j)}. \end{aligned} \quad (6.6)$$

Из (6.6) следует, что информация, доставляемая символом b_j о символе a_i равна информации доставляемой символом a_i о символе b_j . Поэтому $I(a_i, b_j)$ называется *взаимной информацией*.

На практике представляет интерес *средняя взаимная информация*, равная количеству информации о множестве A передаваемых символов, которая в среднем содержится во множестве B принимаемых символов.

$$\begin{aligned} I(A, B) &= \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) I(a_i, b_j) = \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log \frac{P(a_i / b_j)}{P(a_i)} = \\ &= \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log \frac{P(a_i, b_j)}{P(a_i)P(b_j)}. \end{aligned} \quad (6.7)$$

С учетом (6.1) выражение (6.7) перепишем в виде

$$\begin{aligned} I(A, B) &= \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(a_i / b_j) - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(a_i) = \\ &= I(A) - I(A / B) = H(A) - H(A / B), \end{aligned} \quad (6.8)$$

где

$$I(A / B) = - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(a_i / b_j) = H(A / B), \quad (6.9)$$

$$I(A) = - \sum_{i=1}^m P(a_i) \log P(a_i) \sum_{j=1}^n P(b_j / a_i) = - \sum_{i=1}^m P(a_i) \log P(a_i) = H(A), \quad (6.10)$$

так как $\sum_{j=1}^n P(b_j / a_i) = 1$.

Учитывая формулу (6.1), выражение (6.7) можно записать также в виде:

$$I(A, B) = I(B) - I(B / A) = H(B) - H(B / A), \quad (6.11)$$

где

$$H(B) = - \sum_{j=1}^n P(b_j) \log P(b_j), \quad (6.12)$$

$$H(B / A) = - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(b_j / a_i). \quad (6.13)$$

Обратимся к формуле (6.8). Входящее в нее выражение для $I(A)$ определяется формулой (6.10) и равно среднему количеству информации, которая содержится в возможных символах на выходе источника дискретных сообщений. Формула (6.10) идентична выражению для энтропии системы в статистической механике.

Поэтому величину $I(A)$ называют *энтропией* дискретного источника и обозначают $H(A)$.

Выражение $I(A/B)$, входящее в формулу (6.10), определяет среднее количество информации, содержащееся во всевозможных символах на выходе источника дискретных сообщений до момента

приема символа на выходе канала при известном множестве символов на выходе этого канала. Математически величина $I(A/B)$ определяется формулой (6.9). Величину $I(A/B)$ называют *условной энтропией* и обозначают буквой $H(A/B)$. Условную энтропию называют также *ненадежностью* дискретного канала связи.

Рассмотрим, реализации двух дискретных источников сообщений $\{a_i\}$, $\{b_j\}$. Пусть между этими реализациями имеется статистическая связь, определяемая совместной вероятностью $P(a_i, b_j)$. По аналогии с (6.10) определим энтропию множества совместных событий, заключающихся в появлении символов a_i , b_j :

$$H(A, B) = - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(a_i, b_j). \quad (6.14)$$

Подставляя в (6.14) выражение (6.1), получим

$$H(A, B) = H(A) + H(B/A) = H(B) + H(A/B), \quad (6.15)$$

где
$$H(B, A) = - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(b_j / a_i),$$

$$H(A, B) = - \sum_{i=1}^m \sum_{j=1}^n P(a_i, b_j) \log P(b_j / a_i). \quad (6.16)$$

Если символы статистически независимы, то $P(b_j/a_i) = P(b_j)$, $P(a_i/b_j) = P(a_i)$

$$H(A, B) = H(A) + H(B). \quad (6.17)$$

Рассмотрим основные свойства энтропии.

1. Энтропия величина вещественная, ограниченная и неотрицательная. Это следует из (6.10).

2. Энтропия равна нулю, если значение символа (сообщения) известно заранее. В этом случае вероятность одного из символов равна 1, а вероятности остальных символов равны нулю.

3. Энтропия максимальна, если все состояния дискретного источника равновероятны.

Подставляя в (6.10) $P(a_i) = 1/m$, получим

$$H_{\max}(A) = - \sum_{i=1}^m \frac{1}{m} \log \frac{1}{m} = \log m.$$

4. Условные энтропии

$$H(A/B) \leq H(A), \quad H(B/A) \leq H(B).$$

Равенство наблюдается тогда, когда символы a_i , b_j статистически независимы. В этом случае $P(a_i/b_j) = P(a_i)$, $P(b_j/a_i) = P(b_j)$ для всех индексов i, j .

$H(A/B) = 0$, если между символами a_i и b_j наблюдается однозначная связь.

5. Энтропия объединения (6.15) двух зависимых сообщений равна сумме двух энтропий: безусловной и условной.

6. Энтропия объединения (6.17) двух независимых сообщений равна сумме энтропий этих сообщений.

Учитывая приведенные свойства энтропии, проанализируем равенство (6.8). Средняя взаимная информация равна разности энтропий: энтропии дискретного источника $H(A)$ и условной энтропии $H(A/B)$. Поскольку $H(A) \geq 0$, $H(A/B) \geq 0$, то при передаче информации по каналу связи с помехами часть информации теряется. Количественно потери определяются условной энтропией $H(A/B)$. Потери равны нулю, если $H(A/B) = 0$. В этом случае между символами a_i и b_j наблюдается однозначная связь и реализация ансамбля A полностью восстанавливается. Потери максимальны, если $H(A/B) = H(A)$, а символы a_i , b_j являются статистически независимыми.

Аналогично анализируется равенство (6.11). Из этого равенства находим, что энтропия множества сообщений B на выходе канала связи

$$H(B) = I(A, B) + H(B/A). \quad (6.18)$$

Следовательно, при передаче информации по каналу связи с помехами к средней взаимной информации $I(A, B)$ добавляется средняя условная энтропия $H(B/A)$. Эта добавочная энтропия обусловлена только помехами. При отсутствии помех $H(B/A) = 0$ и $H(B) = I(A, B)$. Кроме того, $H(A/B) = 0$ и $H(B) = H(A)$.

6.1.2. Производительность источника дискретных сообщений.

Скорость передачи информации

Энтропия дискретного источника, отнесенная к среднему времени передачи одного символа T , называется *производительностью источника дискретных сообщений*

$$H'(A) = H(A) / T. \quad (6.19)$$

Скорость передачи информации равна количеству взаимной информации, деленному на среднее время передачи одного символа:

$$R = I'(A, B) / T = \frac{H(A) - H(A/B)}{T} = \frac{H(B) - H(B/A)}{T}. \quad (6.20)$$

Вычислим скорость передачи информации по двоичному симметричному каналу (см. рис.5.4). На вход канала поступают символы $a_1 = 0$ и $a_2 = 1$ с вероятностями $P(a_1)$ и $P(a_2)$, причем $P(a_1) + P(a_2) = 1$. На выходе канала присутствуют двоичные символы $b_1 = 0$ и $b_2 = 1$. Вероятность ошибки при передаче любого

символа равна P . Переходные вероятности: $P(b_1/a_1) = 1 - p$; $P(b_1/a_1) = p$; $P(b_2/a_1) = p$.

Вычислим взаимную информацию по формуле (6.11).

Энтропия

$$H(B) = - \sum_{j=1}^2 P(b_j) \log P(b_j) = -P(b_1) \log P(b_1) - P(b_2) \log P(b_2),$$

где $P(b_1) = P(a_1)P(b_1/a_1) + P(a_2)P(b_1/a_2) = P(a_1)(1-p) + p(a_2)p$; $P(b_2) = 1 - P(b_1)$.

Условная энтропия при $P(a_1) = P(a_2) = 0,5$

$$H(B/A) = - \sum_{i=1}^2 P(a_i) \sum_{j=1}^2 P(b_j/a_i) \log P(b_j/a_i) = -p \log p - (1-p) \log(1-p).$$

Подставляя $H(B)$ и $H(B/A)$ в (6.11), получим

$$I(A,B) = 1 + p \log p + (1-p) \log(1-p).$$

Скорость передачи информации определяется по формуле (6.20). При отсутствии помех $p = 0$, $I(A,B) = 1$. Скорость равна $1/T$. При $p = 0,5$, $I(A,B) = 0$, скорость передачи информации равна нулю.

6.2. Эффективное кодирование источников дискретных сообщений

Как показано выше, энтропия $H(A)$ источника дискретных сообщений определяется априорными вероятностями реализации этих сообщений $P(a_i)$ $i = 1, 2, \dots, m$, причем $H(A) \leq H_{\max}(A) = \log m$.

Количество информации, доставляемое этим источником, меньше максимального в $H_{\max}(A)/H(A)$ раз.

Относительная разность

$$\chi = \frac{H_{\max}(A) - H(A)}{H_{\max}(A)} = 1 - \frac{H(A)}{H_{\max}(A)} = 1 - \mu, \quad (6.21)$$

где $\mu = H(A)/H_{\max}(A)$, называется *избыточностью источника дискретных сообщений*.

Одной из задач кодирования является устранение избыточности сообщения.

Кодирование, обеспечивающее устранение избыточности в закодированном сообщении, называется *эффективным*.

Возможность реализации эффективного кодирования была доказана К.Шенноном. Теорема о кодировании сообщения при отсутствии помех утверждает, что *минимальное среднее число элементов кодовой последовательности на выходе кодирующего устройства, соответствующее одному символу дискретного источника, можно сделать сколь угодно близким к $H(A)$* .

Эффективное кодирование реализуется применением *неравномерного* кода, у которого более короткие кодовые комбинации используются для передачи более вероятных символов и более длинные — для передачи менее вероятных символов.

Часто неравномерный код является двоичным. Примером такого кода является *код Шеннона-Фано*.

Задача построения эффективного кода формулируется следующим образом. Для сообщения, записанного с помощью букв некоторого алфавита $A = \{a_1, a_2, \dots, a_m\}$ содержащего m букв, требуется указать правило кодирования, которое устанавливает однозначное соответствие между буквами алфавита a_i и двоичными кодовыми последовательностями, состоящими из символов "0" и "1".

Эффективный код должен обеспечивать однозначное декодирование, т.е. по принятой двоичной кодовой последовательности восстанавливать передаваемый символ (букву алфавита). Кроме того, на передачу одной буквы в среднем должно тратиться минимальное число нулей и единиц двоичной кодовой последовательности.

Для обеспечения однозначности декодирования эффективный код должен удовлетворять следующему требованию: никакое более короткое слово (последовательность) эффективного кода не должно являться началом другого более длинного слова.

Построение кода Шеннона-Фано поясняется табл.6.1. Сообщения располагают в таблице по мере убывания их вероятностей. Далее буквы алфавита a_i разбивают на группы и подгруппы, обеспечивая их равновероятность. Верхней группе присваивается символ "0", а нижней — "1". Каждую из полученных групп, в свою очередь разбивают на две подгруппы с одинаковыми суммарными вероятностями и т.д. Процесс повторяется до тех пор, пока в каждой подгруппе не останется по одному символу.

Таблица 6.1

Сообщения a_i	Вероятность сообщения $P(a_i)$	Группы и подгруппы на этапах разбиения				Код	Число символов в кодовой комбинации
		1-й	2-й	3-й	4-й		
a_1	0,25	0	0	0	0	00	2
a_2	0,25		1			01	2
a_3	0,125	1	0			1	100
a_4	0,125			0		101	3
a_5	0,0625		1	1	1	1100	4
a_6	0,0625				0	1101	4
a_7	0,0625				1	1110	4
a_8	0,0625			1	1111	4	

Вычислим энтропию, используя данные табл.6.1.

$$H(A) = - \sum_{i=1}^8 P(a_i) \log P(a_i) = 2,75 \text{ бит.}$$

Средняя длина двоичной кодовой комбинации

$$n_{\text{ср}} = \sum_{i=1}^8 P(a_i) n_i = 2,75.$$

Энтропия, приходящаяся на один элемент двоичной кодовой последовательности:

$$H(A)/n_{\text{ср}} = 2,75/2,75 = 1.$$

Это значение совпадает с максимальной энтропией на выходе двоичного дискретного источника сообщений, которая равна $H_{\text{max}}(A) = \log 2 = 1$.

6.3. Информационные характеристики источников непрерывных сообщений

6.3.1. Дифференциальная энтропия

Понятия теории информации, относящиеся к дискретным сообщениям, можно распространить и на непрерывные сообщения.

Пусть ξ — непрерывная случайная величина с плотностью распределения вероятности $w(x)$. Эта величина является отсчетом

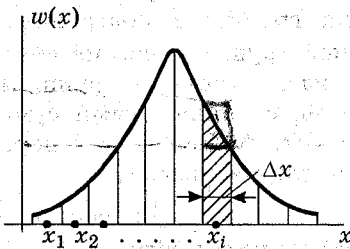


Рис.6.1. Плотность распределения вероятности случайной величины

непрерывной случайной функции в некоторый момент времени t (рис.6.1). Разобьем диапазон изменения непрерывной случайной величины на конечное число N малых интервалов шириной Δx . Поскольку Δx мало, вероятность того, что случайная величина ξ находится в пределах i -го отрезка Δx

$$P_i \left\{ x_i - \frac{\Delta x}{2} < \xi \leq x_i + \frac{\Delta x}{2} \right\} \cong w(x_i) \Delta x$$

Аналогично можно найти вероятность того, что случайная величина ξ будет находиться в пределах любого другого отрезка Δx .

Располагая рядом дискретных значений вероятностей, вычислим энтропию.

$$H_0(\xi) = - \sum_{i=1}^N p_i \log p_i = - \sum_{i=1}^N w(x_i) \Delta x \log [w(x_i) \Delta x],$$

где N — число полученных дискретных значений. Увеличивая число N за счет уменьшения Δx , в пределе получим выражение для энтропии непрерывной случайной величины:

$$\begin{aligned} H(\xi) &= \lim_{\Delta x \rightarrow 0} H_0(\xi) = \lim_{\Delta x \rightarrow 0} \left[- \sum_{i=1}^N w(x) \Delta x \log w(x) \right] - \lim_{\Delta x \rightarrow 0} \sum_{i=1}^N w(x) \Delta x \log \Delta x = \\ &= - \int_{-\infty}^{\infty} w(x) \log w(x) dx + \lim_{\Delta x \rightarrow 0} \log \frac{1}{\Delta x}. \end{aligned} \quad (6.22)$$

Первое слагаемое полученного выражения имеет конечное значение. Обозначим его $h(\xi)$:

$$h(\xi) = - \int_{-\infty}^{\infty} w(x) \log w(x) dx. \quad (6.23)$$

Второе слагаемое зависит от шага квантования и при $\Delta x \rightarrow 0$ стремится к ∞ . Это значит, что среднее количество информации, содержащееся в одном отсчете непрерывной случайной функции (сообщении) также равно бесконечности.

По этой причине $H(\xi)$ не может быть количественной мерой, которая позволила бы оценить объем информации, содержащейся в одном отсчете непрерывной случайной функции (сообщения).

В качестве этой меры принимают первое слагаемое выражения (6.22) и отбрасывают второе. Поскольку первое выражение зависит от дифференциальной плотности распределения вероятности, оно получило название *дифференциальной энтропии*.

Для двух непрерывных случайных величин ξ и γ , которые характеризуются двумерной плотностью распределения вероятности $w(x, y)$, можно по аналогии с дискретными случайными величинами найти взаимную информацию, а следовательно, и скорость передачи $R = \frac{1}{T} I(\xi, \gamma)$:

$$I(\xi, \gamma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y) I(x, y) dx dy, \quad (6.24)$$

где
$$I(x, y) = \log \frac{w(y/x)}{w(y)} = \log \frac{w(x/y)}{w(x)} = \log \frac{w(x, y)}{w(x)w(y)}. \quad (6.25)$$

Ее можно представить также в виде, аналогичном (6.8), (6.11):

$$I(\xi, \gamma) = h(\xi) - h(\xi/\gamma), \quad (6.26)$$

$$I(\gamma, \xi) = h(\gamma) - h(\gamma/\xi), \quad (6.27)$$

где $h(\xi / \gamma) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y) \log w(x / y) dx dy$ — условная дифферен-

циальная энтропия непрерывной случайной величины ξ при известной случайной величине γ ;

$h(\gamma / \xi) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y) \log w(y / x) dx dy$ — условная дифференциаль-

ная энтропия непрерывной случайной величины γ при известной случайной величине ξ .

Дифференциальная энтропия, в отличие от энтропии дискретной случайной величины, является относительной мерой неопределенности и не обладает рядом ее свойств. В частности, она может быть как положительной, так и отрицательной. Значение дифференциальной энтропии зависит от масштаба случайной величины, следовательно от выбора единицы ее измерения.

Изменим масштаб случайной величины ξ в k раз. Если $x_1 = kx$, то в соответствии с формулой (5.26) $w(x_1) = w(x)/k$. Тогда

$$h(\xi_1) = - \int_{-\infty}^{\infty} w(x) \log w(x) dx = - \int_{-\infty}^{\infty} \frac{w(x)}{k} \log \frac{w(x)}{k} k dx = h(\xi) + \log k.$$

Пример 6.1. Для непрерывной случайной величины ξ , распределенной по нормальному закону

$$w(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2 / 2\sigma^2},$$

где a — среднее значение случайной величины, σ^2 — дисперсия, найти дифференциальную энтропию.

Решение. Дифференциальная энтропия (6.23)

$$\begin{aligned} h(\xi) &= \int_{-\infty}^{\infty} w(x) \left[\log \sqrt{2\pi\sigma^2} + \frac{\log e}{2\sigma^2} (x-a)^2 \right] dx = \log \sqrt{2\pi\sigma^2} \times \\ &\times \int_{-\infty}^{\infty} w(x) dx + \frac{\log e}{2\sigma^2} \int_{-\infty}^{\infty} (x-a)^2 w(x) dx. \end{aligned}$$

Первый интеграл равен 1, а второй — дисперсии σ^2 . Поэтому

$$h(\xi) = \log \sqrt{2\pi\sigma^2} + \frac{1}{2} \log e = \log \sqrt{2\pi e\sigma^2}. \quad (6.28)$$

Дифференциальная энтропия гауссовской случайной величины не зависит от ее среднего значения и возрастает с ростом дисперсии.

6.3.2. ϵ -энтропия

Для передачи непрерывного сообщения и соответствующего этому сообщению сигнала с абсолютной точностью требуется бесконечно большое количество информации. Этот вывод следует из результата, полученного выше: энтропия непрерывной случайной величины равна бесконечности. Однако передача бесконечно большого количества информации по реальному каналу связи с ограниченной шириной полосы и за ограниченное время оказывается не только невозможной, но и ненужной. Это объясняется тем, что для восприятия информации абсолютно точного ее воспроизведения не требуется.

Пусть на выходе непрерывного канала связи воспроизводимый сигнал $b^*(t)$ отличается от передаваемого $b(t)$ на величину разности $\Delta\xi(t) = b^*(t) - b(t)$. Средний квадрат этой разности

$$\overline{\epsilon^2(t)} = \frac{1}{T} \int_0^T [b^*(t) - b(t)]^2 dt, \quad (6.29)$$

где $0-T$ — интервал времени, в течение которого передается сигнал.

Если средний квадрат разности не превышает некоторого заданного числа ϵ_0 , то сигналы $b^*(t)$ и $b(t)$ считаются *эквивалентными*.

Это значит, что, если послан один сигнал, а принят другой, эквивалентный ему, то по данному критерию переданный сигнал считается принятым верно.

Минимальное количество информации, содержащееся в принятом сигнале $b^*(t)$ относительно $b(t)$, при котором они еще эквивалентны, называется ϵ -энтропией:

$$h_\epsilon(b) = \min I[b(t), b^*(t)] = h[b(t)] - \max h[b(t)/b^*(t)], \quad (6.30)$$

где $h[b(t)]$ — дифференциальная энтропия сигнала $b(t)$, передаваемого по каналу связи в течении времени T ; $h[b(t)/b^*(t)]$ — условная дифференциальная энтропия этого сигнала.

Максимизация условной энтропии означает, что из всех возможных условных распределений, для которых $\overline{\epsilon^2(t)} < \epsilon_0$, выбирается такое, которое обеспечивает максимум условной энтропии.

Вычисление ϵ -энтропии несложно при следующих допущениях.

1. Известны n -мерные плотности распределения вероятности, определяющие случайные сигналы $b(t)$, $b^*(t)$. Такими плотностями являются:

$w(\bar{x}) = w(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ — безусловная n -мерная плотность распределения первичного случайного сигнала $b(t)$;

$w(\bar{x} / \bar{y}) = w(x_1, x_2, \dots, x_n / y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n)$ — условная n -мерная плотность распределения вероятности первичного случайного сигнала.

2. Выборочные значения входного сигнала $b_1, b_2, \dots, b_n$ являются статистически независимыми случайными величинами.

3. Выборочные значения $b_1^*, b_2^*, \dots, b_n^*$, которые являются оценками выборочных значений входного сигнала, также являются статистически независимыми случайными величинами.

4. Спектры сигналов $b(t)$, $b^*(t)$ ограничены верхней граничной частотой F_c . Поэтому, согласно теореме Котельникова, интервал дискретизации $\Delta t = 1/2F_c$. Число выборочных значений в интервале $[0, T]$ равно $n = T/\Delta t = 2F_c T$.

5. Случайный процесс на входе непрерывного канала подчиняется нормальному закону.

6. Мощность (дисперсия) разности $\xi(t)$ является ограниченной величиной.

Определим дифференциальные энтропии.

С учетом приведенных ограничений дифференциальная энтропия сигнала $b(t)$ запишется по аналогии с формулой (6.23):

$$\begin{aligned} h(b(t)) &= - \int \dots \int_{-\infty}^{\infty} w(\bar{x}) \log w(\bar{x}) d\bar{x} = - \int \dots \int_{-\infty}^{\infty} \prod_{i=1}^n w(x_i) \log \prod_{i=1}^n w(x_i) dx_1 dx_2 \dots dx_n = \\ &= -2F_c T \int_{-\infty}^{\infty} w(x) \log w(x) dx = 2F_c T \log \sqrt{2\pi e \sigma_b^2}, \end{aligned}$$

где $w(x)$ — плотность распределения случайной величины, совпадающей с одним из выборочных значений входного сигнала; σ_b^2 — дисперсия случайной величины, равная мощности этого сигнала.

Условная дифференциальная энтропия

$$\begin{aligned} h[b(t) / b^*(t)] &= - \int \dots \int_{-\infty}^{\infty} w(\bar{x} / \bar{y}) \log w(\bar{x} / \bar{y}) d\bar{x} = \\ &= - \int \dots \int_{-\infty}^{\infty} \prod_{i=1}^n w(x_i / y_i) \log \prod_{i=1}^n w(x_i / y_i) dx_1 dx_2 \dots dx_n, \end{aligned}$$

где $w(x_i / y_i)$ — условная плотность распределения случайной величины (выборки) на входе канала.

Поскольку $b(t_i) = b^*(t_i) - \varepsilon(t_i)$, то условная плотность

$$w(x_i / y_i) = w(z_i),$$

где $z_i = y_i - x_i$.

Известно, что при ограниченной дисперсии случайной величины ε_i ее энтропия максимальна, если она подчиняется нормальному закону. Поэтому

$$\begin{aligned} \max h[b(t)/b^*(t)] &= \max h[\varepsilon(t)] = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^n w(z_i) \log w(z_i) dz_i = \\ &= 2F_c T \log \sqrt{2\pi\sigma_\varepsilon^2}, \end{aligned} \quad (6.32)$$

где σ_ε^2 — дисперсия случайной величины $\varepsilon(t_i)$, равная мощности разности сигналов $\varepsilon(t)$.

Подставляя (6.31), (6.32) в (6.30), получим:

$$h_\varepsilon[b(t)] = 2F_c T [\log \sqrt{2\pi\sigma_B^2} - \log \sqrt{2\pi\sigma_\varepsilon^2}] = F_c T \log \frac{\sigma_B^2}{\sigma_\varepsilon^2} = F_c T \log \rho. \quad (6.33)$$

ε -энтропия первичного сигнала $b(t)$ определяется произведением $F_c T$ и отношением сигнал/шум $\rho = \sigma_B^2 / \sigma_\varepsilon^2$.

Сравнивая формулы (6.33), (1.3), находим, что объем сигнала равен ε -энтропии, если $D_c = \log \rho$.

ε -энтропия, отнесенная к длительности интервала T , в течение которого передается сигнал $b(t)$, называется *производительностью источника непрерывных сообщений*

$$H'_\varepsilon(b) = \frac{H_\varepsilon[b(t)]}{T} = F_c \log \rho. \quad (6.34)$$

6.4. Пропускная способность канала связи

6.4.1. Пропускная способность дискретного канала связи

Под пропускной способностью дискретного канала связи C_0 понимают максимальное количество переданной информации.

При этом из всего множества дискретных источников пропускную способность определяет только такой источник, распределение вероятностей $P(A)$ которого обеспечивает максимальное значение передаваемой информации:

$$C_0 = \max_{P(A)} I(A, B) = \max_{P(A)} [H(B) - H(B/A)]. \quad (6.35)$$

Размерность пропускной способности — бит. Пропускная способность C_0 зависит от энтропии $H(A)$, т.е. от свойств передаваемого сообщения.

Пропускная способность зависит также от энтропий $H(B)$, $H(A/B)$, $H(B/A)$, которые определяются как свойствами дискретного канала связи, так и свойствами передаваемого дискретного сообщения.

Пропускную способность канала связи определяют также в расчете на единицу времени

$$C = \frac{C_0}{T} = \max_{P(A)} I(A, B) = \max_{P(A)} \frac{H(A) - H(A/B)}{T} = \max_{P(A)} \frac{H(B) - H(B/A)}{T}, \quad (6.36)$$

где T — средняя длительность передачи одного символа (сообщения).

Пропускная способность C всегда больше скорости передачи информации по каналу связи.

Пример 6.2. Требуется найти пропускную способность m -ичного симметричного канала без памяти в расчете на один символ. Переходные вероятности заданы уравнениями (5.15).

Решение: Применим формулу (6.35). В этой формуле только энтропия $H(B)$ зависит от распределения вероятности $P(A)$. Поэтому максимизируя $H(B)$, находим максимум и выражения (6.35). Вычислим энтропию

$$H(B) = - \sum_{j=1}^m P(b_j) \log P(b_j),$$

где

$$P(b_j) = \sum_{i=1}^m P(a_i) P(b_j / a_i).$$

Максимальное значение энтропии $H_{\max}(B) = \log m$ обеспечивается, если вероятности реализации символов на входе дискретного канала одинаковы и равны $1/m$.

В этом случае

$$P(b_j) = \frac{1}{m} \sum_{i=1}^m P(b_j / a_i) = \frac{1}{m} \left[1 - p + (m-1) \frac{p}{m-1} \right] = \frac{1}{m},$$

$$H(B) = - \frac{1}{m} \sum_{j=1}^m \log \frac{1}{m} = \log m.$$

Вычисляя условную энтропию, убеждаемся, что она не зависит от распределения вероятности $P(A)$.

$$\begin{aligned} H(B/A) &= - \sum_{j=1}^m P(b_j / a_i) \log p(b_j / a_i) = - \left[\frac{p}{m-1} \log \frac{p}{m-1} + \dots + \right. \\ &\left. + \frac{p}{m-1} \log \frac{p}{m-1} + (1-p) \log(1-p) \right] = - \left[p \log \frac{p}{m-1} + (1-p) \log(1-p) \right]. \end{aligned}$$

Подставляя вычисленные значения для энтропий в (6.35), получим

$$C_0 = \log m + p \log \frac{p}{m-1} + (1-p) \log(1-p). \quad (6.37)$$

Для двоичного симметричного канала

$$C_0 = 1 + p \log p + (1-p) \log(1-p).$$

6.4.2. Пропускная способность непрерывного канала связи

Пусть сигнал на выходе канала $z(t)$ определяется сигналом на его входе $v(t)$ и присутствующим в канале шумом $N(t)$:

$$z(t) = v(t) + N(t), \quad 0 \leq t \leq T, \quad (6.38)$$

а статистические свойства сигнала и шума описываются n -мерными плотностями вероятности:

$$w(\vec{x}) = w(x_1, x_2, \dots, x_n); \quad w(\vec{r}) = w(r_1, r_2, \dots, r_n).$$

Если канал имеет ограниченную ширину полосы пропускания F , то сигнал $z(t)$ и шум $N(t)$ можно в соответствии с теоремой Котельникова представить их выборочными значениями. Число выборочных значений

$$n = T/\Delta t = 2FT.$$

Пропускная способность непрерывного канала

$$C = \frac{1}{T} \max_{w(\vec{x})} I(\vec{v}, \vec{z}), \quad (6.39)$$

где $I(\vec{v}, \vec{z})$ — среднее количество информации о реализации сигнала $v(t)$ длительностью T , которое содержится в принятом сигнале $z(t)$ той же длительности.

Максимум количества информации ищут по всем возможным распределениям непрерывного сигнала $v(t)$.

Выражение (6.39) можно записать через дифференциальные энтропии

$$C = \frac{1}{T} \max_{w(\vec{x})} [h(\vec{z}) - h(\vec{z} / \vec{v})], \quad (6.40)$$

где

$$h(\vec{z}) = - \int \dots \int_{-\infty}^{\infty} w(\vec{y}) \log w(\vec{y}) d\vec{y},$$

$$h(\vec{z} / \vec{v}) = - \int \dots \int_{-\infty}^{\infty} w(\vec{y}, \vec{x}) \log w(\vec{y}, \vec{x}) d\vec{y} d\vec{x}.$$

Условная плотность вероятности $w(\vec{y}, \vec{x})$ равна плотности вероятности одного шума (см. разд. 5.2). Следовательно,

$$C = \frac{1}{T} \max_{w(\vec{x})} [h(\vec{z}) - h(\vec{N})].$$

Если выборочное значение смеси $z(t)$ и шума $N(t)$ статистически независимы, то энтропию смеси сигнала и шума можно записать в виде

$$\begin{aligned}
 h(\bar{z}) &= - \int_{-\infty}^{\infty} \int w(\bar{y}) \log w(\bar{y}) d\bar{y} = - \int_{-\infty}^{\infty} \int \prod_{i=1}^n w(y_i) \log \prod_{i=1}^n w(y_i) dy_1 dy_2 \dots dy_n = \\
 &= -2FT \int_{-\infty}^{\infty} w(y) \log w(y) dy = 2FT h(z),
 \end{aligned}$$

где $h(z)$ — энтропия выборки смеси $z(t)$.

Аналогично получаем

$$h(\bar{z} / \bar{v}) = h(\bar{N}) = 2FT h(N), \quad (6.42)$$

где $h(N)$ — энтропия выборки шума.

Выборочные значения статистически независимы, если: а) случайные реализации ограничены по спектру; б) их спектральные плотности в полосе F равномерны; в) случайные реализации подчиняются нормальному закону.

С учетом формул (6.41), (6.42) выражение для пропускной способности канала записывается в виде:

$$C = 2F \max_{w(x)} [h(z) - h(N)]. \quad (6.43)$$

Найдем максимальное значение среднего количества информации, варьируя плотностью распределения вероятности $w(x)$ выборки сигнала. Максимум стоящего в скобках выражения определяется максимальным значением энтропии $h(z)$. Если мощности (дисперсии) сигнала P_c и шума $P_{ш}$ ограничены, то для статистически независимых случайных функций $v(t)$ и $N(t)$ мощность смеси $z(t)$ также будет ограниченной величиной:

$$D\{z\} = D\{v\} + D\{N\} = P_c + P_{ш}.$$

Учитывая, что из множества возможных плотностей распределения вероятности случайной величины с ограниченной дисперсией наибольшую энтропию имеет плотность, подчиняющаяся нормальному закону, получим

$$\max h(\bar{z}) = 2FT \log \sqrt{2\pi e(P_c + P_{ш})}. \quad (6.44)$$

Пусть присутствующий в канале шум подчиняется нормальному закону, тогда дифференциальная энтропия выборки этого шума

$$h(N) = \log \sqrt{2\pi e P_{ш}}. \quad (6.45)$$

Варируя видом плотности распределения $w(x)$ с целью отыскания максимума энтропии $h(z)$, изменяем и вид плотности распределения вероятности смеси $z(t)$.

Известно, что, если отдельные (две и более) случайные величины распределены по этому закону, то их сумма также распределена по нормальному закону. Поэтому, до тех пор, пока плотность сигнала $w(x)$ не подчиняется нормальному закону,

плотность распределения смеси сигнала и шума $w(y)$ также не подчиняется этому закону, а энтропия $h(z) < h_{\max}(z)$. Когда плотность $w(x)$ подчиняется нормальному закону, тогда и смесь сигнала и шума будет ему подчиняться, а энтропия $h(z)$ достигает максимума и соответствует максимальному значению энтропии на выходе источника непрерывных сообщений. Подставляя формулы (6.44), (6.45) в (6.43), получим

$$C = F \log \frac{P_c + P_{\text{ш}}}{P_{\text{ш}}} = F \log \left(1 + \frac{P_c}{P_{\text{ш}}} \right). \quad (6.46)$$

Формулу (6.46) называют *формулой Шеннона*.

Пропускная способность непрерывного канала связи растет с ростом ширины полосы канала F и отношения сигнал/шум $P_c/P_{\text{ш}}$.

Одинаковые значения для C можно получить или изменением F при фиксированном отношении $P_c/P_{\text{ш}}$, или изменением $P_c/P_{\text{ш}}$ при фиксированном F . Это означает, что в непрерывном канале связи возможен обмен ширины полосы на отношение сигнал/шум и наоборот.

Пропускная способность непрерывного канала неограниченно возрастает с неограниченным ростом полосы F и стремится к предельному значению:

$$C \cong 1,443 \frac{P_c}{N_0}, \quad (6.47)$$

где $N_0 = P_{\text{ш}}/F$.

6.5. Основы теории помехоустойчивого кодирования

6.5.1. Теорема кодирования для канала с помехами

Теорема, известная в технической литературе как основная теорема кодирования К.Шеннона, применительно к дискретному источнику формулируется следующим образом: *если производительность источника сообщений $H'(A)$ меньше пропускной способности канала C , то существует по крайней мере одна процедура кодирования и декодирования, при которой вероятность ошибочного декодирования и ненадежность $H(A/B)$ могут быть сколь угодно малы. Если $H'(A) > C$, то такой процедуры не существует.*

Доказательство теоремы достаточно сложно. Его можно найти, например, в [2]. Ограничимся обсуждением результатов теоремы.

Положения, доказанные Шенноном, оказались неожиданными и расходились с признанными представлениями. В частности,

считалось, что для повышения верности нужно уменьшать скорость передачи сообщения. Это подтверждалось практическими результатами анализа многократной повторной передачи символов по каналу связи, из которых следовало, что весьма малую вероятность ошибки можно обеспечить только при стремлении скорости передачи к нулю.

Однако теорема Шеннона указывает на принципиальную возможность передачи информации со скоростью, сколь угодно близкой к пропускной способности канала и при сколь угодно малой вероятности ошибки.

Это очень важный для теории и практики результат.

К сожалению, теорема Шеннона не конструктивна в том смысле, что она не дает конкретного способа кодирования, существование которого доказывает. Однако, обосновав принципиальную возможность такого кодирования, она мобилизует усилия ученых на поиск высокоэффективных помехоустойчивых кодов.

6.5.2. Принципы построения помехоустойчивых кодов

Целью помехоустойчивого кодирования является *повышение верности передачи сообщения*, которая достигается введением избыточности в передаваемую информационную кодовую комбинацию. В дальнейшем будем рассматривать в основном коды с основанием 2, как наиболее просто реализуемые на практике. Такие коды называются *двоичными*. Рассмотрим передачу двоичной информации по каналу связи (рис.6.2). На вход кодирующего

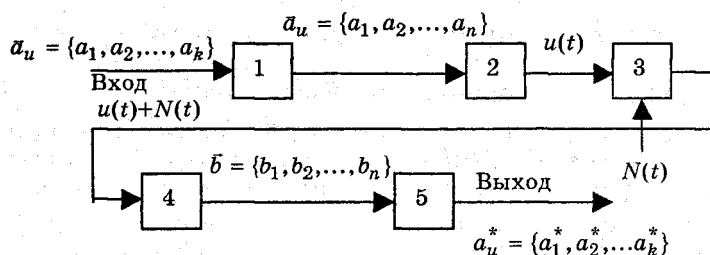


Рис.6.2. Передача помехоустойчивого кода по каналу связи:

- 1 — кодирующее устройство; 2 — модулятор; 3 — линия связи;
4 — демодулятор; 5 — декодирующее устройство

устройства поступают информационные кодовые комбинации $\bar{a}_n$ длиной в " k " символов. Их количество равно 2^k . В кодирующем устройстве производится формирование помехоустойчивого кода добавлением к информационным символам проверочных. В результате длина помехоустойчивого кода становится равной n . От-

сюда число проверочных символов равно $(n - k)$. Количество кодовых комбинаций на выходе кодирующего устройства при этом остается равным количеству кодовых комбинаций на его входе, т.е. 2^k . Эти кодовые комбинации $\bar{a}$ называются *разрешенными*. Разрешенные кодовые комбинации преобразуются модулятором в сигнал $u(t)$, который передается по линии связи. В демодуляторе происходит восстановление этих комбинаций. Восстановленные комбинации отличаются от переданных, так как в линии связи присутствует помеха $N(t)$. Возможны следующие случаи передачи (рис.6.3).

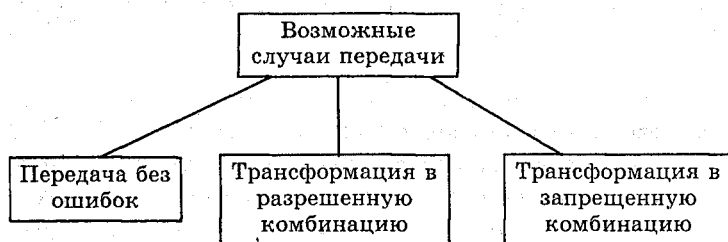


Рис.6.3. Случаи передачи помехоустойчивого кода по каналу связи

1. Кодовая комбинация на выходе линии связи $\bar{b}_j$ равна кодовой комбинации на ее входе $\bar{a}_i$. Количество возможных случаев приема равно 2^k . Эти случаи имеют место, когда помеха в линии связи отсутствует.

2. Кодовая комбинация на выходе линии связи $\bar{b}_j$ не равна кодовой комбинации на ее входе $\bar{a}_i$ но совпадает с одной из других разрешенных комбинаций. Число этих возможных случаев равно $2^k(2^k - 1)$.

3. Кодовая комбинация на выходе линии связи $\bar{b}_j$ не совпадает ни с одной разрешенной кодовой комбинацией на ее входе $\bar{a}_i$. Число этих возможных случаев равно $2^k(2^n - 2^k)$.

Случаи 2,3 имеют место тогда, когда в линии связи действует помеха и некоторые символы входной разрешенной кодовой комбинации принимаются с ошибками. Вместо "1" принимается "0" и наоборот. Входная разрешенная кодовая комбинация в результате этого переходит или в другую разрешенную комбинацию (случай 2) или в запрещенную (случай 3).

На приемном конце известны все разрешенные кодовые комбинации. С этими комбинациями можно сравнить принимаемые

кодовые комбинации и на основании этого определить: принадлежит ли принимаемая комбинация к разрешенной или запрещенной. Таким образом решается задача обнаружения ошибок. В системах с обратной связью эта задача является одной из главных. При ее решении принятая с ошибкой кодовая комбинация для дальнейшей обработки не передается, а на передатчик посылается сигнал запроса на повторение ранее переданной комбинации.

Напомним, что q -кратной ошибке при передаче кодовой комбинации длиной в n двоичных символов соответствует принятая кодовая комбинация, у которой на любых позициях число ошибочно принятых символов равно q . На остальных $n - q$ позициях символы принимаются верно.

Степень отличия переданной и принятой кодовых комбинаций характеризуется *вектором ошибки* (см. разд.5.2) и *кодovým расстоянием* между ними (*расстоянием Хэмминга*). Чтобы получить кодовое расстояние между двумя двоичными комбинациями, нужно подсчитать число единиц в поразрядной сумме этих комбинаций по модулю 2.

Например:

$$\begin{array}{r} 10110111 \\ + \quad 01001011 \\ \hline 11111100 \end{array}$$

Число единиц равно 6. Кодовое расстояние равно 6.

Кратность ошибки, с учетом определения кодового расстояния, есть не что иное, как кодовое расстояние между переданной и принятой кодовыми комбинациями или *вес вектора ошибки*, равный сумме единиц кодовой комбинации этого вектора.

Минимальное кодовое расстояние, взятое по всем парам разрешенных кодовых комбинаций, называется *минимальным кодовым расстоянием*. Обозначим его буквой d .

Обнаруживающая способность кода характеризуется следующей теоремой.

Если код имеет $d > 1$ и используется декодирование по методу обнаружения ошибок, то все ошибки кратностью $q < d$ обнаруживаются. Что же касается ошибок кратности $q \geq d$, то одни из них обнаруживаются, а другие нет.

Доказательство теоремы проведем от противного. Допустим, что ошибки кратностью $q < d$ не обнаруживаются. В этом случае принимаемые кодовые комбинации не выходят за пределы множества разрешенных комбинаций, что возможно при кратности ошибки $q \geq d$. Но, это противоречит исходному неравенству и

свидетельствует о справедливости теоремы. Следовательно, если $q < d$, принятая комбинация не будет разрешенной, а будет принадлежать к запрещенным комбинациям. Ошибка будет обнаружена. При $q \geq d$ принятая комбинация может оказаться разрешенной и ошибка окажется необнаруженной.

Условие обнаружения ошибок кратности q можно записать в виде

$$q \geq d + 1. \quad (6.48)$$

Используя формулы (5.17), (6.48), а также формулировку приведенной теоремы, найдем вероятность необнаруженной ошибки для двоичного симметричного дискретного канала без памяти, считая, что ошибки в приеме символов являются независимыми:

$$P_{\text{но}} \leq \sum_{q=d}^n c_n^q p^q (1-p)^{n-q}, \quad (6.49)$$

где p — вероятность ошибочного приема одного символа.

Из (6.48) следует, что ошибка обнаруживается при $q \leq d - 1$ и не обнаруживается при $q \geq d$. Это позволяет определить нижний предел суммы в выражении (6.49).

Рассмотрим процесс исправления ошибок в симметричном канале без памяти. При построении кода все множество принимаемых последовательностей длиной n разбивается на m непересекающихся подмножеств $M_1, M_2, \dots, M_m$, где m — число разрешенных кодовых комбинаций. При приеме комбинации ее сравнивают с комбинациями указанных подмножеств и таким образом определяют, к какому из подмножеств $M_1, M_2, \dots, M_m$ принадлежит принятая кодовая комбинация.

Очевидно, что при таком правиле декодирования будут исправляться все ошибки, если принимаемые кодовые комбинации не выводятся помехами за пределы указанных подмножеств.

При наличии помех в разделении множества принятых кодовых комбинаций возникают затруднения. Их можно снять, применяя *правило декодирования по методу максимального правдоподобия*.

Сущность метода заключается в следующем. Из рассмотрения алгоритма функционирования m -ичного дискретного симметричного канала без памяти следует, что разрешенная кодовая комбинация $\bar{a}_i$ может перейти в кодовую комбинацию $\bar{b}_j$ с вероятностью

$$P[\bar{b}_j / \bar{a}_i] = [p / (m - 1)]^{d(i,j)} (1 - p)^{n-d(i,j)}. \quad (6.50)$$

У принятой кодовой комбинации $\bar{b}_j[n - d(i, j)]$ символов приняты верно, а остальные $d(i, j)$ символов — с ошибками. Буквой $d(i, j)$ обозначено кодовое расстояние между комбинациями $\bar{a}_i$, $\bar{b}_j$. Вероятность $P[\bar{b}_j / \bar{a}_i]$ зависит от кодового расстояния $d(i, j)$ и тем больше, чем меньше это расстояние.

Допустим требуется принять решение о том какая кодовая комбинация $\bar{a}_i$ передавалась при приеме комбинации $\bar{b}_j$. Для этого нужно вычислить вероятности $P[\bar{b}_i / \bar{a}_j]$ и сравнить их между собой. Та кодовая комбинация $\bar{a}_i$, для которой эта вероятность максимальна, а кодовое расстояние минимально, будет считаться переданной.

Правило принятия решения по максимуму вероятности записывается в виде:

$$\max_i P[\bar{a}_i / \bar{b}_j]. \quad (6.51)$$

Его можно записать также через функцию правдоподобия $L(a_i)$

$$\max_i P[\bar{a}_i / \bar{b}_j] = \max_i L[\bar{a}_i] P[\bar{a}_i] / P[\bar{b}_j] = c L[\bar{a}_i], \quad (6.52)$$

где

$$L[\bar{a}_i] = P[\bar{b}_j / \bar{a}_i]; c — \text{постоянное число.}$$

Как правило полагают, что $P(\bar{a}_i) = 1/m$, а постоянное число находят из условия

$$\sum_{i=1}^m c L[\bar{a}_i] = 1.$$

Приведенный анализ позволяет сформулировать правило разбиения множества принятых кодовых комбинаций на m подмножеств.

В подмножество M_i следует включать все те комбинации $\bar{b}_j$, кодовое расстояние которых по отношению к кодовой комбинации $\bar{a}_i$ минимально. Декодирование по наименьшему кодовому расстоянию оптимально для симметричного дискретного канала. Для других каналов оно может оказаться не оптимальным.

Для двоичного симметричного канала без памяти справедлива следующая теорема, определяющая исправляющую способность кода.

Если код имеет $d > 2$ и используется декодирование с исправлением ошибок по наименьшему расстоянию, то все ошибки

кратностью $q < d/2$ исправляются. Что же касается ошибок большей кратности, то одни из них исправляются, а другие нет.

Для доказательства покажем, что переданная комбинация $\bar{a}_i$ ближе к принятой $\bar{b}_j$ нежели любая другая разрешенная комбинация. Предположим противное, т.е. что существует разрешенная комбинация $\bar{a}_k$ для которой $d(k, j) < d(i, j)$. Отсюда следует, что

$$d(k, i) \leq d(k, j) + d(i, j) < 2d(i, j). \quad (6.53)$$

Последнее неравенство написано на основе известного соотношения (рис. 6.4):

$$d(x, z) + d(yz) \geq d(xy). \quad (6.54)$$

Сумма двух сторон треугольника больше или по крайней мере равна третьей стороне.

Но, по условию теоремы $d(i, j) = q < d/2$. Отсюда $d(k, i) < d$, что противоречит определению d . Это противоречие свидетельствует о справедливости теоремы.

Условие исправления ошибок записывается в виде:

$$d \geq 2q + 1. \quad (6.55)$$

Вероятность ошибочного декодирования при коррекции ошибок записывается по аналогии с формулой (6.49) в виде:

$$P_{\text{од}} \leq \sum_{q=d/2}^n c_n^q p^q (1-p)^{n-q}. \quad (6.56)$$

Поясним процесс разбиения множества кодовых комбинаций. Пусть длина двоичных кодовых комбинаций $n = 3$, а минимальное кодовое расстояние $d = 1$. Разрешенные кодовые комбинации образуют следующее множество: 000, 001, 010, 011, 100, 101, 110, 111. Любая одиночная ошибка при приеме одной из этих кодовых комбинаций трансформирует ее в другую разрешенную комбинацию. Это случай безызбыточного кода, не обладающего обнаруживающей и корректирующей способностью.

Если $d = 2$, то ни одна из разрешенных комбинаций при одиночной ошибке не переходит в другую.

Например, разрешенные комбинации 000, 011, 101, 110, а запрещенные 001, 010, 100, 111. Этот код обнаруживает одиночные ошибки.

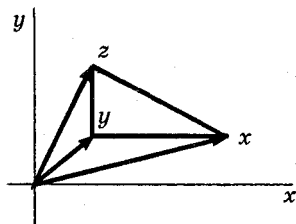


Рис. 6.4. К определению расстояния между тремя векторами x, y, z

Для их исправления нужно обеспечить непересекаемости подмножеств из запрещенных комбинаций. Для этого нужно обеспечить $d = 3$.

Разрешенные комбинации 000, 111. Запрещенные комбинации 001, 010, 100 образуют 1-е подмножество M_1 . Эти комбинации получаются в результате искажений символов разрешенной комбинации 000. Запрещенные комбинации 011, 101, 110 образуют 2-е подмножество M_2 . Они получаются при искажении символов разрешенной кодовой комбинации 111.

Отметим одно важное свойство, с которым мы столкнулись при рассмотрении данного примера. Оно заключается в следующем. С уменьшением числа разрешенных кодовых комбинаций кодовое расстояние между ними может увеличиваться, что позволяет сначала решить задачу обнаружения при $d = 2$, а затем и задачу исправления ошибок при $d = 3$.

Отмеченное свойство справедливо для любой длины кода n . Поэтому, если длина n задана, то можно получить любое число d не превышающее n , уменьшением числа разрешенных кодовых комбинаций m .

Из всевозможных задач помехоустойчивого кодирования главной является *задача поиска наилучшего (оптимального) кода*. Она формулируется следующим образом: при заданных m и n найти код, имеющий наибольшее d .

В общем виде эта задача не решена.

Рассмотренные принципы помехоустойчивого кодирования позволяют определить структурные схемы кодирующего и декодирующего устройств, исправляющие ошибки. В кодирующем устройстве обязательно имеется запоминающее устройство, в памяти которого хранятся разрешенные кодовые комбинации. Декодирующее устройство, кроме запоминающего, содержит также сравнивающее и решающее устройства. Последние предназначены для сравнения принятой кодовой комбинации со всеми разрешенными и нахождения такой комбинации, которая отличается от принятой меньшим числом символов. Однако подобные устройства не нашли применения из-за их сложности, особенно при большой длине кодов.

Ниже рассматриваются такие классы кодов, у которых кодирование и декодирование осуществляется относительно простыми техническими средствами даже при больших длинах кодов ($n > 100$). Это достигается использованием регулярных правил определяемых структурой кодов.

6.5.3. Классификация помехоустойчивых кодов

Помехоустойчивые коды делятся на два больших класса: *блочные* и *непрерывные*. В блочных процедура кодирования заключается в сопоставлении любой буквы алфавита дискретного источника (или последовательности из k символов, соответствующей этой букве), блоку из n символов. Длина блока может быть как постоянной, так и переменной. При постоянной длине коды называются *равномерными*. Коды с переменной длиной называются *неравномерными*. Если символы кодовой последовательности разделяются на информационные и проверочные, то такие блочные коды называются *разделимыми*. Разделимые коды длиной в n символов, из которых " k " символов являются информационными, обозначаются как (n, k) . У неразделимых кодов их символы нельзя разделить на информационные и проверочные. Разделимые коды делятся на *линейные* и *нелинейные*. *Линейными* называются такие коды, у которых поразрядная сумма по модулю 2 любых двух кодовых комбинаций (кодовых слов) также является комбинацией этого же кода.

Если первые k символов любой кодовой комбинации являются информационными, а остальные $n - k$ — проверочными, то такой линейный код называется *систематическим*. Примером простого линейного систематического кода является код $(n, n - 1)$, который содержит один проверочный символ. Этот код является кодом с проверкой на четность и обнаруживает ошибки нечетной кратности.

К линейным относятся *циклические* коды, используемые для обнаружения и исправления ошибок. Разновидностями циклических кодов являются: коды Хэмминга, Боуза-Чоудхури-Хоквингема (БЧХ-коды) и др.

К *нелинейным* относится код Бергера, проверочные символы которого представляют двоичную запись числа единиц в информационной кодовой комбинации.

В непрерывных кодах (в отличие от блочных) разбиение на блоки не производится. К ним относятся сверточные коды. Наконец, коды делятся на *исправляющие независимые ошибки* и *исправляющие пакеты ошибок*.

6.5.4. Линейные двоичные блочные коды

Основными характеристиками этих кодов являются: *длина* n , *число информационных символов* k , *число проверочных символов* $n - k$, *вес кодовой комбинации* (число единиц в ней), *число разрешенных кодовых комбинаций*, равное 2^k , *полное число кодовых*

комбинаций 2^n , кодовое расстояние, избыточность кода. Под избыточностью кода понимают величину

$$x = 1 - k/n,$$

где отношение k/n определяет скорость передачи информации при отсутствии помех. Величину k/n называют *относительной скоростью кода*.

Рассматриваемые блочные двоичные коды являются *линейными, разделимыми и систематическими*.

Проверочные символы кодовых комбинаций этого (n, k) — кода формируются по правилу:

$$a_{i\text{пр}} = \sum_{l=1}^k \gamma_{l,i} a_l, \quad (6.57)$$

где a_l , $l = 1, 2, \dots, k$ — информационные символы; $i = k + 1, \dots, n$; $\gamma_{l,i}$ — коэффициенты, равные 0 или 1. Коэффициент $\gamma_{l,i}$ равен 1, если информационный символ a_l участвует в образовании проверочного символа $a_{i\text{пр}}$. В противном случае $\gamma_{l,i} = 0$.

Суммирование в (6.57) осуществляется по модулю два.

Выражение (6.57) можно записать также в матричном виде

$$\bar{a}_{\text{пр}} = \bar{a}_n \gamma^T = \begin{bmatrix} a_1 & a_2 & \dots & a_k \end{bmatrix} \begin{bmatrix} \gamma_{1,k+1} \dots \gamma_{1n} \\ \gamma_{2,k+1} \dots \gamma_{2n} \\ \dots \\ \gamma_{k,k+1} \dots \gamma_{kn} \end{bmatrix}, \quad (6.58)$$

где $\bar{a}_{\text{пр}} = \{a_{i\text{пр}}\} = a_{k+1}, a_{k+2}, \dots, a_n$ — матрица строка из проверочных символов; $\bar{a}_n$ — матрица-строка заданной кодовой комбинации длиной в k символов; γ^T — транспонированная матрица коэффициентов γ_{li} .

Совокупность коэффициентов γ_{li} , можно также записать в виде матрицы размерности $(n - k) \times k$:

$$\gamma = \begin{bmatrix} \gamma_{1,k+1} \cdot \gamma_{2,k+1} \cdot \dots \cdot \gamma_{k,k+1} \\ \gamma_{1,k+2} \cdot \gamma_{2,k+2} \cdot \dots \cdot \gamma_{k,k+2} \\ \dots \\ \gamma_{1,n} \cdot \gamma_{2,n} \cdot \dots \cdot \gamma_{k,n} \end{bmatrix}. \quad (6.59)$$

Разрешенные кодовые комбинации линейных двоичных блочных кодов длиной в n символов строятся добавлением к каждой входной информационной кодовой комбинации длиной в k символов проверочных элементов длиной $n - k$ символов. Совокупность всех разрешенных кодовых комбинаций длиной n мож-

но записать в виде таблицы, содержащей 2^k строк. Для хранения этой таблицы потребуется память равная $n \cdot 2^k$. С ростом n объем памяти возрастает. Для линейного блочного кода объем памяти можно уменьшить, так как этот код можно получить линейным суммированием по модулю 2 любых k линейно независимых кодовых комбинаций. Матрица, содержащая эти кодовые комбинации, называется *порождающей*. Размерность порождающей матрицы равна $k \cdot n$ двоичных символов и это число определит необходимый объем памяти запоминающего устройства.

Порождающая матрица записывается в виде

$$G = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & \gamma_{1,k+1}, \dots, \gamma_{1,n} \\ 0 & 1 & 0 & \dots & 0 & \gamma_{2,k+1}, \dots, \gamma_{2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \gamma_{k,k+1}, \dots, \gamma_{k,n} \end{bmatrix} = [1_k, \gamma^T], \quad (6.60)$$

где 1_k — единичная матрица порядка k ; γ — матрица из коэффициентов γ_{ij} .

Каждая строка порождающей матрицы является кодовой комбинацией линейного блочного кода, причем в качестве входных информационных кодовых комбинаций длиной в k символов использованы только такие кодовые комбинации, которые содержат лишь один не нулевой символ.

Линейный двоичный блочный код можно задать также с помощью проверочной матрицы:

$$H = \begin{bmatrix} \gamma_{1,k+1} & \gamma_{2,k+1} & \dots & \gamma_{k,k+1} & 1 & 0 & \dots & 0 \\ \gamma_{1,k+2} & \gamma_{2,k+2} & \dots & \gamma_{k,k+2} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{1,n} & \gamma_{2,n} & \dots & \gamma_{k,n} & 0 & 0 & \dots & 1 \end{bmatrix} = [\gamma 1_{n-k}], \quad (6.61)$$

где 1_{n-k} — единичная матрица порядка $n - k$.

Единицы в любой j -й строке подматрицы γ , входящей в проверочную матрицу, указывает, какие информационные символы участвуют в формировании j -го проверочного символа.

Для каждой кодовой комбинации $\bar{a}$ ($n - k$)-кода произведение

$$\bar{a} H^T = 0, \quad (6.62)$$

где H^T — транспонированная проверочная матрица.

Процесс декодирования сводится к выполнению операции:

$$\bar{S} = \bar{b} H^T, \quad (6.63)$$

где $\bar{S}$ — вектор размерностью $(n - k)$, называемый *синдромом*, $\bar{b}$ — матрица строка из принятых кодовых символов длиной n .

Если принятая кодовая комбинация $\bar{b}$ совпадает с одной из разрешенных $\bar{a}$, то $\bar{S} = 0$. В противном случае $\bar{S} \neq 0$, причем вид синдрома зависит от вектора ошибки. Это позволяет указать разряд, где эта ошибка имеется. Действительно,

$$\bar{S} = \bar{b}H^T = (\bar{a} + \bar{e})H^T = \bar{e}H^T, \quad (6.64)$$

где $\bar{e}$ — вектор ошибки.

Число различных синдромов, соответствующих различным ошибкам, равно $2^{n-k} - 1$.

Пример 6.2. Матрица, определяющая проверочные символы (7,4) — кода равна

$$\gamma = \begin{vmatrix} \gamma_{15} = 0 & \gamma_{25} = 1 & \gamma_{35} = 1 & \gamma_{45} = 1 \\ \gamma_{16} = 1 & \gamma_{26} = 0 & \gamma_{36} = 1 & \gamma_{46} = 1 \\ \gamma_{17} = 1 & \gamma_{27} = 1 & \gamma_{37} = 0 & \gamma_{47} = 1 \end{vmatrix}$$

Найти: а) уравнения, определяющие проверочные символы кодовых комбинаций $\bar{a}_{\text{пр}}$; б) уравнения, определяющие символы синдрома $\bar{S}$; в) значения символов синдрома, когда имеется одиночная ошибка в информационных разрядах; г) структурные схемы кодера и декодера.

Решение. Проверочные символы находятся по формуле (6.58):

$$\bar{a}_{\text{пр}} = \bar{a}_n \gamma^T = (a_1 a_2 a_3 a_4) \begin{vmatrix} \gamma_{15} = 0 & \gamma_{16} = 1 & \gamma_{17} = 1 \\ \gamma_{25} = 1 & \gamma_{26} = 0 & \gamma_{27} = 1 \\ \gamma_{35} = 1 & \gamma_{36} = 1 & \gamma_{37} = 0 \\ \gamma_{45} = 1 & \gamma_{46} = 1 & \gamma_{47} = 1 \end{vmatrix} = (a_5 = a_1 \cdot 0 + a_2 \cdot 1 +$$

$$+ a_3 \cdot 1 + a_4 \cdot 1; a_6 = a_1 \cdot 1 + a_2 \cdot 0 + a_3 \cdot 1 + a_4 \cdot 1;$$

$$a_7 = a_1 \cdot 1 + a_2 \cdot 1 + a_3 \cdot 0 + a_4 \cdot 1) = (a_5 = a_2 + a_3 + a_4;$$

$$a_6 = a_1 + a_3 + a_4; a_7 = a_1 + a_2 + a_4).$$

Проверочная матрица H находится добавлением к матрице γ дополнительных элементов из 0 и 1.

$$H = \begin{vmatrix} \gamma_{15} & \gamma_{25} & \gamma_{35} & \gamma_{45} & 1 & 0 & 0 \\ \gamma_{16} & \gamma_{26} & \gamma_{36} & \gamma_{46} & 0 & 1 & 0 \\ \gamma_{17} & \gamma_{27} & \gamma_{37} & \gamma_{47} & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}$$

Условия, определяющие символы синдрома, найдем из (6.63):

$$\bar{S} = \bar{B}H^T = (b_1 b_2 b_3 b_4 b_5 b_6 b_7) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (b_2 + b_3 + b_4 + b_5;$$

$$b_1 + b_3 + b_4 + b_6; b_1 + b_2 + b_4 + b_7). \quad (6.66)$$

Если ошибочно принят первый символ кодовой комбинации, то

$$b_1 = \bar{a}_1, \quad b_2 = a_2, \quad b_3 = a_3, \quad b_4 = a_4, \quad b_5 = a_5, \quad b_6 = a_6, \quad b_7 = a_7.$$

Подставляя эти значения в формулу (6.66) получим

$$\bar{S} = (a_2 + a_3 + a_4 + a_5; \bar{a}_1 + a_3 + a_4 + a_6; \bar{a}_1 + a_2 + a_4 + a_7) = (011).$$

При вычислении символов синдрома приняты во внимание уравнения (6.65) для проверочных символов a_5 , a_6 и a_7 , а также то, что $\bar{a}_1 + a_1 = 1$.

Аналогично вычисляются остальные синдромы. Значения синдромов приведены в табл. 6.2.

Таблица 6.2

Разряды кодовой комбинации						
1-й	2-й	3-й	4-й	5-й	6-й	7-й
011	101	110	111	100	010	001

Реализация блочных кодов возможна с применением ЭВМ и специализированных вычислительных устройств (кодеров и декодеров).

В случае применения ЭВМ вычисление проверочных символов в передатчике ведут по формуле (6.58), а вычисление синдрома в приемнике — по формуле (6.63). Для вычислений в памяти ЭВМ должны храниться строки соответствующих матриц. Например, для хранения порождающей матрицы (6.60) требуемый объем памяти равен $k \cdot n$ бит.

В случае применения специализированных вычислительных устройств (кодеров и декодеров) необходимость в хранении матриц отпадает.

Например, при реализации (7,4) кода с помощью ЭВМ для хранения матрицы, определяющей проверочные символы, требуемый объем памяти равен $3 \cdot 4 = 12$ бит, а для хранения проверочной матрицы объем памяти равен $3 \cdot 7 = 21$ бит (см. пример 6.2).

При реализации кодера (7,4) — кода проверочные символы находятся суммированием по модулю 2 соответствующих элементов (символов) информационного слова. Например, пятый проверочный символ находится суммированием второго, третьего и четвертого элементов информационного слова, как показано в (6.65).

При реализации декодера (7,4) кода символы синдрома также находятся суммированием по модулю 2 соответствующих элементов принятого кодового слова в соответствии (6.66).

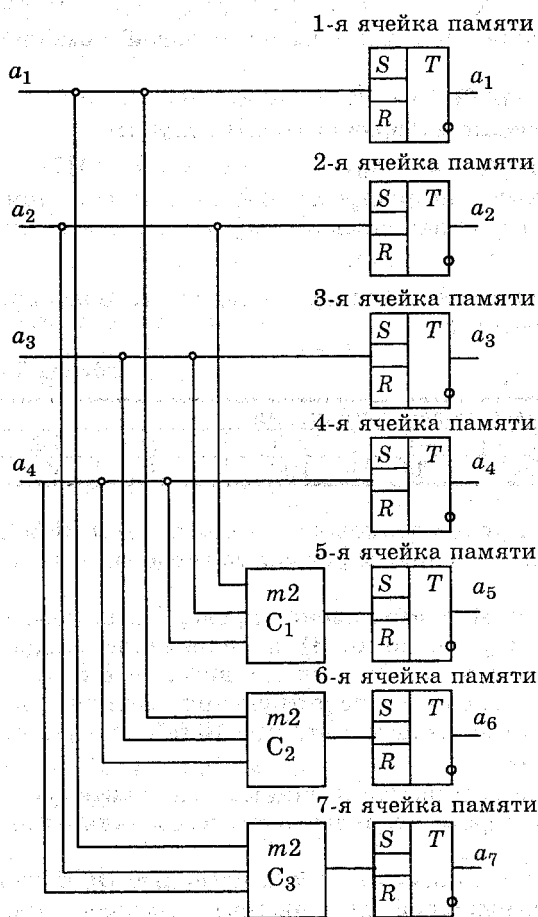


Рис.6.5. Структурная схема кодирующего устройства: T — ячейка памяти; C — сумматор по модулю два

В кодере номера суммируемых элементов кодовых слов определяются только теми символами вертикальных столбцов транспонированной матрицы (6.59), которые равны 1.

Аналогично в декодере номера суммируемых элементов определяются только теми элементами вертикальных столбцов транспонированной матрицы (6.61), которые равны 1.

Поэтому соединения входов сумматоров по модулю 2 с соответствующими разрядами кодовых слов адекватно отражают состояния ячеек памяти, отведенных в запоминающем устройстве ЭВМ для хранения матриц γ и H .

Структурные схемы кодера и декодера приведены на рис.6.5 и рис.6.6.

Кодирующее устройство работает следующим образом. При поступлении импульса синхронизации подлежащая кодированию

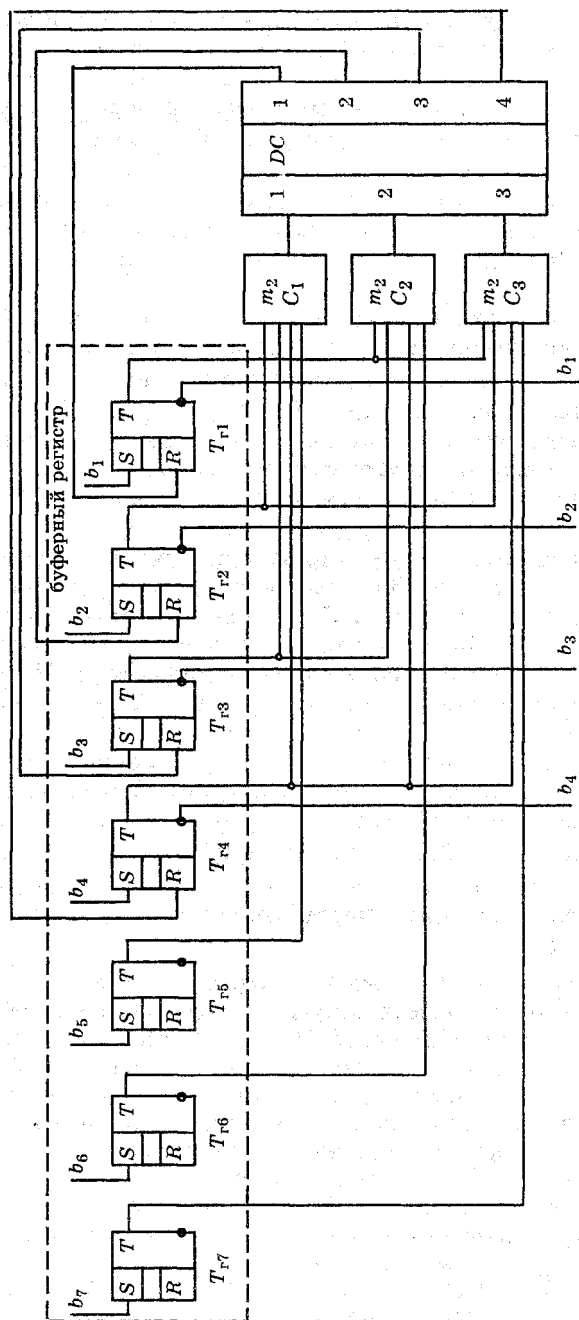


Рис.6.6. Структурная схема декодирующего устройства: T — триггерная ячейка буферного регистра; C — сумматор по модулю 2; DC — дешифратор

четырёхразрядная кодовая комбинация перезаписывается, например, с аналого-цифрового преобразователя в первые четыре ячейки буферной памяти, выполненной на семиразрядном регистре. На выходе сумматора C_1 формируется проверочный элемент a_5 , на выходе сумматора C_2 — проверочный элемент a_6 и на выходе сумматора C_3 — проверочный элемент a_7 . Эти элементы записываются соответственно в пятой, шестой и седьмой ячейках памяти. Сформированная таким образом кодовая комбинация $\{a_i\}$ считывается в канал связи или в параллельном и последовательном коде.

Декодирующее устройство работает следующим образом. Принятая кодовая комбинация поступает на семиразрядный буферный регистр $T_1 - T_7$. Если эта комбинация не содержит ошибок, на всех четырех выходах дешифратора ошибок DC сигналы коррекции отсутствуют. Если принятая кодовая комбинация содержит ошибку в первом информационном символе, на первом выходе дешифратора вырабатывается сигнал коррекции, который воздействует на первый триггер T_1 буферного регистра. Последний изменяет свое состояние на противоположное и ошибка таким образом исправляется. На остальных выходах дешифратора сигналы коррекции отсутствуют. Если ошибка имеется при приеме второго символа кодовой последовательности, то сигнал коррекции присутствует на втором выходе дешифратора. На остальных выходах сигналы коррекции равны нулю. Сигнал коррекции с выхода второго выхода дешифратора, воздействуя на T_2 изменяет его состояние на противоположное. При ошибочном приеме третьего и четвертого символов схема на рис.6.6 работает аналогично. После коррекции состояния буферного регистра, его содержимое считывается с $T_1 - T_4$ для дальнейшей обработки.

6.5.5. Понятие о циклических, сверточных и каскадных кодах

Циклические коды относятся к подклассу линейных двоичных блочных систематических кодов. Они являются разделимыми и обозначаются как (n, k) коды. Название этих кодов происходит от их основного свойства: если исходная кодовая комбинация принадлежит циклическому коду, то все другие кодовые комбинации, полученные циклической перестановкой элементов исходной комбинации, также принадлежат этому коду.

При исследовании свойств циклических двоичных кодов каждая кодовая комбинация может быть записана в виде многочлена (полинома) степени $n - 1$:

$$b(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \dots + b_1x + b_0,$$

где $b_0, b_1, \dots, b_{n-1}$ — кодовая комбинация.

Например, кодовой комбинации 11011 соответствует полином

$$1 \cdot x^4 + 1 \cdot x^3 + 0 \cdot x^2 + 1 \cdot x + 1 \cdot x^0 = x^4 + x^3 + x + 1.$$

Операции над полиномами ведут с использованием следующих правил.

1. Сложение двух полиномов производится сложением по модулю 2 коэффициентов при одинаковых степенях этих полиномов:

$$a(x) + b(x) = (a_{n-1} + b_{n-1})x^{n-1} + (a_{n-2} + b_{n-2})x^{n-2} + \dots + (a_1 + b_1)x + a_0 + b_0 = c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \dots + c_1x + c_0.$$

2. Умножение полиномов производится по модулю некоторого полинома $p(x)$. Это означает, что в качестве результата умножения принимается остаток от деления обычного произведения полиномов на полином $p(x)$. Остаток от деления называется *вычетом по многочлену $p(x)$* . Указанные операции над полиномами не приводят к кодовым комбинациям, длина которых больше длины n исходного кода, т.е. являются замкнутыми.

Многочлен $p(x)$ называется *порождающим*. Он удовлетворяет следующим условиям.

1. Полином не делится на любой другой полином. Поэтому он называется также *неприводимым*.

2. Двучлен вида $x^n + 1$ делится на $p(x)$ без остатка.

Кодовые комбинации циклического кода имеют следующие свойства.

1. Полиномы, соответствующие разрешенным кодовым комбинациям циклического кода, имеют нулевой вычет по модулю $p(x)$, т.е. делятся на этот полином без остатка.

2. Если разрешенные кодовые комбинации принадлежат циклическому коду, то их линейная комбинация будет также принадлежать этому коду, т.е. делиться без остатка на порождающий полином $p(x)$.

Свойство 1 позволяет относительно простыми средствами решить задачу обнаружения и исправления ошибок.

Обнаружение ошибок в циклическом коде производится делением принятой кодовой комбинации на полином $p(x)$. Остаток от деления является синдромом. Если синдром равен нулю, принятая комбинация не содержит ошибок и подвергается дальнейшей обработке. Если синдром не равен нулю, то некоторые символы кодовой комбинации принимаются неверно и принятая кодовая комбинация стирается.

Возможность исправления ошибок в циклических кодах обусловлена тем, что вид синдрома зависит от номера искаженного символа принимаемой кодовой комбинации. Среди множества

порождающих полиномов существуют так называемые примитивные полиномы, степень которых $r = n - k$ связана с длиной n кодовой комбинации зависимостью $n = 2^r - 1$. Такие примитивные полиномы обеспечивают число синдромов, равное n .

Формирование разрешенных кодовых комбинаций циклического кода осуществляется по следующему алгоритму.

1. Многочлен $a(x)$, представляющий информативную часть кодовой комбинации, умножается на x^{n-k} . В результате получается многочлен $a(x)x^{n-k}$. Это n -элементная кодовая комбинация, у которой $n - k$ младшие разряды — нулевые.

2. Многочлен $a(x)x^{n-k}$ делится на примитивный полином $p(x)$. В результате определяется остаток от деления $a(x)x^{n-k} = m(x)p(x) + c(x)$, где $m(x)$ — частное; $c(x)$ — остаток, являющийся вычетом.

3. Многочлен $a(x)x^{n-k}$ складывается с остатком. В результате получается полином $b(x)$, соответствующий n -элементной кодовой комбинации. У этой кодовой комбинации первые k элементы информационные, остальные $n - k$ — проверочные. Проверочные элементы определяются остатком $c(x)$. Полином

$$b(x) = a(x)x^{n-k} + c(x).$$

На рис.6.7 приведена схема кодера циклического (7,4) — кода. У этого кода 4 символа информационные и 3 — проверочные. Порождающий полином равен

$$p(x) = x^3 + x^2 + 1.$$

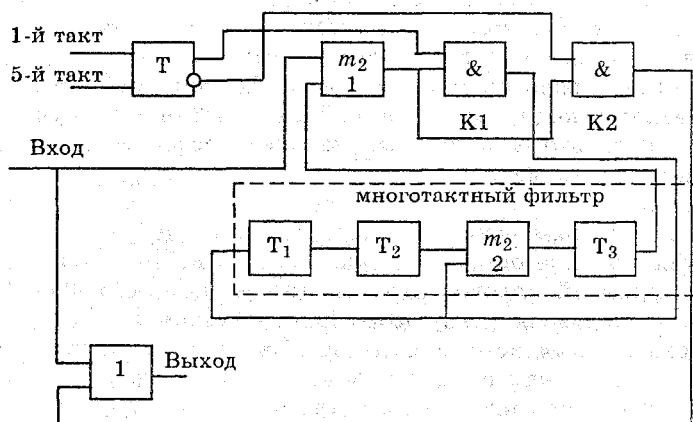


Рис.6.7. Схема кодера циклического (7,4) — кода:

Т — триггер; m_2 — сумматор по модулю два; & — схема И (ключ);
1 — схема ИЛИ

Многочлен, определяющий информационную часть $a(x) = x^3 + 1$. Этому многочлену соответствует кодовая комбинация 1001.

Умножим многочлен $a(x)$ на x^3 . Получим

$$a(x)x^3 = x^6 + x^3.$$

Этому многочлену соответствует кодовая комбинация длиной в 7 символов, у которой первые 4 символа информационные, а остальные 3 символа нулевые: 1001000.

На последних трех позициях будут расположены проверочные элементы.

Кодирующее устройство (рис.6.7) содержит многотактный фильтр, состоящий из трех триггерных ячеек. Число ячеек равно степени порождающего полинома $p(x)$. В многотактном фильтре содержатся также сумматоры по модулю 2. Их число равно числу ненулевых членов порождающего полинома минус 1, т.е. $3 - 1 = 2$ в нашем примере. Сумматоры устанавливаются после ячеек, соответствующих ненулевым членам (после x^2). После члена x^3 сумматор не ставится. Соответствующий члену $x^0 = 1$ сумматор ставится перед первой ячейкой.

Цикл работы кодера состоит из двух этапов и составляет $7\tau_0$, где τ_0 — единичный интервал. Первый этап длится $4\tau_0$. От 1-го тактового импульса триггер T занимает положение, при котором на его прямом выходе появится сигнал "1", а на инверсном — "0". Ключ K_1 откроется, а ключ K_2 закроется. В этом состоянии триггер и ключи находятся в течение времени, равном $4\tau_0$. За это время на выход кодера поступает 4 единичных символа информационной группы. Одновременно через первый — сумматор по модулю 2 и ключ K_1 информационные элементы поступают на устройство деления многочлена $a(x)$ на $p(x)$. Деление осуществляется многотактным фильтром. Процесс формирования кодовой комбинации шаг за шагом представлен в табл.6.3.

Таблица 6.3

Номер такта	Вход	Состояния ячеек регистра			Выход
		1	2	3	
1	1	1	0	1	1
2	0	1	1	1	01
3	0	1	1	0	001
4	1	1	1	0	1001
5	0	—	1	1	01001
6	0	—	—	1	101001
7	0	—	—	—	1101001

При поступлении пятого тактового импульса триггер T изменит свое состояние: на инверсном выходе появится 1, а на прямом — 0. Ключ K_1 закроется, а ключ K_2 откроется. За остальные 3 такта элементы остатка от деления (проверочные символы) через сумматор поступают на выход кодера.

Декодирующее устройство (рис.6.8) для кодов, обнаруживающих ошибки, по существу ничем не отличается от кодирующего (см. рис.6.7). В нем добавляется лишь буферный регистр для хранения информационных символов принимаемой кодовой комбинации в течение времени, пока эта комбинация делится на полином $p(x)$ с помощью декодирующего регистра. Если остаток обнаружен, то информация в буферном регистре уничтожается. Состояния ячеек декодирующего регистра приведено в табл.6.4 при отсутствии ошибок и в табл.6.5 — при ошибке в первом символе кодовой последовательности.

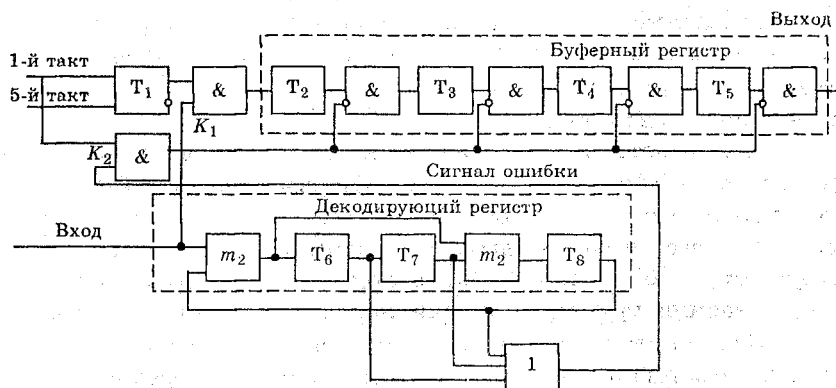


Рис.6.8. Схема декодера циклического (7,4) — кода: T — триггер; m_2 — сумматор по модулю два; $\&$ — схема И (ключ); 1 — схема ИЛИ

Таблица 6.4

Номер такта	Вход	Состояния ячеек многотактного фильтра			Вход буферного регистра
		1	2	3	
1	1	1	0	1	1
2	0	1	1	1	0
3	0	1	1	0	0
4	1	1	1	0	1
5	0	0	1	1	0
6	1	0	0	1	0
7	1	0	0	0	0

Номер такта	Вход	Состояния ячеек многотактного фильтра			Вход буферного регистра
		1	2	3	
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	1	1	0	1	1
5	0	1	1	1	0
6	1	0	1	1	0
7	1	0	0	1	0

В декодирующих устройствах, исправляющих ошибки, определяется номер искаженного символа принятой кодовой последовательности.

Известен метод определения номера символа, основанный на следующем свойстве: остаток от деления принятого многочлена $h(x)$ на $p(x)$ равен остатку от деления многочлена ошибок $e(x)$ на тот же многочлен $p(x)$.

Многочлен $e(x) = b(x) + h(x)$. Эти остатки не зависят от передаваемой информации, а определяются только номерами тех символов, где имеются ошибки. Ошибки зависят от полинома $p(x)$ и номера символа.

Анализируя синдром, можно определить место ошибки. Сначала определяется остаток, соответствующий ошибке в старшем разряде кодовой комбинации $b(x)$. Для этого $h(x)$ делится на полином $p(x)$. Полученный остаток сравнивается с вычисленным, который получается делением многочлена ошибки старшего разряда $e_1(x)$ на полином $p(x)$. Если остатки равны, то ошибка в старшем разряде, а если нет, то проверке подвергается соседний младший разряд. Для этого принятый многочлен $h(x)$ умножается на x , затем находится остаток от деления $xh(x)$ на полином $p(x)$. Полученный остаток сравнивается с остатком, который получается делением многочлена ошибки $e_1(x)$ на полином $p(x)$. Если остатки равны, процесс определения номера разряда (символа) заканчивается, а если нет, то проверке подвергается следующий младший разряд. Для этого принятый полином $h(x)$ умножается на x^2 , затем находится остаток от деления $x^2h(x)$ на полином $p(x)$ и т.д. Алгоритм вычисления продолжается до тех пор, пока вычисляемый остаток не будет равен $e_1(x)$.

Пример 6.3. Принятый код (7,4) равен 1101011. Прimitивный полином $p(x) = x^3 + x^2 + 1$. Определить ошибочно принятый элемент.

$$\begin{array}{r}
 x^6 + x^4 + x^2 + x + 1 \quad | \quad x^3 + x^2 + 1 \\
 \hline
 x^6 + x^5 + x^3 \\
 \hline
 x^5 + x^4 + x^3 + x^2 + x + 1 \\
 x^5 + x^4 + \quad \quad x^2 \\
 \hline
 \quad \quad x^3 + x + 1 \\
 \quad \quad x^3 + x^2 + 1 \\
 \hline
 \quad \quad \quad x^2 + x
 \end{array}$$

Полученный остаток $x^2 + x$ совпадает с $e_1(x)$. Следовательно, ошибка во втором разряде. Принятая кодовая комбинация после исправления ошибки равна 1001011.

Среди других алгоритмов декодирования циклических кодов получили распространение методы, основанные на анализе веса остатка. Обоснование этих алгоритмов можно найти в [33].

Подклассом циклических кодов являются коды Боуза-Чоудхури-Хоквингема (БЧХ). Они отличаются относительно простой реализацией алгоритма декодирования. Доказано, что для любой пары натуральных чисел s и $q < 2^{s-2}$ можно построить двоичный код БЧХ с $n = 2^s - 1$ и $k > n - sq$, исправляющий любую конфигурацию ошибок с кратностью, меньшей или равной q .

Относительно простая реализация декодирующих устройств получается и при *мажоритарном* декодировании некоторых циклических кодов. Мажоритарное декодирование основано на том, что в этих кодах каждый информационный символ определяется несколькими линейными выражениями через другие символы кодовой последовательности.

При отсутствии помех все линейные соотношения дают одинаковые выражения. При наличии помех получаются разные результаты (одни дают результат "0", а другие — "1"). В этом случае окончательное решение о том, какое значение имеет информационный символ, принимается по принципу большинства. Если большее число нулей, принимается решение, что принимаемый информационный символ равен "0". Если большее число "1" — принимается решение о том, что принятый информационный символ равен 1. Такой принцип принятия решения называется *мажоритарным*.

Проиллюстрируем принцип мажоритарного декодирования на примере кода (7,3) с порождающим полиномом $p(x) = x^4 + x^3 + x^2 + 1$ и проверочной матрицей

$$H = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Используя соотношение

$$\bar{a}H^T = 0,$$

найдем следующие выражения для символов a_1 , a_2 и a_3 :

$$a_1 = a_3 + a_4 = a_2 + a_3 + a_5 = a_2 + a_6;$$

$$a_2 = a_1 + a_3 + a_5 = a_1 + a_6 = a_3 + a_7;$$

$$a_3 = a_1 + a_4 = a_1 + a_2 + a_5 = a_2 + a_7.$$

В декодере имеется возможность разными способами вычислить эти символы. Например, первый информационный символ b_1^* определяется следующими уравнениями:

$$b_1^* = b_1; \quad b_1^* = b_3 + b_4; \quad b_1^* = b_2 + b_3 + b_5; \quad b_1^* = b_2 + b_6. \quad (6.67)$$

Они получены из равенства

$$\bar{b}H^T = 0,$$

где $\bar{b} = b_1, b_2, \dots, b_7$ — принятая кодовая комбинация.

Аналогично находятся уравнения для b_2^* и b_3^* .

При отсутствии ошибок все четыре уравнения (6.67) дают одинаковые значения для b_1^* .

Если ошибка происходит в первом разряде, то $b_1^* = \bar{a}_1$, и уравнения записываются в виде:

$$b_1^* = \bar{a}_1, \quad b_1^* = a_3 + a_4, \quad b_1^* = a_2 + a_3 + a_5, \quad b_1^* = a_2 + a_6.$$

Оценки символа, полученные из второго, третьего и четвертого уравнения совпадают и отличаются от оценки, полученной на основе первого уравнения. По большинству принимается решение $b_1^* = a_1$. Аналогично принимаются решения если ошибки происходят во втором и третьем информационных разрядах.

Схема декодера (рис.6.9) довольно проста. Она содержит сдвигающий регистр, схему ИЛИ, три сумматора по модулю 2 и мажоритарный элемент.

После заполнения регистра ключ K размыкается и вычисляется первый кодовый символ a_1 . Далее, путем циклических сдвигов вычисляются второй и третий и кодовые символы a_2 , a_3 .

Кроме блочных, в системах связи применяются *рекуррентные* или *цепные* коды. Информационные символы в этих кодах

чередуются с проверочными. Примером рекуррентного является код

$$a_1 a_{12} a_2 a_{23} a_3 a_{34} a_4 a_{45} a_5,$$

где a_i — i -й информационный символ, равный 0 или 1, а $a_{i,i+1}$ проверочный символ, определяемый уравнением

$$a_{i,i+1} = a_i + a_{i+1}.$$

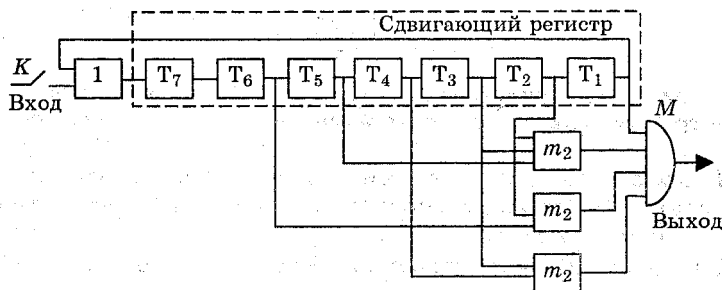


Рис.6.9. Структурная схема декодера циклического мажоритарного (7,3) — кода: Т — триггерная ячейка; m_2 — сумматор по модулю 2, М — мажоритарный элемент

Рекуррентный код обозначают (k/n) , где k — число информационных, а n — общее число символов.

К рекуррентным относятся и *относительный* код, используемый при ОФМ. Другим примером является *сверточный* код.

Сверточный код — это линейный, рекуррентный код. Он формируется следующим образом. В каждый i -й тактовый интервал на вход кодирующего устройства поступает k символов сообщения $a_{i1}a_{i2}...a_{ik}$. Выходные символы $b_{i1}b_{i2}...b_{in}$ формируются с помощью рекуррентного соотношения из k символов сообщения, поступивших в данный и предшествующие тактовые интервалы времени.

Рассмотрим формирование сверточного кода на конкретном примере. Кодирующее устройство сверточного кода содержит сдвигающий регистр и сумматоры по модулю 2. В схеме (рис.6.10) на каждый символ a_i входного сообщения вырабатываются два символа выходной последовательности, которые через ключ последовательно во времени подаются в канал. Выходные символы являются линейными функциями комбинации, записанной в разрядах регистра.

Для сверточных кодов разработаны специальные алгоритмы так называемого последовательного декодирования, позволяющие исправлять ошибки относительно небольшой кратности. Используется также алгоритм Витерби. Описание методов декодирования сверточных кодов приведено в [34].

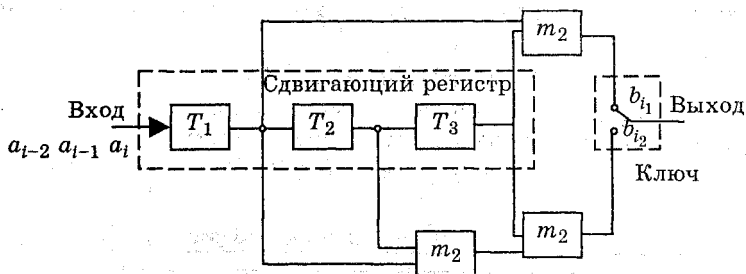


Рис.6.10. Структурная схема кодера сверточного кода:
Т — триггер сдвигового регистра; m_2 — сумматор по модулю два

Коды с длинными блоками и большим кодовым расстоянием получают объединением коротких кодов. К их числу относятся *итеративные* и *каскадные*.

Итеративный код строится следующим образом. Вначале сообщение кодируется (n_1, k_1) кодом первой ступени. При этом любая кодовая комбинация из числа k_2 комбинаций строится по известному правилу: к k_1 информационным символам приписываются $n_1 - k_1$ проверочные символы. Полученные таким образом комбинации записываются одна под другой в виде матрицы из k_2 строк. Далее сообщение кодируется (n_2, k_2) — кодом второй ступени. Каждая комбинация этого кода содержит k_2 информационных символов, совпадающих с символами одного из столбцов матрицы, и $n_2 - k_2$ проверочных символов, которые приписываются к этим информационным символам.

Результирующая матрица, приведенная в табл.6.6, содержит $n_1 n_2$ символов, из которых $k_1 k_2$ являются информационными.

Таблица 6.6

Символы	Информационные 1-й ступени	Проверочные 1-й ступени
Информационные 2-й ступени	$a_{11} \ a_{12} \ \dots \ a_{1k_1}$	$a_{1,k_1+1} \ \dots \ a_{1,n_1}$
	$a_{21} \ a_{22} \ \dots \ a_{2k_1}$	$a_{2,k_1+1} \ \dots \ a_{2,n_1}$
		
	$a_{k_2,1} \ a_{k_2,2} \ \dots \ a_{k_2,k_1}$	$a_{k_2,k_2+1} \ \dots \ a_{k_2,n_1}$
Проверочные 2-й ступени	$a_{k_2+1,1} \ a_{k_2+1,2} \ \dots$	$a_{k_2+1,k_1+1} \ \dots \ a_{k_2+1,n_1}$
		
	$a_{n_2,1} \ a_{n_2,2} \ \dots$	$a_{n_2,k_1+1} \ \dots \ a_{n_2,n_1}$

Декодирование производится в 2 этапа. На первом обнаруживают и исправляют ошибки принятых (n_1, k_1) кодовых комбинаций. На втором вновь исправляют ошибки, но уже (n_2, k_2) кода. На первом этапе обнаружение и исправление ошибок ведут по строкам табл.6.6, а на втором этапе — по столбцам.

Минимальное кодовое расстояние итеративного кода $d = d_1 d_2$, где d_1, d_2 — минимальные кодовые расстояния для кодов первой и второй ступеней.

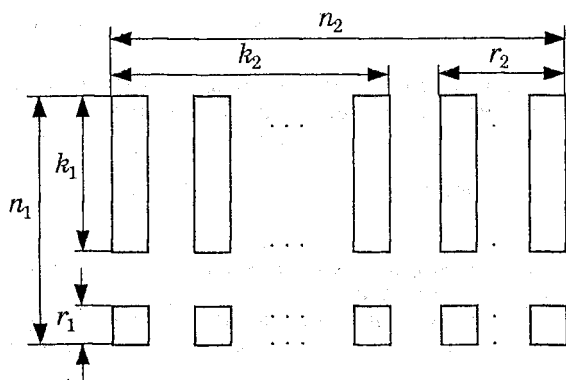


Рис.6.11. К процедуре кодирования каскадным словом

При использовании каскадного кода множество $k_1 k_2$ информационных символов разбивается на k_2 подблоков по k_1 символов (рис.6.11). Каждый подблок из k_1 символов рассматривается как символ из алфавита объемом 2^{k_1} . Затем k_2 подблоков кодируются кодовыми комбинациями внешнего вида длиной n_2 . Каж-

дый из n_2 символов кода второй ступени представляется, как вектор длины k_1 , и кодируется кодом $(n_1 k_1)$ первой ступени. В результате получается n_2 кодовая комбинация длиной $n = n_1 n_2$ каскадного кода.

В качестве внешнего кода часто применяют код Рида-Соломона с основанием 2^{k_1} , а в качестве внутреннего — двоичный $(n_1 k_1)$ — код.

Декодирование осуществляется следующим образом. Вначале декодируется внутренний код. При этом получается n_2 подблоков, содержащих по k_1 символов, которые декодируются внешним кодом. В результате на выходе внешнего декодера появляется k_2 подблоков по k_1 символов. Каскадные коды имеют большую длину и большее кодовое расстояние. Среди блочных кодов они являются одними из наилучших. В частности, для двоичного симметричного канала существует каскадный код, при котором вероятность ошибки может быть сколь угодно малой.

Контрольные вопросы

1. Чем определяется количество информации, заключенное в конкретном сообщении?
2. Как связаны между собой собственная информация и энтропия?
3. Какие свойства энтропии Вам известны?
4. Как связаны между собой энтропия дискретного источника и производительность этого источника?
5. Чем определяется скорость передачи информации по каналу связи?
6. В чем суть основной теоремы кодирования К.Шеннона при отсутствии помех?
7. Что такое дифференциальная энтропия? Сравните ее с энтропией дискретных сообщений.
8. Дайте определение ε -энтропии. Как связаны между собой ε -энтропия и производительность источника непрерывных сообщений?
9. Чему равны пропускные способности дискретного, непрерывного каналов связи?
10. Как формулируется теорема кодирования К.Шеннона для канала с помехами?
11. Сформулируйте теоремы, определяющие обнаруживающую и исправляющую способность помехоустойчивых кодов. Как записывается условие обнаружения (исправления) ошибок?
12. Какие коды называются линейными двоичными блочными кодами? Перечислите их основные свойства. Приведите схемы кодера и декодера.
13. Какие коды называются циклическими? Как они задаются?
14. Как происходит исправление ошибок в случае циклического кода? Приведите примеры схем кодирующего и декодирующего устройств.
15. На каком свойстве циклических кодов основано мажоритарное декодирование?
16. Перечислите основные свойства итеративных кодов. В чем их достоинство?
17. В чем сущность сверточных кодов? Приведите пример кодирования информации сверточными кодами.

Глава 7. ТЕОРИЯ ПОМЕХОУСТОЙЧИВОСТИ СИСТЕМ СВЯЗИ

7.1. Оптимальный прием элементов дискретных сообщений. Критерии оптимального приема

Целью оптимального приема (так же, как и помехоустойчивого кодирования) является *повышение верности принимаемых сообщений*. Если при помехоустойчивом кодировании эта задача решается выбором кода, который определяет структуры кодера и декодера, то при оптимальном приеме задача решается выбором структуры приемника (демодулятора). Структура приемника, удовлетворяющая заданному критерию оптимальности, называется *оптимальной*, а приемник — *оптимальным*.

Приемник анализирует смесь элемента сигнала и помехи

$$z(t) = u_i(t) + N(t), \quad i = 1, 2, \dots, m$$

в течение единичного интервала времени и принимает решение о том, какой из возможных сигналов (символов) присутствует в этой смеси. С приходом следующего элемента анализ повторяется. Этот способ приема получил название *посимвольного*. Возможными видами сигналов могут быть: двоичные АМ, ФМ, ОФМ и ЧМ сигналы; многопозиционные (m -ичные) сигналы с ФМ, ОФМ и с квадратурной амплитудной модуляцией (КАМ). Методы формирования сигналов и их свойства анализировались в гл.3. Обычно полагают, что помеха, возникающая в линии связи, является *нормальным белым шумом*.

При посимвольном приеме решаются следующие задачи:

- 1) задача *обнаружения сигнала*;
- 2) задача *различения одного из m сигналов*.

Первая задача возникает при приеме сигналов, соответствующих элементам двоичного кодового слова, которые равны 0 или 1. Приемник, анализируя смесь сигнала и шума в течение длительности единичного интервала, дает ответ: есть на входе сигнал или его нет.

Вторая задача возникает: а) при приеме элементов двоичных сигналов, которым соответствуют элементы двоичного кодового слова 1, -1; б) при приеме элементов m -ичных сигналов, соответствующих символам m -ичных кодовых слов. Впервые задача оптимального приема была поставлена и решена в 1946 г. выдающимся русским ученым В.А.Котельниковым [1].

Рассмотрим задачу посимвольного приема двоичных (бинарных) сигналов на фоне нормального белого шума. Форма сигналов известна точно, т.е. описывается детерминированными

функциями времени $u_i(t)$, $i = 1, 2$. Сигналы имеют одинаковую длительность $\tau_{\text{и}}$, равную длительности единичного интервала, причем границы, в которых происходит их прием, также известны точно. Выполнение последнего условия обеспечивается с помощью устройства тактовой синхронизации. Вопросы синтеза и анализа устройств синхронизации рассматриваются подробно в специальных дисциплинах. С изложением вопросов теории и техники синхронизации можно познакомиться в [12].

Символы двоичного кодового слова изменяются в соответствии с передаваемой информацией, причем эти изменения до обработки на приемном конце неизвестны. Соответственно неизвестно какой из двух сигналов — $u_1(t)$ или $u_2(t)$ присутствует на входе. Полагают, что априорные (доопытные) вероятности присутствия сигналов на входе равны соответственно $P[u_1(t)]$, $P[u_2(t)]$, причем их сумма равна 1.

Алгоритм работы приемника заключается в вычислении апостериорной (послеопытной) вероятности возможных значений бинарного сигнала на основе анализа смеси сигнала и шума и принятии решения о том, какое возможное значение бинарного сигнала (символа) присутствует на его входе.

Значение бинарного сигнала на выходе приемника $u^*(t)$ из-за присутствия шума не совпадает со значением сигнала на входе $u(t)$, т.е. является его *оценкой*. Оценка обозначена звездочкой.

Вычислим апостериорную вероятность бинарного сигнала. В соответствии с теоремой Котельникова представим случайный процесс $z(t)$ его выборочными значениями, полагая, что спектр квазибелого шума ограничен верхней граничной частотой F , которая во много раз больше ширины спектра дискретного сигнала.

Выборочные значения $z(t_k)$, $k = 1, 2, \dots, n$ следует через интервал $\Delta t = 1/2F$, поэтому число выборок в интервале наблюдения смеси $z(t)$ $n = \tau_{\text{и}}/\Delta t = 2F\tau_{\text{и}}$.

Случайный процесс $z(t)$ считается заданным, если задана его n -мерная плотность вероятности

$$w(y_1, y_2, \dots, y_n) = \frac{d^n F(y_1, y_2, \dots, y_n)}{dy_1 dy_2 \dots dy_n}, \quad (7.1)$$

где $F(y_1, y_2, \dots, y_n) = P[z(t_1) \leq y_1; z(t_2) \leq y_2, \dots, z(t_n) \leq y_n]$.

Здесь $F(y_1, y_2, \dots, y_n)$ — интегральная функция распределения, P — соответствующая вероятность.

Зная n -мерную плотность, можно вычислить вероятность реализации случайного процесса $z(t)$:

$$\begin{aligned} P(A) &= \int \dots \int w(y_1, y_2, \dots, y_n) dy_1 dy_2 \dots dy_n \cong \\ &\cong w(y_1, y_2, \dots, y_n) dy_1 dy_2 \dots dy_n = w(\bar{y}) d\bar{y}, \end{aligned} \quad (7.2)$$

где $\bar{y} = \{y_1, y_2, \dots, y_n\}$ — n -мерный вектор, состоящий из совокупности n чисел; $d\bar{y} = dy_1 dy_2 \dots dy_n$ — объем бесконечно малой области n -мерного пространства; dY — пределы интегрирования, определяемые заданной областью пространства; A — событие, заключающееся в том, что случайный процесс $z(t)$ находится в пределах заданной области n -мерного пространства. Это событие независимо и совместно с другим событием B , которое заключается в присутствии на входе приемника одного из двух сигналов $u_i(t)$, $i = 1, 2$. Математическое описание событий A и B исчерпывается совместной вероятностью:

$$P(A, B) = P(A) P(B/A) = P(B) P(A/B), \quad (7.3)$$

где $P(A)$ — вероятность реализации смеси $z(t)$ в интервале $0 - \tau_n$; $P(B)$ — априорная вероятность бинарного сигнала; $P(B/A)$ — условная вероятность, равная апостериорной вероятности бинарного сигнала; $P(A/B)$ — функция правдоподобия.

Из (7.3) находим выражение для апостериорной вероятности:

$$P(B/A) = \frac{P(B)P(A/B)}{P(A)}. \quad (7.4)$$

Вероятность $P(A/B)$ может быть найдена с помощью n -мерной условной плотности вероятности

$$P(A/B) = \int \dots \int_{dY} w[y_1, y_2, \dots, y_n / u_i(t)] dy_1 dy_2 \dots dy_n \cong w[\bar{y} / u_i(t)] d\bar{y}. \quad (7.5)$$

Эта плотность

$$w[y_1, y_2, \dots, y_n / u_i(t)] = \frac{d^n F(y_1, y_2, \dots, y_n / u_i(t))}{dy_1 dy_2 \dots dy_n}, \quad (7.6)$$

где $F(y_1, y_2, \dots, y_n / u_i(t)) = P[z(t_1) \leq y_1; z(t_2) \leq y_2 \dots z(t_n) \leq y_n / u_i(t)]$.

Отличие n -мерных плотностей (7.1) и (7.6) в том, что если в ансамбль реализаций $z(t)$ первой плотности входят всевозможные реализации (при $u_i(t)$, $i = 1, 2$), то в ансамбль второй плотности входят только те, которые соответствуют или сигналу $u_1(t)$ или сигналу $u_2(t)$.

Подставляя выражения (7.2), (7.5) в (7.4), получим

$$P(B/A) = \frac{P(B)w(\bar{y} / u_i(t)d\bar{y}}{w(\bar{y})d\bar{y}} = P(B) \frac{w(\bar{y} / u_i(t)d\bar{y}}{w(\bar{y})}.$$

Вычисленные апостериорные вероятности сравниваются между собой. Если выполняется неравенство

$$P(B = u_1(t) / A) = P[u_1(t)] \frac{w(\bar{y} / u_1(t))}{w(\bar{y})} > P(B = u_2(t) / A) =$$

$$= P(u_2(t)) \frac{w(\bar{y} / u_2(t))}{w(\bar{y})}, \quad (7.7)$$

то принимается решение, что принят сигнал $u_1(t)$. Если выполняется неравенство

$$P(B = u_2(t) / A) > P(B = u_1(t) / A),$$

то принимается решение, что принят сигнал $u_2(t)$.

Перепишем неравенство (7.7) в виде

$$\Lambda = \frac{w(\bar{y} / u_1(t))}{w(\bar{y} / u_2(t))} > \frac{P[u_2(t)]}{P[u_1(t)]}.$$

Отношение Λ называется *отношением правдоподобия* двух гипотез о том, что передавался сигнал $u_1(t)$ и сигнал $u_2(t)$. Правило принятия решения называется *критерием оптимальности приемника*.

Решение задачи посимвольного приема m -ичного сигнала мало чем отличается от приведенного выше решения для бинарного сигнала: приемник вычисляет апостериорные вероятности и по максимальной из них определяет возможный сигнал на входе.

Правило принятия решения записывается в виде

$$P[u_i(t)] \frac{w[\bar{y} / u_i(t)]}{w(\bar{y})} > P[u_j(t)] \frac{w[\bar{y} / u_i(t)]}{w(\bar{y})}, \quad (7.8)$$

где $P[u_i(t)]$ априорная вероятность i -го сигнала; $w[\bar{y} / u_i(t)]$ — функция правдоподобия; $\sum_{i=1}^m P[u_i(t)] = 1$.

Правило (7.8) можно записать иначе через отношение правдоподобия $\Lambda_{i,j}$ двух гипотез о том, что передавался сигнал $u_i(t)$ и сигнал $u_j(t)$:

$$\Lambda_{i,j} = \frac{w[\bar{y} / u_i(t)]}{w[\bar{y} / u_j(t)]} > \frac{P[u_j(t)]}{P[u_i(t)]}. \quad (7.9)$$

$$j \neq i \quad i, j = 1, 2, 3, \dots, m.$$

Критерий максимума отношения правдоподобия не является единственным. Другими используемыми критериями являются: *критерий минимума среднего риска*, *весовой*, *Неймана-Пирсона*. Поскольку в технике связи преимущественно применяют критерий максимального правдоподобия, мы в дальнейшем будем использовать этот критерий.

7.2. Синтез оптимальных приемников при приеме элементов дискретных сигналов, известных точно

7.2.1. Реализация оптимального приема на базе коррелятора

Под *синтезом* оптимального приемника понимают отыскание его структуры. Задача синтеза формулируется следующим образом. Требуется найти структуру приемника, которая удовлетворяет исходным условиям и ограничениям и при этом обеспечивает совокупность показателей качества, наилучших в смысле заданного критерия оптимальности.

Синтез обычно проводят, сочетая эвристический метод, основанный на использовании накопленного инженерного опыта и знаний по проектированию приемников, и математический анализ.

Математический синтез заключается в математической формулировке совокупности исходных данных и критерия оптимальности, а также в отыскании чисто математическим путем такой структуры приемника, которая удовлетворяет исходным данным, критерию оптимальности и требуемой совокупности показателей качества.

Рассмотрим задачу математического синтеза оптимального приемника сигналов, известных точно.

Исходными данными являются:

- а) вид зависимости сигнала $u_i(t)$, $i = 1, 2, \dots, m$ от его параметров;
- б) вид смеси сигнала и помехи;
- в) вид канала связи;
- г) время анализа смеси сигнала и помехи.

Сигналы $u_i(t)$ $i = 1, 2, \dots, m$ являются детерминированными функциями времени, амплитуды, частоты и начальной фазы. Их начало и окончание заранее известно и совпадает с началом и окончанием единичного интервала. Выполнение последнего условия обеспечивается устройствами тактовой синхронизации.

Помеха является нормальным белым шумом со спектральной плотностью N_0 . Канал связи является дискретно-непрерывным. Время анализа смеси сигнала и шума равно длительности единичного интервала τ_d .

В качестве критерия оптимальности выберем критерий максимума отношения правдоподобия, записанного в виде неравенств (7.9).

Используя результаты, приведенные в разд. 4.3 и связанные с описанием n -мерной плотности распределения вероятности белого шума $N(t)$, а также сведения из 5.2.1, 5.2.3, касающиеся

математического описания непрерывного и дискретно-непрерывного каналов связи, перепишем неравенство (7.9) в виде:

$$\Lambda_{i,j} = \frac{e^{-\frac{1}{N_0} \int_0^{\tau_H} [z(t) - u_i(t)]^2 dt}}{e^{-\frac{1}{N_0} \int_0^{\tau_H} [z(t) - u_j(t)]^2 dt}} > \frac{P[u_j(t)]}{P[u_i(t)]}. \quad (7.10)$$

При написании неравенств (7.10) использована формула (5.18).

Вместо максимума Λ_{ij} можно искать максимум $\ln \Lambda_{ij}$. Тогда, если все сообщения равновероятны и $P[u_i(t)] = P[u_j(t)]$, условие (7.10) можно записать иначе:

$$\ln \Lambda_{ij} = -\frac{1}{N_0} \int_0^{\tau_H} [z(t) - u_i(t)]^2 dt + \frac{1}{N_0} \int_0^{\tau_H} [z(t) - u_j(t)]^2 dt > 0, j \neq i.$$

Полученное неравенство можно переписать также в виде:

$$\int_0^{\tau_H} [z(t) - u_j(t)]^2 dt > \int_0^{\tau_H} [z(t) - u_i(t)]^2 dt, j \neq i.$$

Раскрывая скобки в подинтегральных выражениях и сокращая в обеих частях неравенства слагаемые $\int_0^{\tau_H} z^2(t) dt$, получим следующий алгоритм приема:

$$\int_0^{\tau_H} z(t) u_i(t) dt - 0,5 E_i > \int_0^{\tau_H} z(t) u_j(t) dt - 0,5 E_i, j \neq i. \quad (7.11)$$

где E_i — энергия принимаемого сигнала:

$$E_i = \int_0^{\tau_H} u_i^2(t) dt.$$

Если все сигналы $u_i(t)$, $i = 1, 2, \dots, m$ имеют одинаковые энергии, то неравенство (7.11) упрощается:

$$\left. \begin{aligned} & \int_0^{\tau_H} z(t) u_i(t) dt > \int_0^{\tau_H} z(t) u_j(t) dt, j \neq i \\ & \text{или} \\ & \max_i \left[\int_0^{\tau_H} z(t) u_i(t) dt \right], i = 1, 2, \dots, m. \end{aligned} \right\} \quad (7.12)$$

Структура оптимального приемника, которая определяется уравнениями (7.12), содержит: m корреляторов, вычисляющих

корреляционные интегралы; решающее устройство, в котором результаты вычислений сравниваются между собой. По коррелятору, на выходе которого фиксируется наибольшее значение корреляционного интеграла, определяется сигнал, присутствующий на входе и совпадающий с опорным сигналом этого коррелятора. По найденному сигналу $u_i^*(t)$ находится соответствующий ему символ b_i^* . Структурная схема оптимального приемника построенного на основе m корреляторов, приведена на рис.7.1.

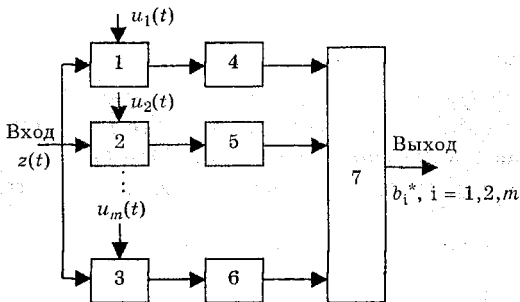


Рис.7.1. Структурная схема оптимального приемника (демодулятора) m сигналов, известных точно: 1,2,3 — перемножители; 4,5,6 — интеграторы; 7 — решающее устройство

Отметим, что приемник синтезирован с целью поэлементного приема сигнала, с числом возможных элементов (букв алфавита) равным m , причем каждому элементу b_i ставится в однозначное соответствие сигнал $u_i(t)$.

Рассмотрим случай приема на фоне помех m разрешенных двоичных кодовых комбинаций, полученных в результате помехоустойчивого кодирования. Поскольку

все эти комбинации известны точно, их прием можно вести аналогично рассмотренному выше. При этом будет реализован прием в целом, практически не отличающийся от схемы на рис.7.1.

При приеме в целом (n, k) — кода в смеси $z(t)$ присутствует сигнал, соответствующий i -й разрешенной кодовой комбинации. Опорными являются сигналы, которые соответствуют всем разрешенным кодовым комбинациям. На выход решающего устройства поступает оценка информационного кодового слова. Для используемых на практике (n, k) -кодов числа n, k могут быть очень большими и оптимальный приемник (демодулятор) становится трудно реализуемым.

Для двоичной системы ($m = 2$) неравенство 7.11 записывается в виде

$$\int_0^{\tau_n} z(t) u_{\Delta}(t) dt > \lambda = 0,5(E_1 - E_2), \quad (7.13)$$

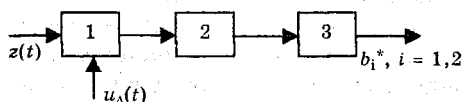


Рис.7.2. Структурная схема оптимального приемника двоичных сигналов: 1 — перемножитель; 2 — интегратор; 3 — решающее устройство

где $u_{\Delta}(t) = u_1(t) - u_2(t)$ — разностный сигнал; λ — пороговый уровень.

Структурная схема оптимального приемника (демодулятора) для приема двоичного сигнала показана на рис.7.2.

Рассмотрим случаи приема сигналов.

7.2.2. Бинарное обнаружение сигналов

Задача бинарного обнаружения сигналов решается: а) при приеме видеосигнала, полученного модуляцией постоянного тока по амплитуде; б) при приеме дискретного АМ радиосигнала.

В первом случае $u_1(t) = a$, $u_2(t) = 0$ в течение длительности единичного интервала τ_n . Энергия сигналов $E_1 = \int_0^{\tau_n} u_1^2(t) dt = a^2 \tau_n$,

$E_2 = 0$, а пороговый уровень $\lambda = a^2 \tau_n / 2$. Условие (7.13) записывается в виде:

$$a \int_0^{\tau_n} z(t) dt > a^2 \tau_n / 2. \quad (7.14)$$

Структура оптимального приемника состоит из интегратора, работающего в течение времени τ_n , и решающего устройства, в котором вычисленное значение интеграла сравнивается с пороговым уровнем, равным $a^2 \tau_n / 2$. Рассмотренный алгоритм (7.14) используют в простейших устройствах проводной связи.

В кабельных линиях и каналах радиосвязи применяют дискретный АМ радиосигнал. Энергия этого сигнала

$$E_1 = a^2 \int_0^{\tau_n} \cos^2(\omega_0 t + \varphi) dt = a^2 \tau_n / 2, E_2 = 0, \text{ а пороговый уровень}$$

$$\lambda = a^2 \tau_n / 4.$$

Структурная схема оптимального приемника для этого случая приведена на рис.7.3.

В состав схемы входит когерентный детектор, содержащий перемножитель и интегратор, а также дискретизатор и решающее устройство. Интегратор является одним из вариантов фильтра нижних частот. Напомним, что приемник, работа которого основывается на знании фазы и частоты несущей принимаемого сигнала, называется когерентным.

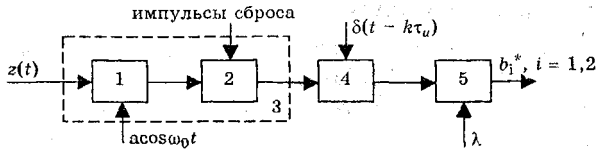


Рис.7.3. Структурная схема оптимального когерентного приемника: 1 — перемножитель; 2 — интегратор; 3 — когерентный детектор; 4 — дискретизатор; 5 — решающее устройство

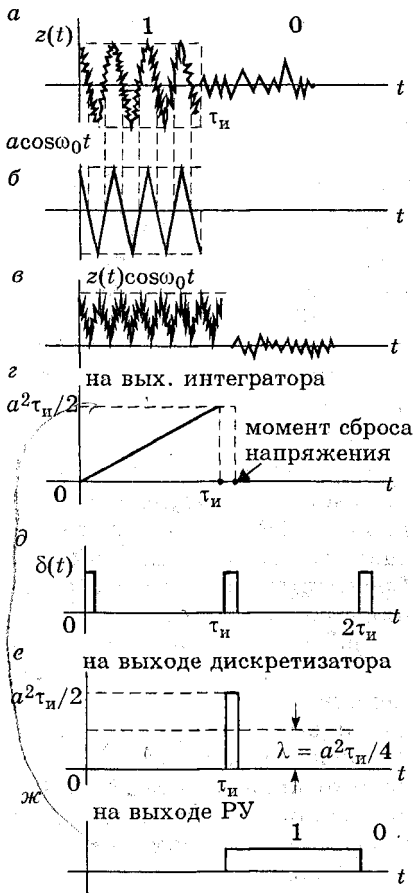


Рис.7.4. Временные диаграммы, поясняющие работу приемника АМ сигналов

Временные диаграммы, поясняющие работу когерентного приемника приведены на рис.7.4.

Вырабатываемый в приемнике опорный гармонический сигнал является непрерыв-

ным колебанием, а его амплитуда, фаза и частота с заданной степенью точности совпадают с аналогичными параметрами принимаемого сигнала. Опорный (рис.7.4, б) и принимаемый (рис.7.4, а) сигналы умножаются в перемножителе, а результат (рис.7.4, в) интегрируется. Напряжение с выхода интегратора (рис.7.4, г) дискретизируется. Выборочные значения (рис.7.4, е) с выхода дискретизатора поступают на решающее устройство, где сравниваются с заданным пороговым уровнем λ . После этого за короткий промежуток времени производится сброс накопленного интегратором напряжения для того, чтобы устранить связь между результатами приема в соседних единичных интервалах. Опорный сигнал и импульсы управления сбросом напряжения на интеграторе вырабатываются устройством синхронизации, которое на рис.7.3 не показано.

Если напряжение на выходе интегратора больше порогового уровня λ , то решающее устройство вырабатывает единичный символ, а если меньше, то нулевой.

7.2.3. Бинарное распознавание сигналов

Эта задача решается при приеме сигналов постоянного тока, модулированных по знаку, ФМ, ОФМ и ЧМ сигналов.

При приеме сигнала постоянного тока, модулированного по знаку разностный сигнал

$$u_{\Delta}(t) = 2u_1(t), \quad \text{т.к. } u_2(t) = -u_1(t);$$

энергия $E_1 = E_2$, а уровень порога $\lambda = 0$.

Правило принятия решения равно:

$$\int_0^{\tau_H} z(t)u(t)dt > 0.$$

При бинарном распознавании противоположных ФМ, ОФМ сигналов

$$u_{\Delta}(t) = 2u_1(t) = 2a \cos(\omega_0 t + \varphi), \quad E_1 = E_2, \quad \lambda = 0,$$

а правило принятия решения равно:

$$\int_0^{\tau_H} z(t) \cos(\omega_0 t + \varphi) dt > 0. \quad (7.15)$$

Неравенство (7.15) позволяет оценить знак принимаемого сигнала. Если корреляционный интеграл больше нуля, то в решающем устройстве вырабатывается символ $b^* = 1$. В противном случае $b^* = -1$.

Структурная схема оптимального демодулятора бинарных ФМ сигналов совпадает со схемой на рис.7.3. Временные диаграммы, поясняющие его работу, приведены на рис.7.5. Так же, как и при приеме дискретного АМ сигнала, известного точно, для нормальной работы оптимального демодулятора необходимо производить сброс накопленного напряжения на интеграторе в конце интервала наблюдения.

Структурная схема оптимального демодулятора сигнала с ОФМ приведена на

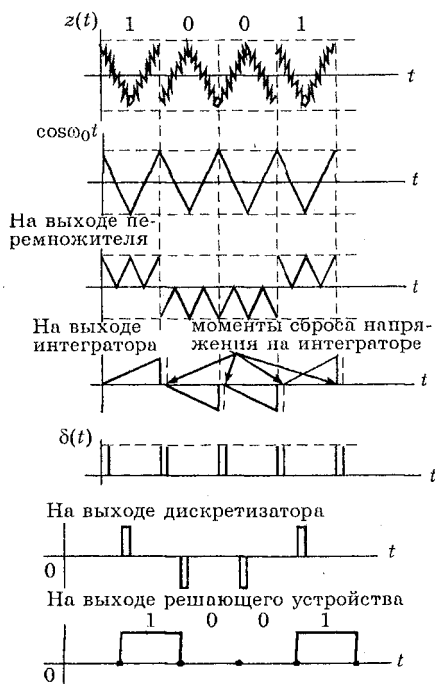


Рис.7.5. Временные диаграммы, поясняющие работу оптимального приемника (демодулятора) бинарных ФМ сигналов

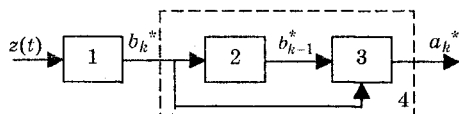


Рис.7.6. Структурная схема оптимального демодулятора ОФМ сигнала:

1 — демодулятор; 2 — линия задержки на один единичный интервал; 3 — сумматор по модулю 2; 4 — преобразователь

рис.7.6. Она содержит оптимальный демодулятор ФМ сигнала и преобразователь относительного двоичного кодового слова $\{b_k^*\}$ в информационное кодовое слово $\{a_k^*\}$. Работа преобразователя поясняется временными диаграммами (рис.7.7) и основана на суммировании по модулю 2 принятого символа b_k^* с принятым в преды-

дущем тактовом интервале символом b_{k-1}^* :

$$a_k^* = b_k^* + b_{k-1}^*.$$

При бинарном распознавании ортогональных сигналов (ЧМ)

$$u_1(t) = a \cos(\omega_1 t + \varphi_1), \quad u_2(t) = a \cos(\omega_2 t + \varphi_2),$$

разностный сигнал

$$u_{\Delta}(t) = a \cos(\omega_1 t + \varphi_1) - a \cos(\omega_2 t + \varphi_2),$$

$$E_1 = E_2, \lambda = 0.$$

Правило решения равно:

$$\int_0^{\tau_{\Pi}} z(t) [\cos(\omega_1 t + \varphi_1) - \cos(\omega_2 t + \varphi_2)] dt > 0. \quad (7.16)$$

Его можно переписать в виде:

$$\begin{aligned} & \int_0^{\tau_{\Pi}} z(t) \cos(\omega_1 t + \varphi_1) dt > \\ & > \int_0^{\tau_{\Pi}} z(t) \cos(\omega_2 t + \varphi_2) dt. \end{aligned} \quad (7.17)$$

При выполнении неравенства принимается решение о том, что принятый символ $b^* = 1$. Если

неравенство не выполняется, то принимается решение, что принятый символ $b^* = 0$.

Оптимальный демодулятор можно строить, используя правило (7.16) или правило (7.17).

Для правила (7.16) схема приведена на рис.7.2, при $\lambda = 0$ и

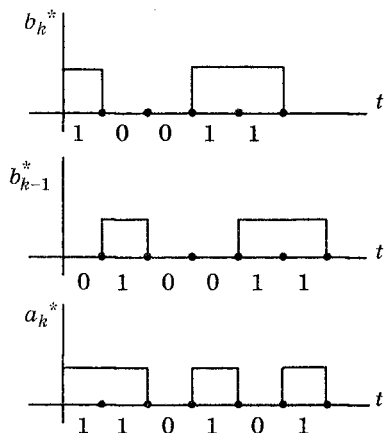


Рис.7.7. Временные диаграммы, поясняющие работу преобразователя

$$u_{\Delta}(t) = \cos(\omega_1 t + \varphi) - \cos(\omega_2 t + \varphi_2).$$

Эта схема содержит один коррелятор. Для правила (7.17) схема содержит два коррелятора и является частным случаем схемы рис. 7.1 при $m = 2$.

7.2.4. Реализация алгоритма оптимального приема на основе согласованных фильтров

Алгоритм оптимального приема сигналов, известных точно, можно реализовать также на основе согласованного фильтра.

Под *согласованным фильтром* понимают линейный четырехполюсник, коэффициент передачи которого

$$K(j\omega) = cS^*(j\omega)e^{-j\omega t_0}, \quad (7.18)$$

где c — постоянное число; $S^*(j\omega)$ — функция, комплексно сопряженная со спектральной плотностью принимаемого сигнала; $e^{-j\omega t_0}$ — фазовый множитель, совпадающий с коэффициентом передачи идеальной линии задержки.

Спектральная плотность сигнала на выходе согласованного фильтра

$$S_{\text{вых}}(j\omega) = S(j\omega)K(j\omega) = cS(j\omega)S^*(j\omega)e^{-j\omega t_0} = c|S(j\omega)|^2 e^{-j\omega t_0},$$

где $S(j\omega)$ — спектральная плотность сигнала на входе фильтра.

Отклик согласованного фильтра на входной сигнал с учетом формулы (2.23) равен:

$$u_{\text{вых}}(\tau - t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} c|S(j\omega)|^2 e^{-j\omega t_0} e^{j\omega \tau} d\omega = cR(\tau - t_0).$$

При $\tau - t_0 = 0$ значение отклика согласованного фильтра с точностью до постоянного числа совпадает со значением автокорреляционной функции входного сигнала $R(0)$.

Приведем другие свойства согласованного фильтра:

а) импульсная характеристика согласованного фильтра является зеркальным отображением сигнала $u(t)$.

Вычислим реакцию согласованного фильтра на входной сигнал, используя интеграл Дюамеля,

$$u_{\text{вых}}(t) = \int_0^{\infty} u(\tau)h(t - \tau)d\tau,$$

где $h(t)$ — импульсная характеристика.

Если фильтр согласован с входным сигналом, то, как показано выше, $u_{\text{вых}}(t) = R(t)$.

Следовательно,

$$R(t) = \int_0^{\infty} u(\tau)h(t - \tau)d\tau. \quad (7.19)$$

Сравнивая (7.19) с (2.21) находим, что

$$u(\tau + t) = h(t - \tau);$$

б) отношение мгновенной мощности сигнальной составляющей к средней мощности шума на выходе линейного фильтра равно $2E/N_0 = h^2$, где E — энергия сигнала, N_0 — мощность шума, приходящаяся на единицу полосы.

Мгновенная мощность сигнала на выходе согласованного фильтра в момент времени t_0 :

$$u_{\text{вых}}^2(t_0) = \left[\frac{c}{2\pi} \int_{-\infty}^{\infty} S(j\omega) S^*(j\omega) e^{-j\omega t_0} e^{j\omega t} d\omega \right]^2 = c^2 \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} |S(j\omega)|^2 d\omega \right]^2 = c^2 E^2.$$

Средняя мощность шума на выходе согласованного фильтра

$$P_{\text{ш}} = \frac{N_0}{2} \frac{1}{2\pi} \int_{-\infty}^{\infty} |c S^*(j\omega) e^{-j\omega t_0}|^2 d\omega = \frac{c^2 N_0 E}{2}.$$

Отношение

$$\rho_{\text{пик}} = \frac{u_{\text{вых}}(t_0)}{P_{\text{ш}}} = \frac{c^2 E^2}{c^2 N_0 E / 2} = \frac{2E}{N_0} = 2h^2.$$

С помощью неравенства Буняковского-Шварца можно доказать, что максимальное отношение, равное $2h^2$ реализуется на выходе линейного четырехполосника, если его коэффициент передачи определен выражением (7.18) [3].

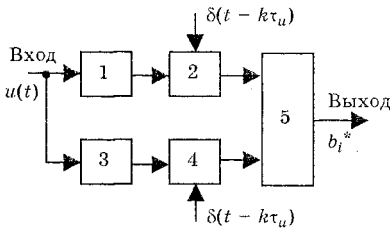


Рис.7.8. Структурная схема демодулятора ЧМ сигнала:

- 1,3 — согласованные фильтры;
- 2,4 — дискретизаторы (ключи);
- 5 — решающее устройство

сигналом, который показан в первом единичном интервале (рис.7.9, а), а фильтр нижнего канала — с сигналом, который показан во втором единичном интервале. Поэтому в течение первого единичного интервала на входной сигнал откликается только первый (рис.7.9, б), а в течение второго интервала — только

Рассмотрим работу оптимальных демодуляторов, реализованных на согласованных фильтрах. Пусть на входе присутствует ЧМ сигнал. Соответствующая этому случаю приема схема приведена на рис.7.8. Она содержит два одинаковых канала (верхний и нижний) и решающее устройство. В каждом канале содержится согласованный фильтр и дискретизатор (ключ). Согласованные фильтры откликаются только на те сигналы, с которыми они согласованы. Например, фильтр верхнего канала согласован с

второй фильтр (рис.7.9, в). Ключи, управляемые короткими импульсами (рис.7.9, з), пропускают отклики фильтров на входы решающего устройства только в течение действия этих импульсов. Таким образом в конце первого единичного интервала напряжение на выходе первого ключа не равно нулю, а на выходе второго — равно нулю. В конце второго единичного интервала напряжение на выходе первого ключа равно нулю, а на выходе второго — равно нулю. В зависимости от этого решающее устройство выдает на свой выход символ b_i , равный 0 или 1.

Пусть на входе присутствует АМ или ФМ сигнал. Соответствующая этому случаю приема схема оптимального демодулятора приведена на рис.7.10. В отличие от схемы на рис.7.8, она содержит всего один канал. При помощи ключа осуществляется

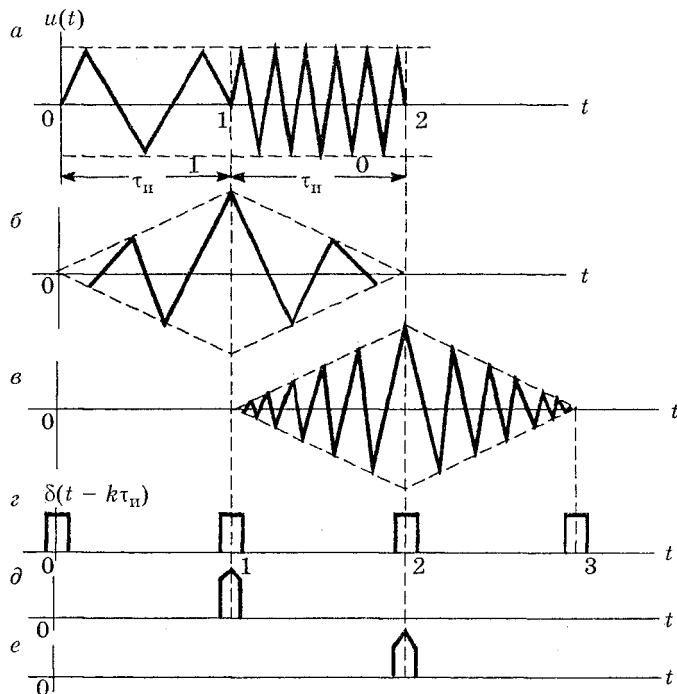


Рис.7.9. Временные диаграммы, поясняющие работу демодулятора ЧМ сигнала

дискретизация напряжения на выходе согласованного фильтра, затем полученные в конце каждого единичного интервала выборочные значения сравниваются в вычитающем устройстве с зара-

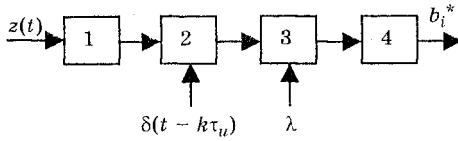


Рис.7.10. Структурная схема оптимального демодулятора АМ сигнала: 1 — согласованный фильтр; 2 — дискретизатор (управляемый ключ); 3 — сравнивающее устройство; 4 — решающее устройство

Если схему оптимального демодулятора ФМ сигнала (рис.7.10) дополнить преобразователем относительного кодового

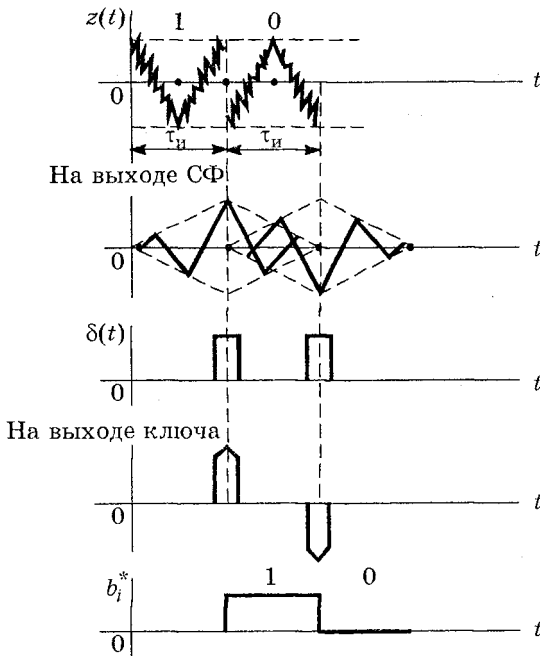


Рис.7.11. Временные диаграммы, поясняющие работу оптимального демодулятора ФМ сигналов

нее заданным порогом λ . Для АМ $\lambda = E/2$, где E — энергия сигнала, принятого в течение единичного интервала времени. Для ФМ $\lambda = 0$. В зависимости от результата сравнения на выходе решающего устройства формируется символ b_i^* , равный 0 или 1. Временные диаграммы, поясняющие работу демодулятора ФМ сигналов, приведены на рис.7.11.

Если схему оптимального демодулятора ФМ сигнала дополнить преобразователем относительного кодового слова (см. рис.7.6), то легко получить и схему оптимального демодулятора ОФМ сигнала.

В заключение отметим следующие особенности приемников с согласованными фильтрами.

1. Форма сигнала на выходе согласованного фильтра существенно отличается от формы сигнала на его входе. Например, при прямоугольной форме радиоимпульса на входе получается ромбовидный импульс на выходе. Однако, в сохранении формы нет необходимости, так как для принятия решения о символе достаточно сравнить выборочные значения откликов этих фильтров с соответствующими порогами.

2. В схемах оптимальных приемников с согласованными фильтрами нет опорных генераторов, следовательно, отпадает необходимость согласования фаз опорных с фазами принимаемых сигналов.

3. В схемах с согласованными фильтрами сохраняется существенный недостаток, связанный с необходимостью высокоточных когерентных отсчетов значений откликов в моменты окончания единичных интервалов.

7.3. Оптимальный прием при неопределенной фазе и амплитуде элементов дискретного сигнала

7.3.1. Оптимальный прием при неопределенной фазе дискретного сигнала

Этот способ некогерентного приема применяют в случаях, если: а) реализация приема сигнала, известного точно, невозможна и экономически нецелесообразна из-за большой сложности демодулятора; б) оценку фазы и амплитуды получить достаточно трудно из-за их быстрых флуктуаций.

Пусть на входе демодулятора в смеси с нормальным белым шумом присутствует один из возможных сигналов $u_i(t)$, $i = 1, 2, \dots, m$ с параметрами, известными точно, за исключением начальных фаз, которые в интервале $0 - 2\pi$ независимы и распределены по равномерному закону.

Определим структуру оптимального демодулятора, приняв во внимание:

- а) модель канала связи с неопределенной фазой сигнала;
- б) выражение для функции правдоподобия;
- в) правило принятия решения.

Модель канала связи рассмотрена в разд. 5.2. Используя эту модель, запишем выражение для функции правдоподобия:

$$w[\bar{y} / u_i, \varphi_i] = \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^n e^{-\frac{1}{N_0} \int_0^{\tau_n} [z(t) - A_i \cos(\omega_i t + \varphi_i)]^2 dt}, \quad (7.20)$$

где сигнал $u_i(t)$ определяется гармоническим колебанием с известной амплитудой A_i и частотой ω_i и неизвестной (случайной) начальной фазой φ_i .

Чтобы записать правило принятия решения, рассмотрим вероятность совместных событий:

- а) событие A заключается в том, что случайный процесс $z(t)$ находится в заданной области; б) событие B заключается в реализации одного из возможных сигналов $u_i(t)$; в) событие C заклю-

чается в том, что начальная фаза сигнала $u_i(t)$ находится в бесконечно малом интервале $d\varphi_i$. Эта вероятность

$$P(A, B, C) = P(A)P(B, C / A) = P(BC)P(A / BC),$$

где $P(A)$ — вероятность события A , определяемая равенством (7.2); $P(B, C)$ — априорная, совместная вероятность событий B и C реализации сигнала $u_i(t)$ и начальной фазы φ_i , попадающей в заданную область.

Эта вероятность $P(B, C) = P(B)P(C) = P[u_i(t)1/2\pi d\varphi_i P(B, C/A)$ — условная вероятность реализации двух событий B и C (реализации сигнала $u_i(t)$ и реализации его начальной фазы в заданной области $d\varphi_i$; $P(A/BC)$ — функция правдоподобия, равная условной вероятности реализации смеси сигнала и шума, при условии, что в этой смеси присутствует сигнал $u_i(t)$, а его фаза находится в области $d\varphi_i$.

Используя полученное выражение, находим:

$$P(B, C / A) = \frac{P(B)P(C)}{P(A)} P(A / B, C). \quad (7.21)$$

Найдем вероятность реализации сигнала $u_i(t)$ при условии реализации смеси сигнала и шума $z(t)$, суммируя (7.21) по всем возможным значениям начальной фазы.

$$\begin{aligned} P(B / A) &= \sum_{\varphi_i} \frac{P(B)P(C)}{P(A)} P(A / B, C) = \sum_{\varphi_i} \frac{P[u_i(t)] \frac{1}{2\pi} d\varphi_i}{w(\bar{y}) d\bar{y}} w[\bar{y} / u_i, \varphi_i] d\bar{y} = \\ &= \frac{P[u_i(t)]}{2\pi\omega(\bar{y})} \int_0^{2\pi} w[\bar{y} / u_i, \varphi_i] d\varphi_i. \end{aligned}$$

Правило принятия решения по критерию максимальной обратной вероятности записывается аналогично (7.8):

$$\frac{P[u_i(t)]}{2\pi\omega(\bar{y})} \int_0^{2\pi} w[\bar{y} / u_i, \varphi_i] d\varphi_i > \frac{P[u_j(t)]}{2\pi\omega(\bar{y})} \int_0^{2\pi} w[\bar{y} / u_j, \varphi_j] d\varphi_j, \quad j \neq i, j = 1, 2, 3, \dots, m.$$

Сократив в левой и правой частях неравенства одинаковые выражения, получим:

$$\frac{\int_0^{2\pi} w[\bar{y} / u_i, \varphi_i] d\varphi_i}{\int_0^{2\pi} w[\bar{y} / u_j, \varphi_j] d\varphi_j} > \frac{P[u_j(t)]}{P[u_i(t)]}, \quad j \neq i, j = 1, 2, \dots, m. \quad (7.22)$$

Подставляя в полученное неравенство выражение для функции правдоподобия (7.20) и выполнив необходимые преобразования, получим:

$$\frac{e^{-E_i / N_0} \int_0^{2\pi} e^{\frac{2A_i}{N_0} \int_0^{\tau_H} z(t) \cos(\omega_i t + \varphi_i) dt} d\varphi_i}{e^{-E_j / N_0} \int_0^{2\pi} e^{\frac{2A_j}{N_0} \int_0^{\tau_H} z(t) \cos(\omega_j t + \varphi_j) dt} d\varphi_j} > \frac{P[u_j(t)]}{P[u_i(t)]}, j \neq i, j = 1, 2, \dots, m.$$

где $E_i = \int_0^{\tau_H} A_i^2 \cos^2(\omega_i t + \varphi_i) dt$ — энергия сигнала $u_i(t)$; N_0 — мощность шума, приходящаяся на единицу полосы.

Преобразуем выражение

$$\int_0^{\tau_H} z(t) \cos(\omega_i t + \varphi_i) dt,$$

приняв во внимание формулу

$$\cos(\omega_i t + \varphi_i) = \cos \omega_i t \cos \varphi_i + \sin \omega_i t \sin \varphi_i.$$

$$\begin{aligned} \int_0^{\tau_H} z(t) \cos(\omega_i t + \varphi_i) dt &= \left(\int_0^{\tau_H} z(t) \cos \omega_i t dt \right) \cos \varphi_i + \\ &+ \left(\int_0^{\tau_H} z(t) \sin \omega_i t dt \right) \sin \varphi_i = X \cos \varphi_i + Y \sin \varphi_i, \end{aligned} \quad (7.24)$$

где

$$X = \int_0^{\tau_H} z(t) \cos \omega_i t dt, \quad Y = \int_0^{\tau_H} z(t) \sin \omega_i t dt.$$

Считая значения X, Y проекциями вектора $\vec{M}$ на оси x, y , запишем

$$X = M \cos \theta_i, Y = M \sin \theta_i, \quad (7.25)$$

$$M = \sqrt{X^2 + Y^2} = \sqrt{\left(\int_0^{\tau_H} z(t) \cos \omega_i t dt \right)^2 + \left(\int_0^{\tau_H} z(t) \sin \omega_i t dt \right)^2}. \quad (7.26)$$

Подставляя (7.25) в (7.24), получим

$$\int_0^{\tau_H} z(t) \cos(\omega_i t + \varphi_i) dt = M \cos(\theta_i + \varphi_i). \quad (7.27)$$

Подставляя (7.27) в (7.23), получим

$$\frac{e^{-\frac{E_i}{N_0}} I_0\left(\frac{2A_i}{N_0} M_i\right)}{e^{-\frac{E_j}{N_0}} I_0\left(\frac{2A_j}{N_0} M_j\right)} > \frac{P[u_j(t)]}{P[u_i(t)]}, j \neq i, j = 1, 2, \dots, m, \quad (7.28)$$

где

$$I_0\left(\frac{2A_i}{N_0} M_i\right) = \int_0^{2\pi} e^{\frac{2A_i}{N_0} M_i \cos(\theta_i + \varphi_i)} d\varphi_i$$

модифицированная функция Бесселя нулевого порядка. Логарифмируя левые и правые части неравенства (7.28), перепишем его в виде:

$$\ln \left[\frac{I_0(2A_i M_i / N_0)}{I_0(2A_j M_j / N_0)} \right] > \ln \frac{P[u_j(t)]}{P[u_i(t)]} + \frac{E_i - E_j}{N_0}, j \neq i, j = 1, 2, \dots, m. \quad (7.29)$$

Для двоичной системы равновероятных сигналов правило оптимального некогерентного приема определяется неравенством

$$\ln I_0\left(\frac{2A_1 M_1}{N_0}\right) - \frac{E_1}{N_0} > \ln I_0\left(\frac{2A_2 M_2}{N_0}\right) - \frac{E_2}{N_0}. \quad (7.30)$$

При выполнении этого неравенства на выходе регистрируется 1, в противном случае — 0.

Структурная схема оптимального демодулятора, реализующего алгоритм (7.30), приведена на рис.7.12. Она содержит два

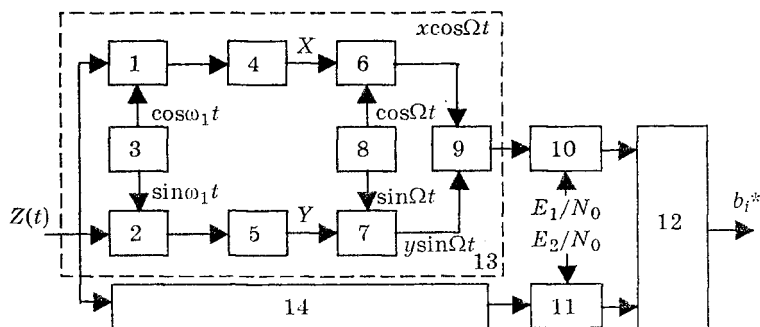


Рис.7.12. Структурная схема оптимального демодулятора двоичных сигналов с неопределенной начальной фазой:
1,2,6,7 — умножители; 4,5 — интеграторы; 3,8 — опорные генераторы; 9 — устройство, вычисляющее обгибающую M;
10,11 — сравнивающие устройства; 12 — решающее устройство; 13 — 1-й канал; 14 — 2-й канал

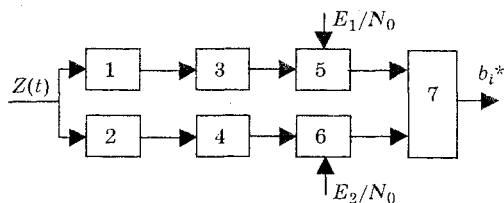


Рис.7.13. Структурная схема оптимального демодулятора, построенного на согласованных фильтрах: 1,2 — согласованные фильтры; 3,4 — амплитудные детекторы; 5,6 — сравнивающие устройства; 7 — решающее устройство

амплитудного детектора рассмотрена в разд.3.9. Структурная схема оптимального демодулятора построенная на согласованных фильтрах, приведена на рис.7.13.

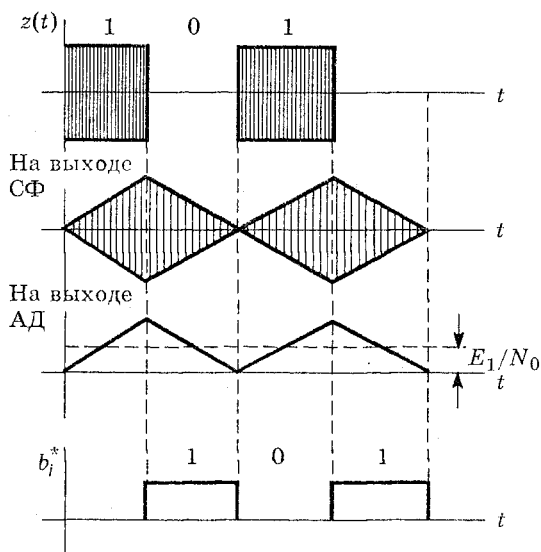


Рис.7.14. Временные диаграммы, поясняющие работу оптимального демодулятора АМ сигнала со случайной начальной фазой

идентичных канала и является довольно сложной. Однако, ее можно упростить, если применить согласованные фильтры. Величину M , входящую в выражение (7.26), можно получить, пропуская колебание $z(t)$ через согласованный фильтр и выделяя огибающую на выходе фильтра амплитудным детектором с характеристикой $\ln I_0(x)$, где $x = 2AM / N_0$. Работа

Рассмотрим случаи приема:

- а) двоичного АМ сигнала со случайной начальной фазой;
- б) двоичного ЧМ сигнала со случайной начальной фазой;
- в) ОФМ сигнала.

Оптимальный демодулятор АМ сигнала со случайной начальной фазой содержит всего один канал. Для этого случая $u_1(t) \neq 0$, $u_2(t) = 0$, $E_2 = 0$. Временные диаграммы, поясняющие работу этого оптимального демодулятора, приведены на рис.7.14.

Оптимальный демодулятор ЧМ сигнала

содержит два канала (рис.7.13). Работа каналов поясняется временными диаграммами, приведенными на рис.7.15.

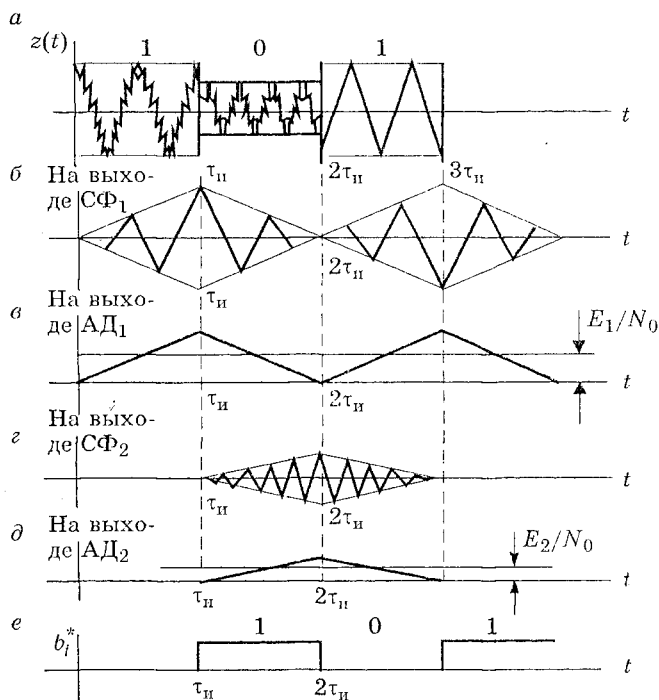


Рис.7.15. Временные диаграммы, поясняющие работу оптимального демодулятора ЧМ сигнала со случайной начальной фазой

Работа каналов при приеме АМ и ЧМ сигналов идентична. Смесь сигнала и шума $z(t)$ фильтруется согласованным фильтром, а затем выделяется огибающая сигнала на выходе этого фильтра. Огибающая сравнивается с пороговым уровнем, величина которого при равных априорных вероятностях $P(u_1(t) = P(u_2(t))$ определяется отношением E_i / N_0 . Если эти вероятности не равны, пороговый уровень, согласно (7.29), изменится на $\ln[P(u_1(t)/P(u_2(t))]$.

При превышении порогового уровня в схеме с АМ сигналом принимается решение $b_i = 1$, в противном случае принимается решение $b_i = 0$.

При приеме ЧМ сигнала, если порог превышает в верхнем канале, то принимается решение $b_i^* = 1$, а если в нижнем, то — решение $b_i^* = 0$.

Некогерентный прием ОФМ сигнала оказывается возможным, если начальная фаза изменяется медленно и сохраняется практически постоянной в двух соседних единичных интервалах.

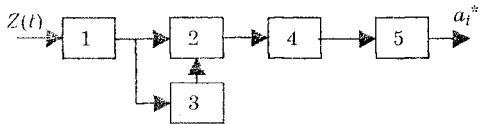


Рис.7.16. Схема оптимального некогерентного приема сигналов с ОФМ: 1 — согласованный фильтр; 2 — перемножитель; 3 — линия задержки на один единичный интервал; 4 — ФНЧ; 5 — решающее устройство

перемножителем и линией задержки.

Схема работает следующим образом. Отклик согласованного фильтра умножается на задержанный отклик, а результат фильтруется фильтром нижних частот. Знакопеременное напряжение с выхода ФНЧ сравнивается с пороговым уровнем. По результату сравнения принимается решение о переданном информационном символе.

7.3.2. Оптимальный прием при неопределенной фазе и амплитуде элементов сигнала

Задача синтеза оптимального демодулятора m дискретных сигналов с неопределенной фазой и амплитудой решается аналогично задаче синтеза сигналов с неопределенной фазой. Однако условия приема несколько отличаются. Математической моделью дискретно-непрерывного канала связи является модель гауссовского канала с замираниями сигнала. Сигнал на выходе канала флуктуирует как по начальной фазе, так и по амплитуде. Это приводит: а) к изменению выражения для функции правдоподобия; б) к изменению выражения для правила принятия решения.

Функция правдоподобия для этого случая равна

$$w[\bar{y} / u_i, k_i, \varphi_i] = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{N_0} \int_0^{T_H} z(t) - k_i v_i \cos(\omega_i t + \varphi_i) dt},$$

где k_i — коэффициент, характеризующий ослабление амплитуды v_i сигнала на входе канала связи.

Правило принятия решения равно:

Структурная схема оптимального демодулятора ОФМ сигнала приведена на рис.7.16. Она отличается от схемы на рис.7.6 тем, что в ней нет преобразователя относительной кодовой комбинации b_i^* в комбинацию a_i^* . Преобразование $a_i^* = f(b_i^*)$ обеспечивается двумя элементами:

$$\frac{\int_0^\infty \int_0^{2\pi} w(k_i) w(\bar{y} / u_i, k_i, \varphi_i) dk_i d\varphi_i}{\int_0^\infty \int_0^{2\pi} w(k_j) w(\bar{y} / u_i, k_j, \varphi_i) dk_j d\varphi_i} > \frac{P[u_j(t)]}{P[u_i(t)]}, j \neq i, i = 1, 2, \dots, m, \quad (7.31)$$

где $w(k_i)$ — плотность распределения коэффициента ослабления сигнала.

Принимая во внимание выражение (7.28), это правило можно переписать в виде

$$\frac{\int_0^\infty w(k_i) e^{-\frac{k_i^2 E_0}{N_0}} I_0\left(\frac{2k_i v_i}{N_0} M_i\right) dk_i}{\int_0^\infty w(k_j) e^{-\frac{k_j^2 E_0}{N_0}} I_0\left(\frac{2k_j v_j}{N_0} M_j\right) dk_j} > \frac{P[u_j(t)]}{P[u_i(t)]}, j \neq i, j = 1, 2, \dots, m. \quad (7.32)$$

Вычисляя интегралы в выражении (7.32) для заданной плотности распределения вероятности $w(k_i)$ можно определить структуру оптимального демодулятора. Она совпадает со структурой оптимального приемника дискретных сигналов с неопределенной начальной фазой. Изменяются только значения пороговых уровней на входах устройств сравнения.

7.4. Неоптимальные методы приема дискретных сообщений

Неоптимальные приемники обычно строятся так, чтобы при несколько меньшей помехоустойчивости получить простое техническое решение, следовательно, и меньшую стоимость аппаратуры, или еще какое-то преимущество.

Рассмотрим примеры приема двоичного частотно-модулированного сигнала на фоне нормального белого шума. Оптимальный прием реализуется на базе коррелятора (см. рис. 7.1 для $m = 2$) или согласованного фильтра (см. рис. 7.8). Для нормальной работы оптимального корреляционного приемника необходимо иметь источник опорных гармонических сигналов неравных несущих частот с непрерывной подстройкой их начальных фаз. Для аналогичной работы оптимального приемника с согласованными фильтрами требуется источник коротких периодически повторяющихся видеопульсов с непрерывной подстройкой их временного положения.

Для осуществления указанных подстроек в источнике опорных сигналов необходимо:

а) непрерывно измерять начальные фазы радиоимпульсов частотно-модулированного сигнала на фоне нормального белого шума;

б) по результатам измерений подстраивать начальные фазы радиоимпульсов опорного частотно-модулированного сигнала.

Для подстройки временного положения коротких видеоимпульсов на выходе источника этих импульсов необходимо:

а) непрерывно измерять временное положение единичных интервалов принимаемого частотно-модулированного сигнала на фоне нормального белого шума;

б) по результатам измерений подстраивать временное положение коротких видеоимпульсов.

Источники указанных сигналов оказываются сложными, следовательно, дорогими.

Высокая стоимость аппаратуры во многих случаях становится решающим фактором в выборе структур приемников. Схема оптимального приемника частично или полностью заменяется неоптимальной в зависимости от дополнительных требований на проектируемую аппаратуру.

Например, схему (рис.7.1, $m = 2$) можно полностью заменить схемой, (рис.7.12), в которой источник опорных сигналов хотя и сохраняется, но начальные фазы опорных радиосигналов на его выходах не подстраиваются. Это существенно упрощает и удешевляет приемник.

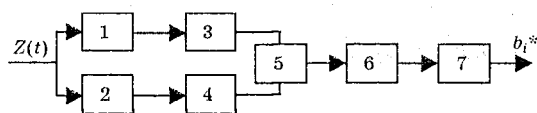


Рис.7.17. Схема неоптимального некогерентного приема ЧМ сигнала:

- 1,2 — полосовые фильтры;
3,4 — амплитудные детекторы;
5 — разностное устройство; 6 — ФНЧ;
7 — решающее устройство

Схему на рис.7.13 часто заменяют схемой (рис.7.17), что позволяет снизить стоимость аппаратуры за счет:

а) замены согласованных фильтров полосовыми;

б) введения последетекторной обработки сигналов.

При замене согласованного фильтра более простым полосовым коэффициент передачи последнего оказывается лишь частично согласованным со спектральной плотностью входного сигнала. За счет этого отношение сигнал/шум на выходе приемника уменьшается. Подобные фильтры называются *квазиоптимальными*.

При последетекторной обработке выборки высокочастотного заполнения радиоимпульса заменяются выборками его огибающей, что позволяет использовать более простой источник коротких видеоимпульсов со значительно большим разбросом их пе-

риода повторения. Можно вообще отказаться от этого источника, если исключить выборку огибающей и изменить правило принятия решения в решающем устройстве. Схема, в которой вместо согласованных включены полосовые фильтры и применяется последовательная обработка, приведена на рис.7.17. Принимаемый сигнал после его фильтрации полосовыми фильтрами разделяется на два. Огибающие этих сигналов выделяются с помощью амплитудных детекторов (рис.7.18, *г, д*) и затем вычитаются (рис.7.18, *е*). Далее можно было бы взять отсчетные значения полученной разности в моменты времени $t_i, i = 0, 1, 2, 3, \dots$ (рис.7.18, *ж*) и по их знаку производить оценку символов b_i^* кодового слова. За счет того, что скорость изменения огибающей около экстремумов функции (рис.7.18, *е*) во много раз меньше скорости изменения высокочастотного заполнения, в этом случае можно было бы существенно ослабить требования к стабильности длительности отсчетных интервалов Δt (рис.7.18, *ж*). Очевидно, что источник коротких видеопульсов был бы существенно упрощен.

В схеме (рис.7.17) реализован еще более выгодный вариант (с точки зрения аппаратных затрат), в котором источник коротких видеопульсов отсутствует, а оценка символа b_i^* кодового

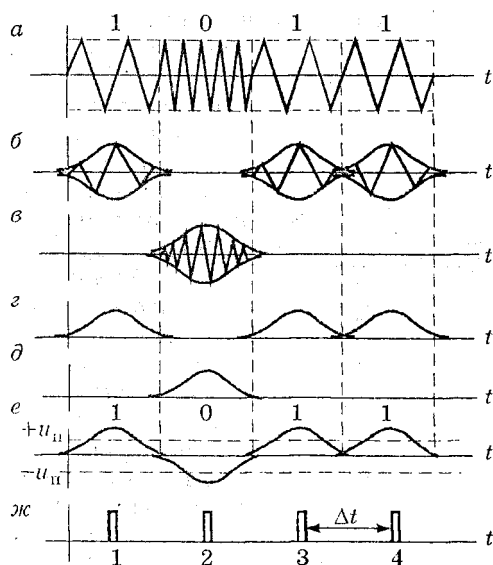


Рис.7.18. Временные диаграммы, поясняющие работу неоптимального приемника ЧМ сигнала

слова производится по знаку напряжения на выходе разностного устройства и ФНЧ.

Один из вариантов оценки символа иллюстрируется рис.7.18, *е*. Если напряжение на выходе ФНЧ пересекает заранее заданный положительный пороговый уровень $+u_{п}$, то на выходе решающего устройства формируется символ $b_i^* = 1$, а если напряжение пересекает отрицательный пороговый уровень $-u_{п}$, то символ $b_i^* = 0$.

Описанные выше замечания приводят:

а) к изменению алгоритма

обработки сигнала;

б) к изменению условий и ограничений накладываемых на сигнал;

в) к уменьшению помехоустойчивости приемника.

Потери возникают за счет:

а) неиспользования информации о фазе принимаемого сигнала (эта информация безвозвратно теряется при амплитудном детектировании);

б) неиспользования информации о положении максимума (минимума) напряжения на выходе ФНЧ;

в) замены согласованных фильтров полосовыми.

7.5. Цифровые методы обработки сигналов в приемнике

Основные принципы цифровой обработки сигналов, изложенные в разд.3.11, в полной мере относятся и к смеси сигнала и шума $z(t)$, действующей на входе приемника. Оптимальную обработку этой смеси можно осуществить: а) с использованием программных и аппаратных средств персональных ЭВМ (ПЭВМ) и микропроцессоров (МП); б) с использованием специализированных интегральных схем, ориентированных на построение цифровой обработки случайных сигналов и помех; в) с использованием микроэлектронной элементной базы общего назначения.

Выбор того или иного способа реализации определяется технико-экономическими соображениями.

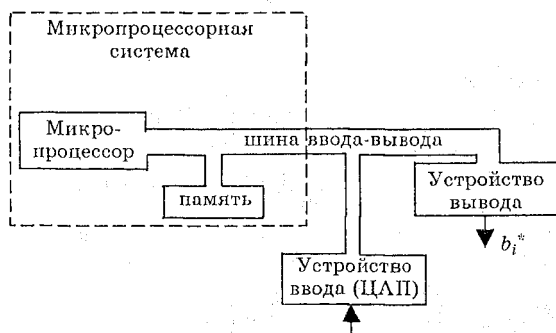


Рис.7.19. Структурная схема демодулятора сигнала, реализованного на МП

Ограничимся рассмотрением варианта оптимальной обработки дискретных сигналов с использованием МП. Структурная схема демодулятора приведена на рис.7.19. На его входе присутствует полезный дискретно-модулированный сигнал (таким сигналом может быть АМ, ФМ, ОФМ, ЧМ сигнал) на фоне нормального бе-

лого шума. В устройстве ввода производится аналого-цифровое преобразование смеси сигнала и шума в их выборочные значения, представленные в виде двоичных чисел. Выборочные значения берутся через интервал $\Delta t \approx 1/2F$, где F — верхняя граничная

частота спектра преобразуемого сигнала. С выхода аналого-цифрового преобразователя (АЦП) случайные двоичные числа через шину ввода-вывода поступают в память МП. Управляющая программа МП реализует заданный алгоритм оптимальной обработки.

В результате вычислений определяется оценка информационного элемента b_i^* , которая через устройство ввода-вывода поступает на вход последующего блока обработки. Синхронизация, необходимая для нормальной работы демодулятора, выполняется с помощью дополнительной управляющей программы, составленной на основе алгоритма работы системы синхронизации.

Наиболее просто и с наименьшими программными и аппаратными ресурсами реализуется алгоритм приема импульсов постоянного тока. Прием радиосигналов требует как большей памяти, так и быстродействия. Реализация демодулятора на МП при приеме радиосигналов может оказаться технически и экономически оправданной, если несущая частота сравнительно невелика.

Достоинством рассматриваемой системы является возможность легко изменять алгоритм обработки изменением управляющей программы, что делает ее универсальной для различных форм применяемых сигналов.

7.6. Методы приема дискретных сообщений в каналах с рассеянием

Если на вход канала с рассеянием подается радиосигнал с постоянной амплитудой и начальной фазой, то на его выходе образуется более сложный сигнал. Этот сигнал, равный сумме нескольких флуктуирующих по амплитуде и начальной фазе сигналов, также является флуктуирующим. Причиной флуктуаций являются нестационарность среды распространения и многопутевое распространение электромагнитных колебаний. Вероятность ошибок при приеме дискретных сообщений можно уменьшить с помощью разнесенного приема. Возможными разновидностями разнесенного приема являются: 1) прием отдельных, наиболее интенсивных по амплитуде составляющих сигнала; 2) одновременный прием всех составляющих сигнала.

При приеме отдельных сигналов используют две разнесенные в пространстве приемные антенны и переключатель (рис. 7.20), с помощью которого на вход приемника отбирают наиболее интенсивный сигнал. В случае, если принятые сигналы различаются по поляризации, они могут быть приняты двумя антеннами с разной поляризацией. Иногда кодовые комбинации передают одновременно на разных частотах (рис. 7.21), не используя два передатчика и два приемника. С помощью блока выбора

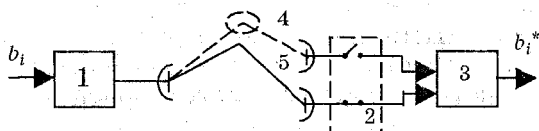


Рис.7.20. Разнесенный в пространстве прием сигнала: 1 — передатчик;
2 — переключатель; 3 — приемник;
4,5 — пути прохождения сигнала

принимает решение о принятом символе не по одному, а по нескольким сигналам, несущим одну и ту же информацию. Разновидностями этого вида приема являются: 1) прием нескольких задержанных во времени дискретных сигналов с применением метода накопления; 2) прием аналогичных сигналов с использованием корреляторов и согласованных фильтров.

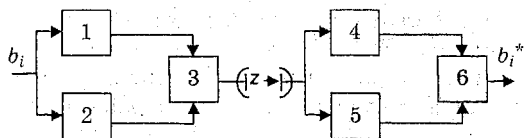


Рис.7.21. Прием сигналов, разнесенных по частоте: 1 — передатчик, работающий на частоте f_1 ; 2 — передатчик, работающий на частоте f_2 ; 3 — сумматор сигналов;
4,5 — приемники, работающие соответственно на частотах f_1 и f_2 ;
6 — блок выбора сигналов

с применением корреляторов и согласованных фильтров позволяет уменьшить влияние флуктуаций на результат обработки и увеличить помехоустойчивость приема за счет использования энергии всех составляющих сигнала. Наиболее полно возможности этого способа реализуются при передаче дискретной информации с помощью шумоподобных сигналов. Основное преимущество заключается в том, что задержанные во времени шумоподобные сигналы разделяются корреляторами и согласованными фильтрами.

Принцип разделения поясняется на примере приема двух лучей (рис.7.22). На входе согласованного фильтра присутствует сумма двух шумоподобных сигналов, одинаковых по форме, но отличающихся по амплитуде и задержке (рис.7.22, б). Каждый из этих сигналов имеет длительность равную длительности единичного интервала τ_n , одинаковые несущую частоту и закон модуляции фазы. Эта фаза равна или нулю или π в течение дли-

определяют наиболее интенсивный сигнал, который и определяет оценку передаваемой кодовой комбинации.

Сущность одновременного приема всех составляющих сигнала заключается в том, что демодулятор при-

Метод накопления связан с повторением передачи одного и того же символа несколько раз. Приемник принимает решение о переданном символе, суммируя (накапливая) результаты обработки смеси сигнала и шума в течение нескольких интервалов наблюдения.

Одновременный прием составляющих сиг-

тельности τ_0 одного элемента двоичной последовательности $a(t)$, кодированной кодом 111-1-11-1 (рис.7.22, а).

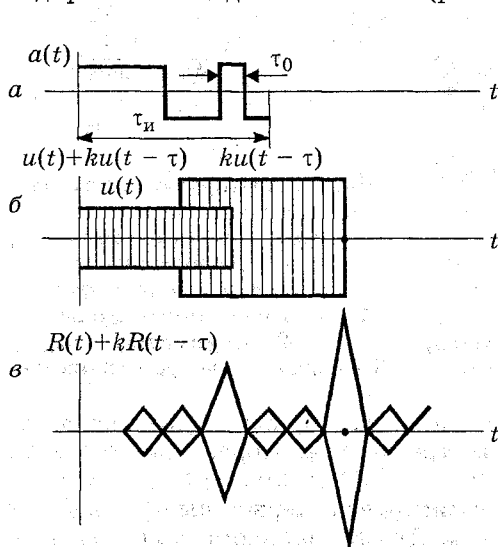


Рис.7.22. Разделение двух лучей с помощью шумоподобных сигналов

При передаче информации символу 1 соответствует присутствие шумоподобного сигнала в единичном интервале $0 - \tau_{и}$, а символу 0 — его отсутствие.

Разделение откликов согласованного фильтра, как это видно из рис.7.22, в, зависит как от величины задержки τ между сигналами, так и от длительности одного дискрета τ_0 шумоподобного сигнала. Чем меньше τ_0 , тем лучше различаются сигналы. Но это требует более широкой полосы канала связи.

Пусть на вход демодулятора поступает n одинаковых по форме шумо-

подобных сигналов, задержанных относительно друг друга во времени и имеется возможность их когерентного приема на фоне нормального белого шума.

Оптимальный алгоритм приема определяется неравенством (7.13). Если передатчик передает АМ шумоподобный сигнал, то в (7.13). $E_2 = 0$, а $E_1 = \sum_{i=1}^n E_{0i}$, где E_{0i} — энергия i -го входного ШС.

Оптимальный демодулятор вычисляет корреляционный интеграл

$$\begin{aligned} & \int z(t)[k_1 u(t - \tau_1) + k_2 u(t - \tau_2) + \dots + k_n u(t - \tau_n)] dt = \\ & = [\int z(t) k_1 u_1(t - \tau_1) dt + \int z(t) k_2 u_2(t - \tau_2) dt + \dots + \int z(t) k_n u(t - \tau_n) dt] > 0,5 E_1. \end{aligned}$$

и сравнивает его с пороговым уровнем, равным

$$0,5 E_1 = 0,5 \sum_{i=1}^n \int_{\tau_{и}}^{\tau_{п}} k_i^2 u^2(t - \tau_i) dt.$$

Структурная схема оптимального демодулятора, реализованная на корреляторах, приведена на рис.7.23. Для когерентного сложения необходимо обеспечить: а) совпадение начальных фаз опорных сигналов $u_i(t - \tau_i)$, $i = 1, 2, \dots, n$ с заданной точностью; б) совпадение задержки опорных с задержкой принимаемых сигналов.

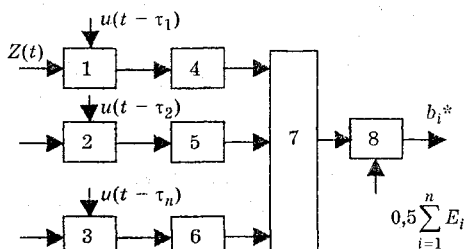


Рис.7.23. Структурная схема оптимального демодулятора когерентного шумоподобного сигнала:
1,2,3 — умножители;
4,5,6 — интеграторы; 7 — сумматор;
8 — решающее устройство

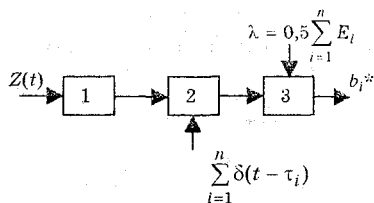


Рис.7.24. Структурная схема оптимального демодулятора шумоподобных сигналов, построенного на согласованном фильтре:
1 — согласованный фильтр;
2 — сумматор;
3 — сравнивающее устройство

Структурная схема оптимального демодулятора, реализованная на согласованном фильтре и сумматоре, приведена на рис.7.24. Сумматор в схеме на рис.7.24 выполняет ту же роль, что и в схеме на рис.7.23, — суммирует задержанные во времени отклики согласованного фильтра. Для выполнения этой операции необходимо: а) иметь в составе сумматора запоминающее устройство для хранения значений откликов согласованного фильтра; б) обеспечить когерентный отсчет откликов фильтра.

7.7. Способы борьбы с импульсными, сосредоточенными и межсимвольными помехами

Понятие о импульсных, сосредоточенных и межсимвольных помехах дано выше, в гл.5.

Для уменьшения влияния помех на качество передачи сообщений используют мероприятия по их подавлению в местах возникновения, защите каналов связи от их проникновения, решают задачу по выбору структуры системы связи и ее отдельных узлов и блоков при проектировании.

Для подавления помех в местах возникновения используют экранирование их источников; применяют искрогасящие конденсаторы, снижают уровень внеполосных излучений передатчиков и т.д.

Для защиты каналов от проникновения в них помех совершенствуют аппаратуру связи, уменьшая как влияние внешних помех, так и взаимные помехи между соседними каналами. В радиосвязи эту задачу решают рациональным выбором несущих частот передатчиков с учетом их расположения на местности и условий распространения радиоволн. Вопросы взаимного влияния различных систем друг на друга изучаются в теории электромагнитной совместимости [21].

При разработке мероприятий по подавлению помех учитывают рекомендации международного союза электросвязи: МСЭ-Т (сектор стандартизации электросвязи) и МСЭ-Р (сектор радиосвязи).

Рассмотрим кратко некоторые способы подавления помех, связанные с рациональным выбором сигналов и методами их обработки.

Из известных методов защиты от импульсных помех отметим метод академика А.Н.Шукина, основанный на ограничении импульсной помехи в широкополосном тракте приемника и узкополосной фильтрации ограниченного сигнала. Система, в которой применяется этот принцип, получила название ШОУ (широкополосной фильтр, ограничитель, узкополосный фильтр).

Другой известный способ борьбы с импульсными помехами основывается на запираании приемника на время их действия.

Для борьбы с узкополосными помехами применяют частотно-разнесенный прием, когда одинаковая информация передается в смежных частотных участках диапазона. При приеме к устройству обработки и решающему устройству подключаются только те частотные участки, которые не поражены сосредоточенными помехами. Межсимвольные помехи ослабляют выбором формы сигнала, структуры демодулятора и скорости передачи дискретной информации.

Достаточно эффективной мерой защиты от импульсной и сосредоточенной помех является разнесенный по времени и частоте прием сигнала. Из ветвей частотного разнесения выбирают те, которые не подвержены воздействию сосредоточенных помех, а из ветвей временного разнесения выбирают непораженные импульсными помехами ветви. Поскольку помехи появляются в заданных участках диапазона и интервалах времени случайным образом, выбор ветвей производится автоматически. При этом структура приемника адаптируется, приспосабливается к помеховой обстановке. Такой прием называется *адаптивным*.

Для борьбы с импульсными, сосредоточенными и межсимвольными помехами применяют также шумоподобные сигналы.

Применение шумоподобных сигналов в сочетании с использованием оптимальных методов их обработки на фоне помех позволяет успешно решать задачу борьбы с межсимвольными помехами. Принципы такой обработки рассмотрены выше, в разд.7.6. Использование шумоподобных сигналов и оптимальных методов их обработки на фоне нормального белого шума позволяет успешно вести борьбу также с импульсными и сосредоточенными помехами. Импульсная и сосредоточенная помеха при прохождении через коррелятор и согласованный фильтр существенно ослабляются. Отношение сигнал/помеха при этом возрастает в B раз, где B — база сигнала, определяемая формулой: (3.57).

Вместе с корреляционной обработкой и согласованной фильтрацией шумоподобного сигнала для борьбы с сосредоточенной

помехой в ряде случаев применяют и *предварительную фильтрацию* смеси сигнала шума и помехи режекторным фильтром. С помощью этого фильтра "вырезают" ту часть спектральных компонент, которые перекрываются узкополосными помехами, а по оставшейся части спектра шумоподобного сигнала восстанавливают переданную информацию.

7.8. Анализ помехоустойчивости систем связи с различными методами приема сигналов

7.8.1. Потенциальная помехоустойчивость двоичной системы при приеме элементов сигналов, известных точно

При приеме равновероятных сигналов $u_1(t)$, $u_2(t)$ на фоне нормального белого шума алгоритм обработки определяется неравенством:

$$\int_0^{\tau_n} [u_1(t) + N(t)] u_{\Delta}(t) dt > 0,5(E_1 - E_2), \quad (7.33)$$

где $u_{\Delta}(t) = u_1(t) - u_2(t)$.

Это неравенство получено на основе неравенства (7.13).

Производя элементарные преобразования, неравенство (7.33) перепишем в виде:

$$\int_0^{\tau_n} N(t) u_{\Delta}(t) dt > -0,5(E_1 + E_2) + \int_0^{\tau_n} u_1(t) u_2(t) dt.$$

При его выполнении принимается верное решение о принятом сигнале $u_1(t)$, а при невыполнении — возникает ошибка. Вероятность принятия ошибочного решения определяется вероятностью события, заключающегося в том, что значение корреляционного интеграла будет меньше некоторой заданной величины λ , т.е.

$$\xi = \int_0^{\tau_n} N(t) u_{\Delta}(t) dt < \lambda = -0,5(E_1 + E_2) + \int_0^{\tau_n} u_1(t) u_2(t) dt.$$

По условию реализация шума $N(t)$ подчиняется нормальному закону. Поскольку интегральное преобразование является линейным, то и случайная величина — ξ также подчиняется нормальному закону.

Среднее значение случайной величины ξ равно нулю, так как среднее значение белого шума равно нулю. Доказано [3], что дисперсия

$$D\{\xi\} = \sigma^2 = \frac{N_0}{2} F_3, \quad (7.34)$$

где

$$E_3 = \int_0^{\tau_n} [u_1(t) - u_2(t)]^2 dt.$$

С учетом этих замечаний выражение для плотности распределения вероятности случайной величины ξ можно записать в виде:

$$w(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2},$$

где σ^2 — дисперсия, определяемая выражением (7.34).

Зная закон распределения случайной величины ξ и уровень порога λ , нетрудно найти выражение для вероятности ошибки:

$$p = \int_{-\infty}^{\lambda} w(x) dx = \int_{-\infty}^{\lambda} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx.$$

Из возможных сигналов, известных точно рассмотрим двоичные дискретные АМ, ФМ и ЧМ сигналы с одинаковой энергией E . Для АМ сигнала

$$u_{\Delta}(t) = u_1(t) - u_2(t) = u_1(t), \text{ так как } u_2(t) = 0, \text{ поэтому } \lambda = -0,5E.$$

Для ЧМ сигнала

$$u_{\Delta}(t) = u_1(t) - u_2(t), \quad \int_0^{\tau_H} u_1(t)u_2(t)dt = 0, \quad \lambda = -E.$$

Для ФМ сигнала $u_1(t) = -u_2(t)$, $\lambda = -2E$.

Учитывая значения для уровней порога λ , используя функцию Крампа

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-y^2/2} dy,$$

запишем выражения для вероятностей ошибок:

Для АМ сигнала

$$E_s = E, \quad \sigma = \sqrt{N_0 E / 2},$$

$$P = \frac{1}{2} [1 - \Phi(h / \sqrt{2})], \quad (7.35)$$

где $h = \sqrt{E / N_0}$;

для ЧМ сигнала

$$E_s = 2E, \quad \sigma = \sqrt{N_0 E};$$

$$P = \frac{1}{2} [1 - \Phi(h)]; \quad (7.36)$$

для ФМ сигнала

$$E_s = 4E, \quad \sigma = \sqrt{2N_0 E};$$

$$p = \frac{1}{2} [1 - \Phi(\sqrt{2}h)]. \quad (7.37)$$

Анализируя выражения для вероятности ошибки, приходим к выводу, что наилучшие результаты по критерию минимальной вероятности ошибки обеспечиваются для ФМ сигналов. Переход от системы с ФМ сигналами к системе ЧМ и АМ сигналов ухудшает помехоустойчивость. Для получения одинаковых вероятностей ошибки в системе с ЧМ сигналами нужно увеличить энергию сигналов в 2 раза, а в системе с АМ сигналами — в 4 раза. Следовательно, при использовании ФМ сигналов передатчик будет излучать наименьшую среднюю мощность по сравнению с сигналами ЧМ и АМ.

Выражение для вероятности ошибки при приеме известного точно ОФМ сигнала найдем на основании оценки вероятностей следующих двух несовместимых событий: а) знак данного элемента принят ошибочно, а знак предыдущего верно; б) знак данного элемента принят верно, а предыдущего — ошибочно. Каждое из этих событий имеет вероятность $p_{\text{фм}}(1 - p_{\text{фм}})$. Поэтому

$$p_{\text{офм}} = 2p_{\text{фм}}(1 - p_{\text{фм}}) = 0,5[1 - \Phi^2(\sqrt{2}h)]. \quad (7.38)$$

Пусть требуется воспроизводить символы с весьма малой ошибкой. В этом случае $p_{\text{фм}} \ll 1$ и

$$p_{\text{офм}} = 2p_{\text{фм}} = 1 - \Phi(\sqrt{2}h). \quad (7.39)$$

"Платой" за устранение обратной работы, как следует из последнего выражения, является удвоение вероятности ошибки по сравнению со случаем приема ФМ сигналов.

7.8.2. Помехоустойчивость приема m сигналов, известных точно

Вероятность принятия ошибочного решения оптимальным демодулятором m сигналов, известных точно:

$$p_{\text{ош}} = 1 - p_n,$$

где p_n — вероятность принятия правильного решения.

Вероятность p_n определяется вероятностью события, заключающегося в том, что значения случайных величин ξ_j , вычисленных корреляторами с номерами $j \neq i$ в схеме, приведенный на рис. 7.1, будут меньше значения $\xi_i = z$, вычисленного коррелятором с номером i .

Пусть сигналы $u_i(t)$ ортогональны и имеют одинаковые энергии. Можно доказать, что вследствие ортогональности сигналов, случайные величины ξ_i , $i = 1, 2, \dots, m$ статистически независимы [29]. Поэтому вероятность того, что все $(m - 1)$ случайные величины окажутся меньше z , равна

$$[p(\xi_j < z)]^{m-1}, j \neq i.$$

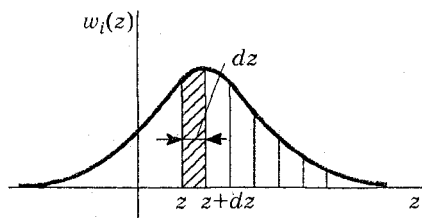


Рис.7.25. Плотность вероятности случайной величины

Вероятность того, что случайная величина ξ_i находится в пределах от z до $z + dz$, равна

$$w_i(z)dz,$$

где $w_i(z)$ — плотность распределения вероятности случайной величины ξ_i (рис.7.25).

Величина ξ_i независима по отношению ко всем остальным величинам ξ_j , $j \neq i$, следовательно

вероятность события, заключающегося в том, что $\xi_j < z$ ($j \neq i$), а ξ_i находится в интервале dz , равна

$$dP = [P(\xi_j < z)]^{m-1} w_i(z) dz.$$

Так как случайная величина ξ_i может находиться в пределах любого бесконечно малого значения dz (рис.7.25), то вероятность того, что значения случайных величин ξ_j , $j \neq i$ будут меньше значения $\xi_i = z$, определится выражением

$$p_n = \int_{-\infty}^{\infty} dp = \int_{-\infty}^{\infty} [P(\xi_j < z)]^{m-1} w_i(z) dz. \quad (7.40)$$

При условии приема сигналов на фоне нормального белого шума с нулевым средним значением плотность распределения вероятности $w_i(z)$ подчиняется нормальному закону со средним значением, равным энергии сигнала E и дисперсией равной $N_0 E / 2 = \sigma^2$, т.е.

$$w_i(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(z-E)^2 / 2\sigma^2}. \quad (7.41)$$

Случайные величины ξ_i , $j \neq i$ также подчинены нормальному закону с нулевым средним значением и дисперсией $N_0 E / 2 = \sigma^2$.

Поэтому вероятность

$$P(\xi_i < z) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^z e^{-x^2 / 2\sigma^2} dx. \quad (7.42)$$

Подставляя (7.41), (7.42) в (7.40), получим:

$$p_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[1 - V \left(\sqrt{2E / N_0} + y \right) \right]^{n-1} e^{-y^2 / 2} dy. \quad (7.43)$$

где

$$V(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt.$$

Полученный интеграл можно вычислить только приближенными методами.

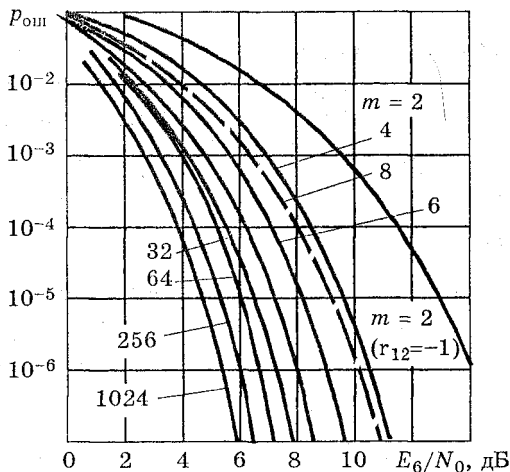


Рис.7.26. Зависимость вероятности ошибки от отношения h^2 при приеме m ортогональных сигналов, известных точно; $E_b = E/\log_2 m$ — энергия, затрачиваемая на один бит информации

Результаты расчетов приведены на рис.7.26 в виде графиков зависимости

$$1 - p_n = p_{ош} = f(h_0^2),$$

где $h_0^2 = h^2 / \log_2 m$.

Анализ этих зависимостей позволяет сделать следующий вывод.

Системы ортогональных сигналов с $m > 2$ при одинаковой скорости передачи информации и одинаковой вероятности ошибки позволяют обеспечить существенный выигрыш в энергии сигнала по сравнению с двоичными сигналами. Например, при $m = 32$ и $p_{ош} = 10^{-5}$ выигрыш равен двум.

Это дает возможность упростить передатчик системы за счет уменьшения его мощности. Однако сложность демодулятора возрастает (увеличивается число корреляторов). Усложняется при этом и канал связи (из-за увеличения ширины спектра сигнала). При относительно больших отношениях сигнал/шум формулу (7.43) можно привести к следующему приближенному выражению

$$p_n \cong 1 - \frac{m-1}{2} [1 - \Phi(h)].$$

Отсюда выражение для вероятности ошибки

$$p_{ош} \cong \frac{m-1}{2} [1 - \Phi(h)]. \quad (7.44)$$

При $m = 2$ последняя формула переходит в формулу (7.36).

7.8.3. Потенциальная помехоустойчивость оптимального приемника ЧМ сигналов с неизвестной начальной фазой

Рассмотрим прием ЧМ сигналов с неизвестной начальной фазой (см. рис. 7.15). ЧМ сигналы являются ортогональными и для них выполняется условие:

$$\int_0^{\tau_H} u_1(t)u_2(t)dt = 0. \quad (7.45)$$

Рассмотрим случай, когда на входе присутствует сигнал $u_1(t)$. Отклик первого согласованного фильтра в момент окончания первого единичного интервала не равен нулю (см. рис. 7.1, б), а отклик второго согласованного фильтра в этот же момент времени равен нулю, так как ЧМ сигналы ортогональны и отклик второго согласованного фильтра на сигнал $u_1(t)$, определяемый уравнением (7.45), равен нулю.

Огибающая смеси сигнала и шума на выходе первого детектора больше огибающей смеси сигнала и шума на выходе второго тот же момент времени τ_H , т.е.

$$A_1 M_1 = \xi_1 > A_2 M_2 = \xi_2.$$

При выполнении последнего неравенства демодулятор принимает правильное решение, а при его невыполнении возникает ошибка.

Пусть случайные величины ξ_1, ξ_2 определяются двумерной плотностью вероятности $w(\vartheta_1, \vartheta_2)$. Зная эту плотность и условие, при котором возникает ошибка, можно вычислить ее вероятность:

$$P(u_2 / u_1) = \int_0^\infty \int_{\vartheta_1}^\infty w(\vartheta_1, \vartheta_2) d\vartheta_1 d\vartheta_2 = \int_0^\infty w(\vartheta_1) d\vartheta_1 \int_{\vartheta_1}^\infty w(\vartheta_2 / \vartheta_1) d\vartheta_2 = \int_0^\infty dP. \quad (7.46)$$

Вероятность события, заключающегося в том, что реализация огибающей смеси сигнала и шума на выходе первого канала меньше реализации огибающей шума на выходе второго канала

$$dP = w(\vartheta_1) d\vartheta_1 \int_{\vartheta_1}^\infty w(\vartheta_2 / \vartheta_1) d\vartheta_2.$$

Вероятность ошибки определяется всеми подобными событиями. Плотность вероятности $w(\vartheta_1)$ определяется формулой (4.44), а плотность $w(\vartheta_2 / \vartheta_1)$ — формулой (4.43). Входящая в эти

формулы дисперсия $\sigma^2 = 0,5 N_0 E$, где $E = \int_0^\infty u_i^2(t) dt$ — энергия сигнала.

Подставляя формулы (4.43), (4.44) в формулу (7.46) и производя вычисления, получим выражение для вероятности ошибки

$$P(u_2 / u_1) = 0,5e^{-0,5h^2}, \quad (7.47)$$

где $h^2 = E/N_0$ — отношение энергии элемента к спектральной плотности мощности шума.

Аналогичное выражение имеет и вероятность $P(u_1/u_2)$. Окончательное выражение для вероятности ошибки при приеме равновероятных ЧМ сигналов со случайной начальной фазой

$$p = P(u_1)P(u_2 / u_1) + P(u_2)P(u_1 / u_2) = 0,5e^{-0,5h^2}. \quad (7.48)$$

Эта формула справедлива и для других ортогональных сигналов с одинаковой энергией.

7.8.4. Потенциальная помехоустойчивость оптимального приемника АМ сигнала с неизвестной начальной фазой

В данном случае вероятность ошибки

$$\begin{aligned} p &= P(u_1)P(u_2 / u_1) + P(u_2)P(u_1 / u_2) = P(1)P(u_2 / 1) + P(0)P(u_1 / 0) = \\ &= P(1) \int_0^{u_{\text{п}}} w(\vartheta / 1) d\vartheta + P(0) \int_{u_{\text{п}}}^{\infty} w(\vartheta / 0) d\vartheta, \end{aligned}$$

где $P(u_1) = P(1)$; $P(u_2) = P(0)$;

$w(\vartheta / 1)$ — плотность распределения вероятности огибающей на выходе оптимального приемника (амплитудного детектора) в момент времени $t = \tau_{\text{и}}$, совпадающий с окончанием единичного интервала (эта плотность определяется формулой (4.44)); $w(\vartheta / 0)$ — аналогичная плотность распределения вероятности, определяемая формулой (4.43); $u_{\text{п}}$ — пороговый уровень.

Для $P(1) = P(0) = 0,5$ выражение для вероятности ошибки [3]:

$$p = \frac{1}{2} \left\{ \int_0^{u_{\text{п}}} \frac{2\vartheta}{EN_0} \exp \left[-\frac{\vartheta^2 + E}{EN_0} \right] I_0 \left(\frac{2\vartheta}{N_0} \right) d\vartheta + \exp \left[-\frac{u_{\text{п}}^2}{EN_0} \right] \right\}. \quad (7.49)$$

Минимальная вероятность ошибки получается при оптимальном значении $u_{\text{п}}$. Оно находится из уравнения

$$dp / du_{\text{п}} = 0.$$

Можно показать, что

$$u_{\text{п опт}} = \begin{cases} E/2 & \text{при больших отношениях сигнал/шум,} \\ \sqrt{EN_0} & \text{при малых отношениях сигнал/шум.} \end{cases}$$

Подставляя, оптимальное значение порогового уровня в (7.49), получим окончательное выражение для вероятности ошибки при большом отношении сигнал/шум.

$$p = \frac{1}{2} e^{-h^2/4}. \quad (7.50)$$

7.8.5. Потенциальная помехоустойчивость приема дискретных сообщений при замираниях сигнала

Определим помехоустойчивость приема двоичных ортогональных с активной паузой и одинаковой энергией сигналов при медленных замираниях, когда на протяжении единичного интервала амплитуда остается постоянной, но меняющейся случайным образом от интервала к интервалу.

Вероятность ошибки

$$p = \int_0^{\infty} P(k) w(k) dk,$$

где $P(k)$ — вероятность ошибки при приеме сигнала с фиксированной амплитудой, следовательно и с фиксированным k и случайной начальной фазой; k — множитель, определяющий ослабление или усиление сигнала; $P(k)$ — плотность распределения вероятности. Вероятность $P(k)$ определяется формулой (7.48). Считая, что плотность распределения вероятности $w(k)$ подчиняется рэлеевскому закону (5.5) и, принимая во внимание, что между значениями h , k и математическими ожиданиями их квадратов существует зависимость $h^2 = k^2 \bar{h}^2 / \bar{k}^2$, найдем выражение для вероятности ошибки

$$\begin{aligned} p &= \int_0^{\infty} 0,5 \exp(-0,5 k^2 \bar{h}^2 / \bar{k}^2) \frac{2k}{\bar{k}^2} \exp(-k^2 / \bar{k}^2) dk = \\ &= \int_0^{\infty} \left(\frac{k}{\bar{k}^2} \right) \exp \left[-k^2 (1 + 0,5 \bar{h}^2) / \bar{k}^2 \right] dk = \int_0^{\infty} x \exp[-x^2 (1 + 0,5 \bar{h}^2)] dx = 1 / (\bar{h}^2 + 2), \end{aligned}$$

где $x = k^2 / \bar{k}^2$.

Если плотность распределения амплитуды подчиняется обобщенному рэлеевскому закону (5.6), то вероятность ошибки

$$p = \frac{1 + q^2}{2 + 2q^2 + \bar{h}^2} \exp \left(- \frac{\bar{h}^2 q^2}{2 + 2q^2 + \bar{h}^2} \right), \quad (7.51)$$

где q^2 — отношение мощностей регулярной и флуктуирующей составляющих.

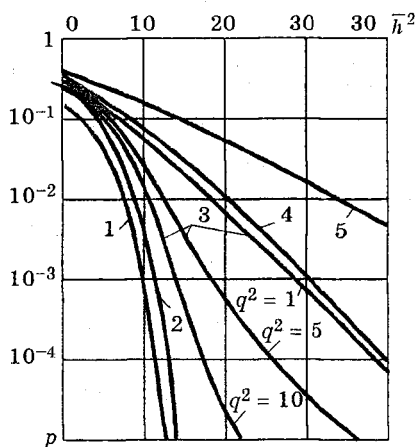


Рис.7.27. Зависимости вероятности ошибки в двоичной системе (ортогональной,

с активной паузой) от $\bar{h}^2$:

- 1 — когерентный прием;
- 2 — некогерентный прием;
- 3 — обобщенные рэлеевские замирания; 4 — рэлеевские замирания; 5 — односторонне-нормальные замирания

двоичных каналов не передается. При h^2 , стремящихся к бесконечности, вероятность ошибки в любой системе стремится к нулю. Следовательно, увеличением энергии (мощности) сигнала можно достичь сколь угодно малой вероятности ошибки.

7.8.6. Помехоустойчивость приема оптических сигналов

Принцип приема оптических дискретных сигналов в цифровых волоконно-оптических системах передачи поясняется рис.7.28 и 7.29.

При передаче единичного символа кодовой последовательности (рис.7.28, а) в пределах единичного интервала t_0 , равного периоду тактовой частоты, формируется оптический импульс длительностью $t_0/2$. При передаче нулевого символа кодовой последовательности оптический импульс не формируется (рис.7.28, б).

На рис.7.27 показаны зависимости $p = f(\bar{h}^2)$ для различных способов приема, построенные по приведенным выше формулам.

Они позволяют сравнить различные способы приема по помехоустойчивости. Например, при одинаковой энергии сигнала и мощности шума на входе ($h^2 = 10$) в канале с постоянными параметрами (когерентный прием) обеспечивается наименьшая вероятность ошибки (10^{-3}). В канале с неопределенной фазой ошибка больше. В канале со случайной фазой и амплитудой ошибка становится еще больше и т.д.

При h^2 стремящемся к нулю вероятности ошибок принимают значения 0,5. Это означает, что из-за шума информация ни в одном из

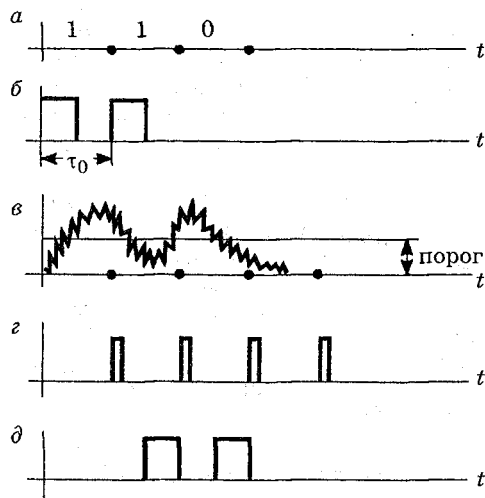


Рис.7.28. Временные диаграммы, поясняющие прием оптических дискретных сигналов

В процессе распространения по оптическому кабелю световые импульсы ослабляются и искажаются.

За счет дисперсии длительность принимаемого импульса увеличивается, а вследствие ограниченной полосы пропускания приемного оптоэлектронного модуля происходит "затягивание" его переднего и заднего фронтов.

Преобразование оптических сигналов в электрический ток в приемном оптоэлектронном модуле сопровождается флуктуационными шумами.

Составляющими этих шумов являются *дробовой*

(квантовый) шум, обусловленный вероятностным характером преобразования оптического сигнала в электрический ток и *тепловой шум* и шум усилителя в приемнике, обусловленные тепловым движением молекул и электронов.

Искаженный средой распространения и шумами сигнал на выходе приемного оптоэлектронного модуля показан на рис.7.28, в.

Обработка принимаемого сигнала сводится: к фильтрации сигнала фильтром нижних частот (этот фильтр входит в состав приемноэлектронного модуля и к стробированию в заданные моменты времени; сравнению полученных выборочных значений с пороговым уровнем.

Стробующие короткие импульсы, вырабатываемые устройством тактовой синхронизации, показаны на рис.7.28, г.

Если пороговый уровень превышает, то формирователь сигнала формирует импульсы тока (напряжения), показанные на рис.7.28, д.

Как следует из описания принципа приема оптических дискретных сигналов, в приведенной на рис.7.29 схеме решается задача обнаружения.

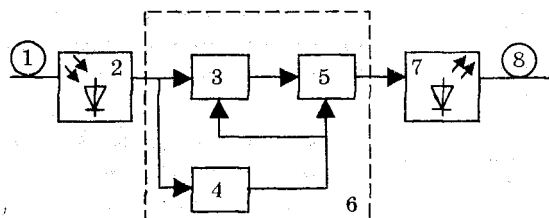


Рис.7.29. Структурная схема оптоволоконной системы связи: 1,8 — оптический кабель; 2 — приемный оптоэлектронный модуль; 3 — решающее устройство; 4 — устройство тактовой синхронизации; 5 — формирователь сигнала; 6 — электронный регенератор; 7 — передающий оптоэлектронный модуль

Пусть "квантовый" шум пренебрежимо мал по сравнению с тепловым, а априорные вероятности равны, т.е. $P(0) = P(1) = 0,5$. Плотность распределения вероятности теплового шума подчиняется нормальному закону с нулевым средним значением:

$$w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2},$$

где σ^2 — дисперсия

флуктуаций шума на выходе фильтра приемника оптоэлектронного модуля. Эта плотность определяет непрерывную случайную величину ξ на выходе стробирующего устройства, когда передаются символ "0" (кривая 1 на рис.7.30).

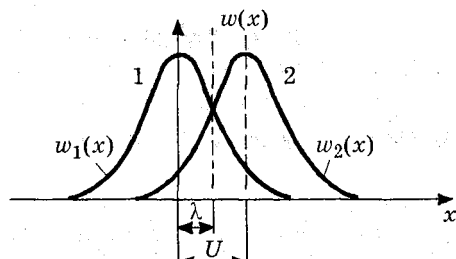


Рис.7.30. К определению вероятности ошибки при приеме оптических сигналов

Случайная величина ξ на выходе стробирующего устройства, когда передается символ "1" (кривая 2), также определяется нормальным законом с той же дисперсией σ^2 и средним значением, равным амплитуде U принимаемого сигнала.

Если уровень порога равен λ , то вероятность ошибки при приеме оптических сигналов определяется выражением

$$p = 0,5P(1/0) + 0,5P(0/1) = 0,5 \left(\int_{\lambda}^{\infty} w_1(x) dx + \int_{-\infty}^{\lambda} w_2(x) dx \right), \quad (7.52)$$

где $P(1/0) = \int_{\lambda}^{\infty} w_1(x) dx$ — вероятность того, что случайная величина ξ превысит уровень порога равный λ , при условии приема символа "0"; $P(0/1) = \int_{-\infty}^{\lambda} w_2(x) dx$ — вероятность того, что случайная величина ξ не превысит порог, равный λ при условии приема символа "1".

Как видно из рис.7.29, если $\lambda = U/2$, то

$$\begin{aligned} \int_{\lambda}^{\infty} w_1(x) dx &= \int_{-\infty}^{\lambda} w_2(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{\lambda}^{\infty} e^{-x^2/2\sigma^2} dx = \frac{1}{2} \left[\frac{2}{\sqrt{2\pi}} \int_{\lambda/\sigma}^{\infty} e^{-z^2/2} dz \right] = \\ &= 0,5 \left[1 - \Phi\left(\frac{\lambda}{\sigma}\right) \right] = 0,5 \left[1 - \Phi\left(\frac{U}{2\sigma}\right) \right]. \end{aligned}$$

где $\Phi(x)$ — функция Крампа.

Подставляя (7.53) в (7.52), получим выражение для вероятности ошибки:

$$p = 0,5 \left[1 - \Phi\left(\frac{U}{2\sigma}\right) \right]. \quad (7.54)$$

Сравнивая схему (рис.7.29) обработки оптического сигнала с оптимальным обнаружителем бинарного сигнала (см. рис.7.10), отметим следующее.

Схема на рис.7.29 весьма близка к оптимальной, т.к. содержит те же устройства, что и в схеме на рис.7.10. Степень близости схем определяется приближением линейного фильтра, содержащегося в приемном оптоэлектронном модуле, к согласованному фильтру.

Если бы характеристики линейного фильтра удалось получить точно такие же, что и у согласованного фильтра, то в этом случае дисперсия оказалась бы равной $\sqrt{N_0 E}/2$, уровень сигнала U равнялся бы энергии E , а выражение для вероятности ошибки (7.54) совпало бы с выражением (7.35).

Рассмотрим случай приема оптических сигналов, когда тепловые шумы отсутствуют и единственным источником шума является "квантовый" шум.

При приеме символа "1" фотодиод возбуждается световым стационарным потоком и на его выходе генерируются пары носителей заряда. Число k созданных пар является случайной величиной, подчиняющейся пуассоновскому распределению вероятности:

$$p(k/N) = \frac{N^k \exp(-N)}{k!},$$

где N — среднее число пар.

Среднеквадратичное число пар носителей заряда равно среднему числу N .

В соответствии с числом пар носителей заряда будет изменяться и ток на выходе фотодиода.

При передаче символа "0" световая энергия отсутствует, а сигнал на выходе фотодиода не генерируется. Это значит, что ошибка в регистрации символа "0" в этом случае будет равна нулю. Ошибка возникает при передаче символа "1", когда при

подаче на фотодиод светового импульса на его выходе сигнал не генерируется. Число созданных пар носителей заряда в этом случае равно нулю. Вероятность этого события определит вероятность ошибки:

$$p = \frac{1}{2} \left[\frac{N^0 \exp(-N)}{0!} + 0 \right] = \frac{1}{2} \exp(-N).$$

Полученная формула определяет помехоустойчивость приема оптических сигналов при наличии только одного квантового шума.

Задаваясь величиной ошибки p и, принимая во внимание следующую зависимость между средним числом пар носителей заряда и средней оптической энергией E :

$$N = \eta \frac{E}{E_{\Phi}} = \eta \frac{\bar{E} \lambda}{hC},$$

где η — квантовая эффективность взаимодействия, E_{Φ} — энергия фотона, λ — длина волны, h — постоянная Планка, C — скорость света, можно определить минимальную среднюю энергию сигнала на входе фотоприемника, которая будет характеризовать абсолютный квантовый предел детектируемости светового сигнала.

7.9. Оптимальный прием непрерывных сообщений

7.9.1. Показатели качества передачи

Непрерывное сообщение $a(t)$ является реализацией некоторого непрерывного случайного процесса, изменяющегося во времени. Сигнал, совпадающий с точностью до постоянной величины с этим сообщением, называется первичным сигналом. Первичный сигнал $b(t)$ передается по каналу связи или непосредственно, как, например, в телефонии, или преобразуется посредством модуляции, как в радиовещании и телевидении.

В первом случае

$$b(t) = ka(t),$$

где k — коэффициент пропорциональности.

Во втором случае первичный сигнал преобразуется в сигнал $u[t, p(t)]$, зависящий от времени t и от модулируемого параметра $p(t)$, причем $p(t) = cb(t)$, где c — коэффициент пропорциональности. Для простоты примем $c = 1$.

При передаче по каналу связи на сигнал накладывается помеха $N(t)$. Для аддитивной помехи сумма сигнала и шума

$$z(t) = u[t, b(t)] + N(t).$$

Задачей приемника является выделение первичного сигнала с минимальной ошибкой.

Мерой помехоустойчивости, как упоминалось во введении, является среднеквадратичное отклонение или средний квадрат ошибки. Разность $\varepsilon(t) = b^*(t) - b(t)$ является помехой на выходе приемника. Величина $M\{\varepsilon^2(t)\} = P_\varepsilon$ является мощностью помехи. Мощность первичного сигнала (сообщения) $b(t)$ равна $P_b = M[b^2(t)]$. Вместо среднеквадратичного отклонения часто используют другой показатель — *отношение мощности сигнала к мощности шума* (на выходе приемника)

$$\rho_{\text{вых}} = P_b / P_\varepsilon.$$

Если отношение сигнал/шум на входе приемника равно $\rho_{\text{вх}} = P_c / P_{\text{ш}}$, то отношение

$$g = \rho_{\text{вых}} / \rho_{\text{вх}} = \frac{P_b}{P_\varepsilon} / \frac{P_c}{P_{\text{ш}}} = \frac{P_b}{P_c} / \frac{P_\varepsilon}{P_{\text{ш}}} \quad (7.55)$$

характеризует выигрыш (или проигрыш) в помехоустойчивости за счет применения того иного способа обработки сигнала.

Полезным показателем является также "*обобщенный выигрыш*" равный

$$g' = \frac{\rho'_{\text{вых}}}{\rho'_{\text{вх}}} = \frac{g}{\alpha}, \quad (7.56)$$

где $\rho'_{\text{вых}} = \frac{P_b}{P_\varepsilon / F_c}$, F_c — ширина спектра сообщения;

$\rho'_{\text{вх}} = \frac{P_c}{P_{\text{ш}} / F}$, F — ширина спектра сигнала; $\alpha = F / F_c$.

7.9.2. Прием отдельных значений непрерывных сообщений.

Критерии оптимальности

Впервые эта задача была решена В.А.Котельниковым [1] при следующих исходных данных.

Сообщение является непрерывной случайной величиной, которая в интервале наблюдения $[0, T]$ постоянна и от одного интервала к другому меняется случайным образом. Плотность распределения вероятности случайной величины в пределах возможного ее изменения является постоянной, т.е. $w(x) = \text{const}$. Отношение сигнал/шум велико.

Оптимальный приемник вычисляет апостериорную плотность распределения вероятности и выдает то значение сообщения, при котором апостериорная плотность максимальна. Этот критерий называется *критерием максимума апостериорной вероятности*.

Апостериорная плотность распределения вероятности определяется формулой

$$w_{\bar{y}}(x) = cw(x)w(\bar{y}/x), \quad (7.57)$$

где c — постоянная величина, $w(\bar{y}/x)$ — функция правдоподобия.

Постоянная величина находится из условия нормировки:

$$\int_X w_{\bar{y}}(x) dx = 1,$$

где X — область возможных значений x .

Интегрируя выражение (7.57), получим

$$\int_X w_{\bar{y}}(x) dx = c \int_X w(x)w(\bar{y}/x) dx = 1.$$

Из последнего уравнения находим

$$c = \frac{1}{\int_X w(x)w(\bar{y}/x) dx}.$$

Функция правдоподобия определяется формулой:

$$w(\bar{y}/x) = \left[\frac{1}{\sqrt{2\pi}\sigma} \right]^n e^{-\frac{1}{N_0} \int_0^T [z(t) - u_x(t)]^2 dt}, \quad (7.58)$$

где $u_x(t)$ — сигнал, все параметры которого известны точно, за исключением модулируемого сообщением параметра.

Если априорная плотность распределения $w(x) = \text{const}$, то как видно из формулы (7.57), апостериорная плотность с точностью до постоянной величины совпадает с функцией правдоподобия.

Поэтому эквивалентным критерием оптимальности является критерий максимума функции правдоподобия:

$$w(\bar{y}/x) = \max.$$

Оценка значения непрерывной случайной величины определяется из уравнения

$$\frac{dw(\bar{y}/x)}{dx} = 0.$$

Эта оценка называется максимально правдоподобной.

Часто в качестве критерия оптимальности выбирают критерий минимума среднеквадратичной ошибки

$$M\{\varepsilon^2(x)\} = \int_X (x - x^*)^2 w(x/\bar{y}) dx = \min, \quad (7.59)$$

где X — заданная область интегрирования.

Оптимальная оценка находится из уравнения:

$$\frac{dM\{\varepsilon^2(x)\}}{dx^*} = 0.$$

Производя необходимые вычисления с учетом того, что

$$\int_X w(x / \bar{y}) dx = 1,$$

получим

$$x^* = \int_X x w(x / \bar{y}) dx.$$

Оптимальной оценкой в этом случае является *математическое ожидание апостериорного распределения*.

Обобщенной оценкой является *байесовская оценка*:

$$M\{L(x - x^*)\} = \int_X L(x - x^*) w(x / \bar{y}) dx,$$

где $L(x - x^*)$ — некоторая весовая функция потерь.

Критерий

$$\int_X L(x - x^*) w(x / \bar{y}) dx = \min$$

называется *критерием минимума среднего риска*.

При $L(x - x^*) = (x - x^*)^2$ байесовская оценка переходит в математическое ожидание квадрата разности (7.59), а критерий минимума среднего риска — в критерий минимума среднеквадратичной ошибки.

Если отношение сигнал/шум велико, а апостериорное распределение представляет симметричную кривую относительно значения x^* , все приведенные выше критерии совпадают.

Найдем значение сообщения, при котором функция правдоподобия максимальна. С этой целью вычислим производную выражения (7.58) и приравняем ее к нулю. В результате получим уравнение

$$\frac{\partial \xi(x)}{\partial x} = 0,$$

где $\xi(x) = \frac{1}{N_0} \int_0^T [z(t) - u_x(t)]^2 dt.$

Наивероятное значение сигнала $u_{x_H}(t)$, следовательно, и наивероятное значение x_H , находятся из уравнения

$$\frac{2}{N_0} \int_0^T [z(t) - u_x(t)] \frac{\partial u_x(t)}{\partial x} dt = 0. \quad (7.60)$$

При x близком к x_H выражение для сигнала можно записать в виде двух слагаемых ряда:

$$u_x(t) = u_{x_H}(t) + (x - x_H) \left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H},$$

где $\left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H}$ — значение производной в точке $x = x_H$.

Подставляя последнее выражение в (7.58) и учитывая условие (7.60), получим

$$\begin{aligned} w(\bar{y}/x) &= \left[\frac{1}{\sqrt{2\pi\sigma}} \right]^n e^{-\frac{1}{N_0} \int_0^T \left[z(t) - u_{x_H}(t) - (x - x_H) \left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} \right]^2 dt} = \\ &= \left[\frac{1}{\sqrt{2\pi\sigma}} \right]^n e^{-\frac{1}{N_0} \int_0^T [z(t) - u_{x_H}(t)]^2 dt} e^{-\frac{2}{N_0} (x - x_H) \int_0^T [z(t) - u_{x_H}(t)] \left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} dt} \times \\ &\times e^{-\frac{1}{N_0} (x - x_H)^2 \int_0^T \left[\left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} \right]^2 dt} = \left[\frac{1}{\sqrt{2\pi\sigma}} \right]^n e^{-\frac{1}{N_0} \int_0^T [z(t) - u_x(t)]^2 dt} \times \\ &\times e^{-\frac{1}{N_0} (x - x_H)^2 \int_0^T \left[\left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} \right]^2 dt} = c_1 c_2 e^{-\frac{(x - x_H)^2}{2\sigma^2}} = c \cdot e^{-\frac{(x - x_H)^2}{2\sigma^2}}, \end{aligned}$$

где c_1, c_2, c — независимые от аргумента x постоянные величины, причем

$$c_1 = \left[\frac{1}{\sqrt{2\pi\sigma}} \right]^n, c_2 = e^{-\frac{1}{N_0} \int_0^T [z(t) - u_x(t)]^2 dt}, \text{ а производная прибли-}$$

зительно равна значению этой производной при $x = x_H$.

$$\left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} \cong \left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H}.$$

Постоянная c определяется из условия нормировки:

$$\int_{-\infty}^{\infty} w(\bar{y}/x) dx = 1.$$

Дисперсия равна

$$\sigma^2 = \frac{N_0}{2T} \cdot \frac{1}{\frac{1}{T} \int_0^T \left[\left. \frac{\partial u_x(t)}{\partial x} \right|_{x_H} \right]^2 dt}. \quad (7.61)$$

Выражение (7.61) определяет среднеквадратичную ошибку при воспроизведении сообщения.

В случае амплитудной модуляции сигнал определяется формулой (3.31). Полагая, что $b(t) = Bb_0(t)$, где B — амплитуда

сообщения, $m = B/A$, а $b_0(t) = x$, выражение (3.31) перепишем в виде

$$u_{AM}(t) = u_x(t) = (1 + mx)A \cos \omega_0 t.$$

Производная

$$\frac{\partial u_x(t)}{\partial x} = mA \cos \omega_0 t.$$

С учетом последней формулы выражение для дисперсии:

$$\sigma^2 = \frac{N_0}{m^2 A^2 T} = \left(\frac{N_0}{E} \right) \frac{1}{2m^2}, \quad (7.62)$$

где $E = A^2 \int_0^T \cos^2 \omega_0 t = A^2 T / 2$ — энергия несущего колебания.

Аналогичные вычисления для ЧМ сигнала дают следующий результат:

$$\sigma^2 = \frac{6N_0}{E\Omega^2 T^2}, \quad (7.63)$$

где Ω — частота модуляции.

Сравнивая формулы (7.63) и (7.62), находим, что при большом отношении сигнал/шум и $\Omega T \gg 1$ применение ЧМ дает в

$$(\Omega T)^2 / 12 m^2$$

раз меньшую ошибку по сравнению с АМ сигналом. Но при $\Omega T \gg 1$ полоса частот, необходимая для передачи ЧМ сигнала, оказывается значительно большей, чем для передачи АМ сигнала.

7.9.3. Прием сигналов, модулированных сообщением

Рассмотрим более сложный случай приема, когда сообщение в интервале наблюдения $[0, T]$ не постоянно, а определяется реализацией некоторого случайного процесса.

Полагая, что первичный сигнал совпадает с передаваемым сообщением, представим его в виде ряда Фурье:

$$b(t) = \sum_{k=m_1}^{m_2} \left(\sqrt{2} \lambda_{2k-1} \sin \frac{2\pi k}{T} t + \sqrt{2} \lambda_{2k} \cos \frac{2\pi k}{T} t \right) = \sum_{k=1}^v \lambda_k \varphi_k(t), \quad (7.64)$$

где λ_k — коэффициенты разложения, являющиеся действующими случайными значениями спектральных составляющих первичного сигнала; $\varphi_k(t)$ — базисные функции типа $\sin x$, $\cos x$.

Число спектральных составляющих равно $v = 2(m_2 - m_1 + 1) = 2F_c T$, где F_c — верхняя граничная частота спектра первичного сигнала.

Сигнал на выходе передатчика зависит от времени и модулируемого параметра, т.е. является функцией $u[t, b(t)]$. С учетом (7.64) выражение для сигнала может быть записано в виде функции времени и значений λ_k , $k = 1, 2, \dots, \nu$ которые в интервале $[0, T]$ остаются постоянными и от интервала к интервалу меняются случайным образом.

$$u[t, \bar{\lambda}], \quad \bar{\lambda} = \{\lambda_1, \lambda_2, \dots, \lambda_\nu\}.$$

Задача оптимального приемника в этом случае сводится к задаче совместной оценки значений $\bar{\lambda}$ сигнала. При отсутствии шума $\bar{\lambda}^* = \bar{\lambda}$, т.е. оценки совпадают с коэффициентами разложения. При приеме сигнала на фоне шума $\bar{\lambda}^* \neq \bar{\lambda}$. В результате оценка сообщения

$$b^*(t) = \sum_{k=1}^{\nu} (\lambda_k + \Delta\lambda_k) \varphi_k(t) = b(t) + \varepsilon(t), \quad (7.65)$$

где $\varepsilon(t)$ — погрешность воспроизведения сообщения, $\Delta\lambda_k = \lambda_k - \lambda_k^*$.

Оптимальный приемник, как и в случае приема отдельного значения непрерывного сообщения, вычисляет апостериорную плотность распределения вероятности и на основе одного из критериев (например, критерия максимума) определяет наиболее вероятную реализацию. Апостериорная плотность с учетом предыдущих замечаний о принимаемом сигнале:

$$w_{\bar{y}}(\bar{x}) = cw(\bar{y} / \bar{x}) = c_1 e^{-\frac{1}{N_0} \int_0^T [z(t) - u(t, \bar{\lambda})]^2 dt} = c_1 e^{-TR/N_0},$$

где c, c_1 — постоянные числа;

$$R = \frac{1}{T} \int_0^T [z(t) - u(t, \bar{\lambda})]^2 dt.$$

Максимум функции $w_{\bar{y}}(\bar{x})$ обеспечивается при выполнении условий:

$$\frac{\partial R}{\partial \lambda_k} = \frac{2}{T} \int_0^T [z(t) - u(t, \bar{\lambda})] \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} dt = 0; \quad k = 1, 2, \dots, \nu. \quad (7.66)$$

При λ_k , близком к наиболее вероятному значению λ_{k_H} , выражение для сигнала имеет вид:

$$\begin{aligned} u(t, \bar{\lambda}) &= u(t, \lambda_1, \lambda_2, \dots, \lambda_{k_H}, \dots) + (\lambda_k - \lambda_{k_H}) \left. \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \right|_{\lambda_k = \lambda_{k_H}} \\ &= u(t, \lambda_1, \lambda_2, \dots, \lambda_{k_H}) + \Delta\lambda_k \left. \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \right|_{\lambda_k = \lambda_{k_H}} \end{aligned} \quad (7.67)$$

Подставляя (7.67) в (7.66) и, полагая, что $\frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \cong \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \Big|_{\lambda_k = \lambda_{kH}}$,

получим:

$$\frac{1}{T} \int_0^T \left[\varepsilon(t) - \Delta \lambda_k \left(\frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \right)_{\lambda_k = \lambda_{kH}} \right] \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} dt = 0,$$

$$\Delta \lambda_k = \frac{\frac{1}{T} \int_0^T \varepsilon(t) \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} dt}{\frac{1}{T} \int_0^T \left[\frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \right]^2 dt} = \frac{\frac{1}{T} \int_0^T \varepsilon(t) D_k(t) dt}{\tilde{D}_k^2},$$

где

$$D_k(t) = \frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k}, \quad \tilde{D}_k^2 = \frac{1}{T} \int_0^T \left[\frac{\partial u(t, \bar{\lambda})}{\partial \lambda_k} \right]^2 dt.$$

Среднеквадратичное значение равно

$$M(\Delta \lambda_k)^2 = \frac{\frac{1}{T^2} \int_0^T \int_0^T M\{\xi(t)\xi(t')\} D_k(t) D_k(t') dt dt'}{\tilde{D}_k^4} = \frac{N_0}{2T \tilde{D}_k^2},$$

где $M\{\xi(t)\xi(t')\} = \frac{N_0}{2} \delta(t - t')$ — корреляционная функция шума на входе приемника.

Из последнего выражения определяется спектральная плотность мощности помехи на выходе оптимального приемника, которая характеризует *потенциальную верность передачи*:

$$G(f) = N_0 / \tilde{D}_k^2. \quad (7.68)$$

Эта формула найдена с учетом следующего. Второе слагаемое в формуле (7.65) на частоте $f_k = k/T$ равно сумме случайных квадратурных составляющих (см. 7.64). Дисперсии этих составляющих одинаковы (см. разд. 4.5) и равны $N_0 / 2T \tilde{D}_k^2$. Следовательно, дисперсия помехи на частоте f_k равна $N_0 / T \tilde{D}_k^2$. Поскольку соседние спектральные компоненты первичного сигнала отстоят друг от друга на величину, равную $1/T$ (см. 7.64), то спектральная плотность мощности помехи на выходе приемника определяется выражением (7.68). Оно находится делением дисперсии помехи на частотный интервал $1/T$:

$$G(f) = \sigma^2 T = N_0 T / T \tilde{D}_k^2.$$

7.9.4. Сравнительная оценка помехоустойчивости оптимальных систем с различными видами модуляции

Выражение (7.68) используется при сравнении различных видов модуляции.

Ненормированное сообщение $b(t)$ имеет среднюю мощность

$$P_b = M[b^2(t)] = b_{\text{макс}}^2 M[b_0^2(t)] = b_{\text{макс}}^2 / \Pi^2,$$

где $b_0(t) = \frac{b(t)}{b_{\text{макс}}}$, $\Pi = b_{\text{макс}} / \sqrt{P_b}$ — пикфактор сообщения,

$$M[b_0^2(t)] = P_b / b_{\text{макс}}^2 = 1 / \Pi^2.$$

Если сообщение нормировано, то $b_{\text{макс}} = 1$ и $P_b = 1 / \Pi^2$.

Отношение мощностей сигнала и шума на выходе приемника

$$\rho_{\text{вых}} = P_b / P_{\varepsilon} = \frac{1}{\Pi^2 \int_0^{F_c} G(f) df}. \quad (7.69)$$

Сравним системы связи с различными видами модуляции. При амплитудной модуляции сигнал

$$u[t, b_0(t)] = A[1 + mb_0(t)] \cos \omega_0 t, \quad m \leq 1.$$

Найдем производную

$$D(t) = \frac{\partial u[t, b_0]}{\partial b_0} = mA \cos \omega_0 t,$$

а также среднее значение от квадрата этой производной

$$\bar{D}^2(t) = \frac{1}{T} \int_0^T \left(\frac{\partial u(t, b_0)}{\partial b_0} \right)^2 dt = \frac{1}{T} \int_0^T m^2 A^2 \cos^2 \omega_c t dt = \frac{1}{2} m^2 A^2.$$

Спектральная плотность помехи на выходе приемника определяется с помощью формулы (7.68):

$$G(f) = 2N_0 / m^2 A^2. \quad (7.70)$$

Подставляя (7.70) в (7.69), получим

$$\rho_{\text{вых}} = \frac{m^2 A^2}{2 \Pi^2 N_0 F_c}. \quad (7.71)$$

Из формулы видно, что, помимо очевидного способа повышения верности передаваемого сообщения, при АМ имеется и другой способ, связанный с уменьшением пикфактора передаваемого сообщения. Обычно он используется в форме сжатия динамического диапазона при передаче сообщения и обратного преобразования при приеме этого сообщения.

Выигрыш в помехоустойчивости определяется отношением:

$$g = \rho_{\text{вых}} / \rho_{\text{вх}} = \frac{2m^2}{\Pi^2 + m^2}, \quad (7.72)$$

так как мощность входного АМ сигнала равна $P_c = \frac{A^2}{2} [1 + m^2 / \Pi^2]$,

а, мощность шума в полосе, равной $2F_c$, равна $2N_0F_c$.

Обобщенный выигрыш

$$g' = g / \alpha = \frac{m^2}{m^2 + \Pi^2}, \quad (7.73)$$

так как $\alpha = 2$.

При балансной модуляции (АМ-ПН)

$$u[t, b(t)] = Ab(t) \cos \omega_0 t.$$

Средняя мощность сигнала равна $P_c = A^2 / 2 \Pi^2$, $\bar{D}^2(t) = \frac{1}{2} A^2$,

$$G(f) = 2N_0 / A^2.$$

Отношение сигнал/шум на выходе приемника:

$$\rho_{\text{вых}} = \frac{A^2}{\Pi^2 2N_0 F_c}. \quad (7.74)$$

Отношение сигнал/шум на входе приемника:

$$\rho_{\text{вх}} = \frac{A^2}{2\Pi^2 2N_0 F_c}. \quad (7.75)$$

Выигрыш:

$$g = \rho_{\text{вых}} / \rho_{\text{вх}} = 2, \quad (7.76)$$

а обобщенный выигрыш $g' = 1$.

При однополосной модуляции производится перенос спектра из области низких частот в область высоких частот без его изменения. Поэтому для данного вида модуляции

$$\left. \begin{aligned} \rho_{\text{вых}} &= \rho_{\text{вх}} \\ g &= g' = 1 \end{aligned} \right\} \quad (7.77)$$

При этом полоса частот в 2 раза меньше, чем при АМ-ПН.

При фазовой модуляции сигнал

$$u[t, b(t)] = A \cos[\omega_0 t + m_\phi b_0(t)],$$

где m_ϕ — индекс фазовой модуляции.

Производная

$$\frac{\partial u[t, b_0(t)]}{\partial b_0} = -Am_\phi \sin[\omega_0 t + m_\phi b_0(t)];$$

$$\tilde{D}^2(t) = \frac{1}{T} \int_0^T \left\{ \frac{\partial u[t, b_0(t)]}{\partial b_0} \right\}^2 dt = \frac{A^2}{2} m_{\Phi}^2;$$

спектральная плотность шума на выходе

$$G(f) = \frac{2N_0}{m_{\Phi}^2 A^2}.$$

Мощность шума на выходе приемника равна

$$G(f)F_c = \frac{2N_0 F_c}{m_{\Phi}^2 A^2}.$$

Мощность сигнала на выходе приемника равна $1/\Pi^2$.

Отношение сигнал/шум

$$\rho_{\text{вых}} = \frac{A^2 m_{\Phi}^2}{2\Pi^2 N_0 F_c}; \quad (7.78)$$

мощность сигнала на входе приемника

$$P_c = -\frac{1}{T} \int_0^T u^2[t, b(t)] dt = \frac{A^2}{2};$$

мощность шума на входе приемника

$$P_{\text{ш}} = N_0 F,$$

где F — ширина спектра сигнала;

отношение сигнал/шум на входе приемника

$$\rho_{\text{вх}} = \frac{A^2}{2N_0 F}; \quad (7.79)$$

выигрыш

$$g_{\text{фм}} = \rho_{\text{вых}} / \rho_{\text{вх}} = m_{\Phi}^2 \alpha / \Pi^2, \quad (7.80)$$

где $\alpha = F/F_c$;

обобщенный выигрыш

$$g'_{\text{фм}} = m_{\Phi}^2 / \Pi^2. \quad (7.81)$$

Из последних формул следует, что выигрыш увеличивается с увеличением индекса модуляции и уменьшением пикфактора сообщения.

При частотной модуляции

$$u[t, b(t)] = A \cos[\omega_0 t + \Delta\omega a(t)], \quad (7.82)$$

где $\Delta\omega$ — девиация частоты;

$$a(t) = \int_0^t b_0(t) dt;$$

Производная

$$\frac{\partial u}{\partial a} = -A\Delta\omega \sin[\omega_0 t + \Delta\omega a(t)]; \quad \tilde{D}^2(t) = \frac{1}{T} \int_0^T \left\{ \frac{\partial u}{\partial a} \right\}^2 dt = \frac{1}{2} \Delta\omega^2 A^2.$$

Спектральная плотность шума на выходе приемника

$$G(f) = 2 \left(\frac{2\Pi f}{\Delta\omega} \right)^2 \frac{N_0}{A^2}. \quad (7.83)$$

Последняя формула получена с учетом того, что сообщение $b_0(t)$ входит в выражение сигнала (7.82) под знаком интеграла. Поэтому спектральная плотность шума на выходе приемника определяется как спектральная плотность производной $\xi'(t)$ [27].

Мощность шума на выходе приемника:

$$\int_0^{F_c} G(f) df = \frac{2}{3} \frac{N_0}{A^2} \frac{F_c^3}{\Delta f^2}.$$

Выигрыш в помехоустойчивости:

$$g_{\text{ЧМ}} = \frac{3}{\Pi^2} \left(\frac{\Delta f}{F_c} \right)^2 \cdot \frac{F}{F_c} = \frac{3}{\Pi^2} m_{\text{ЧМ}}^2 \cdot \alpha, \quad (7.84)$$

где $m_{\text{ЧМ}} = \Delta f / F_c$ — индекс частотной модуляции.

Обобщенный выигрыш:

$$g'_{\text{ЧМ}} = \frac{3F^2}{4\Pi^2 F_c} = \frac{2m_{\text{ЧМ}}^2}{\Pi^2}. \quad (7.85)$$

При ЧМ, так же, как и при ФМ, выигрыш в помехоустойчивости увеличивается с увеличением индекса модуляции и уменьшением пикфактора сообщения. При ФМ и ЧМ индекс модуляции может быть больше 1, тогда как при различных видах АМ индекс модуляции меньше 1. Это свидетельствует о том, что при малом уровне помех ЧМ и ФМ сигналы имеют преимущество в помехоустойчивости по сравнению с АМ сигналами.

7.9.5. Порог помехоустойчивости в системах с широкополосными видами модуляции

Рассмотрим передачу сообщения по непрерывному каналу связи. Сообщение является непрерывной реализацией гауссовского случайного процесса с равномерным спектром в полосе частот $0-F_c$, а в канале действует помеха в виде белого нормального шума с заданной спектральной плотностью. Мощность шума и мощность сигнала в полосе $0-F$ ограничены.

Передача сообщения по каналу связи с заданным отношением сигнал/шум P_b/P_c возможна согласно теореме Шеннона в том

случае, если ε — производительность источника меньше пропускной способности канала, т.е. при

$$H'_\varepsilon(b) < C,$$

где $H'_\varepsilon(b)$ — производительность источника; C — пропускная способность канала связи.

Эти величины определяются формулами (6.34) и (6.46) соответственно.

В идеальной системе связи $H'_\varepsilon(b) = C$, поэтому

$$F_c \log \frac{P_b}{P_\varepsilon} = F \log(1 + P_c / P_{\text{ш}}). \quad (7.86)$$

Принимая во внимание введенные ранее обозначения:

$$F/F_c = \alpha; \quad P_b/P_\varepsilon = \rho_{\text{вых}}; \quad P_c/P_{\text{ш}} = \rho_{\text{вх}},$$

запишем формулу (7.86) в виде:

$$\log \rho_{\text{вых}} = \alpha \log(1 + \rho_{\text{вх}}).$$

Из полученного равенства найдем

$$\rho_{\text{вых}} = (1 + \rho_{\text{вх}})^\alpha. \quad (7.87)$$

Для больших отношений сигнал/шум на входе $\rho_{\text{вх}} \gg 1$, следовательно

$$\rho_{\text{вых}} \approx \rho_{\text{вх}}^\alpha.$$

Выигрыш в идеальной системе связи (при условии $(\rho_{\text{вх}} \gg 1)$ равен

$$g = \frac{\rho_{\text{вых}}}{\rho_{\text{вх}}} = \rho_{\text{вх}}^{\alpha-1}, \quad (7.88)$$

а обобщенный выигрыш

$$g' = \frac{\rho_{\text{вых}}}{\rho_{\text{вх}}^\alpha} = \frac{1}{\alpha} \rho_{\text{вх}}^{\alpha-1}. \quad (7.89)$$

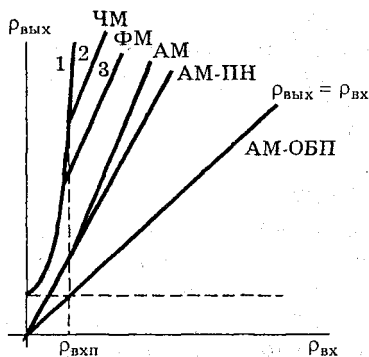


Рис. 7.31. Графики зависимостей отношения сигнал/шум на выходе от отношения сигнал/шум на входе для различных систем связи

График зависимости $\rho_{\text{вых}} = f(\rho_{\text{вх}})$, построенный по формуле (7.87) для заданного значения α , приведен на рис. 7.31 (кривая 1). Здесь же приведены аналогичные зависимости для оптимальных непрерывных систем связи с ФМ и ЧМ, рассчитанные по формулам (7.80) и (7.84) при том же значении α . Это (прямые 3 и 2). Для ФМ сигнала:

$$\rho_{\text{вых}} = \rho_{\text{вх}} \frac{m_\Phi^2 \alpha}{\Pi^2},$$

а для ЧМ сигнала

$$\rho_{\text{вых}} = \rho_{\text{вх}} \frac{3}{\Pi^2} m_{\text{ЧМ}}^2 \alpha.$$

Линии 2, 3 в некоторой области значений входного отношения сигнал/шум $[\rho_{\text{вхл}}, \infty]$ лежат ниже кривой 1. Это означает, что оптимальные непрерывные системы связи с ФМ и ЧМ сигналами проигрывают по помехоустойчивости идеальной по Шеннону системе связи. В области $[0, \rho_{\text{вхл}}]$ линии 2,3 не могут располагаться выше кривой 1, так как это означало бы, что непрерывные системы с ФМ и ЧМ модуляцией при малых входных отношениях сигнал/шум $\rho_{\text{вх}}$ лучше по помехоустойчивости идеальной по Шеннону системы связи, что противоречит и теории и экспериментальным данным. Поэтому, начиная с некоторого значения отношения $\rho_{\text{вхл}}$, которое называется *пороговым*, выходное отношение сигнал/шум резко уменьшается.

С уменьшением $\rho_{\text{вых}}$ также резко падает и выигрыш систем, определяемый формулами (7.80), (7.84).

На рис.7.31 приведены также аналогичные зависимости, построенные по формулам (7.72), (7.76), (7.77) для АМ, АМ-ПН и АМ-ОБП сигналов. Эти прямые расположены ниже кривой для идеальной по Шеннону системы связи, следовательно помехоустойчивость оптимальных систем с этими видами модуляций хуже помехоустойчивости идеальной системы. В области больших значений $\rho_{\text{вх}}$ системы связи с ФМ и ЧМ выигрывают по помехоустойчивости по сравнению с системами АМ, но начиная с некоторого порогового уровня этот выигрыш резко уменьшается. При малом отношении $\rho_{\text{вх}}$ оптимальные системы связи с АМ имеют даже выигрыш по сравнению с системами с ФМ и ЧМ, поскольку, как видно из рис.7.31, системы с АМ не имеют пороговых явлений.

Пороговые значения $\rho_{\text{вхл}}$, полученные для оптимальных методов приема непрерывных сигналов, являются предельными для реальных систем, т.е. помехоустойчивость реальных систем связи хуже оптимальных. Например, порог в реальной системе с ЧМ наступает при равенстве пиковых значений сигнала и помехи, что соответствует, такому входному отношению $\rho_{\text{вх}}$, которое существенно выше теоретического. Это означает, что имеется возможность совершенствования приемников с целью повышения их помехоустойчивости. Например, для снижения порога при ЧМ применяют различные схемы демодуляторов с обратной связью по частоте.

7.10. Оптимальная линейная фильтрация.

Понятие о нелинейной фильтрации

7.10.1. Линейная фильтрация сигнала

Задача фильтрации сигнала из шумов в достаточно общем виде ставится следующим образом. В интервале времени $[0, T]$ принимается смесь сигнала $u(t)$ и шума $N(t)$.

$$z(t) = f\{u[t, p_i(t)], N(t)\}, i = 1, 2, \dots, k.$$

Сигнал является функцией времени и зависит от параметров $p_i(t)$, которые изменяются случайным образом. Один из них, например j -й, изменяется в соответствии с передаваемым сообщением $b(t)$. Шум может быть произвольным, сигнал и шум не обязательно образуют аддитивную смесь. Функция известна, также как и некоторые статистические характеристики смеси $z(t)$.

На основании приема этой смеси, нужно выделить реализацию сообщения $b(t)$ с наименьшей ошибкой. Этот процесс выделения называется *фильтрацией*. Для количественной оценки качества фильтрации используют заранее сформулированные критерии. Такими критериями могут быть: *критерий минимума среднеквадратичной ошибки*; *критерий максимума апостериорной вероятности*.

В результате решения задачи получают уравнения, определяющие структуру *оптимального фильтра*.

Ряд практически интересных задач решены для следующих условий:

1) сигнал и шум являются стационарными случайными процессами;

2) сигнал и шум взаимодействуют аддитивно;

3) сигнал, входящий в смесь, является линейной функцией сообщения, т.е. $u(t) = kb(t)$, где k — постоянный коэффициент. Фильтрация такого сигнала называется *линейной*;

4) сигнал, входящий в смесь, является нелинейной функцией сообщения, т.е. $u(t) = [f, b(t)]$. Соответствующая этому сигналу фильтрация называется *нелинейной*.

Рассмотрим задачу оптимальной линейной фильтрации.

Пусть на вход линейного четырехполосника с импульсной характеристикой $h(t)$ подается аддитивная смесь сигнала и помехи

$$z(t) = u(t) + N(t).$$

Сигнал $u(t)$ и помеха $N(t)$ являются стационарными взаимно-некоррелированными процессами с известными спектральными плотностями.

На выходе воспроизводится отклик $u^*(t)$ линейного четырехполюсника на входное колебание $z(t)$. Этот отклик является оценкой входного сигнала $u(t)$, следовательно и сообщения. Требуется найти линейный четырехполюсник с такой импульсной характеристикой $h(t)$, которая минимизирует средний квадрат ошибки

$$M\{\varepsilon^2(t)\} = M\{[u^*(t) - u(t)]^2\},$$

где среднее значение берется по ансамблю сигналов и помех.

Задача линейной фильтрации была независимо решена А.Н.Колмогоровым (1939 г.) и Н.Винером (1942 г.) и поэтому линейный фильтр, удовлетворяющий сформулированным выше условиям, называется *фильтром Колмогорова-Винера*.

Необходимым и достаточным условием оптимальной линейной фильтрации является условие

$$M\{\varepsilon(t)z(t - \tau)\} = 0 \quad (7.90)$$

для всех τ из заданной области интегрирования γ входной смеси сигнала и помехи $z(t)$.

Это условие означает, что ошибка в оптимальном линейном фильтре не коррелирована со входным сигналом.

Докажем справедливость условия (7.90).

Пусть $h_1(t)$ — импульсная характеристика оптимального линейного фильтра, а $h_2(t)$ — импульсная характеристика другого (неоптимального) фильтра. Пусть отклики фильтров равны $u_1^*(t)$ и $u_2^*(t)$ и каждый из этих откликов определяется интегралом Дюамеля:

$$u^*(t) = \int_{\gamma} z(t - \tau)h(\tau)d\tau, \quad (7.91)$$

где область интегрирования γ для физически реализуемого фильтра равна интервалу $[0, \infty]$.

Для физически нереализуемого фильтра интервал γ равен $[-\infty, \infty]$.

Найдем среднеквадратичное значение разности

$$M\{[u(t) - u_2^*(t)]^2\} = M\{u(t) - u_1^*(t) + u_1^*(t) - u_2^*(t)\}^2 = M\{[u(t) - u_1^*(t)]^2\} + M\{[u_1^*(t) - u_2^*(t)]^2\} + 2[u(t) - u_1^*(t)][u_1^*(t) - u_2^*(t)].$$

Для оптимального линейного фильтра функция $u(t) - u_1^*(t) = \varepsilon(t)$ удовлетворяет условию (7.90).

Поэтому

$$M\{[u(t) - u_1^*(t)][u_1^*(t) - u_2^*(t)]\} = M\{\varepsilon(t) \int_{\gamma} z(t - \tau)[h_1(\tau) - h_2(\tau)d\tau] = \int_{\gamma} M\{\varepsilon(t)z(t - \tau)[h_1(\tau) - h_2(\tau)]d\tau = 0.$$

Следовательно,

$$M\{u(t) - u_2^*(t)\}^2 = M\{\varepsilon^2(t)\} + M\{[u_1^*(t) - u_2^*(t)]^2\}.$$

Последнее выражение будет минимальным в том случае, если

$$u_2^*(t) = u_1^*(t).$$

Условие (7.90) запишем в виде

$$M\{[u(t) - u^*(t)]z(t - \tau)\} = 0,$$

или с учетом (7.91) в виде:

$$\begin{aligned} M\{[u(t)z(t - \tau)]\} &= M\{u^*(t)z(t - \tau)\} = \int_{\gamma} M\{z(t - x)z(t - \tau)\}h(x)dx = \\ &= \int_{\gamma} B_z(\tau - x)h(x)dx = B_{uz}(\tau), \end{aligned}$$

где $B_z(\tau) = M\{z(t - x)z(t - \tau)\}$ — корреляционная функция смеси сигнала и помехи на входе фильтра; $B_{uz}(\tau)$ — взаимнокорреляционная функция между сигналом $u(t)$ и смесью сигнала и помехи $z(t)$.

Если сигнал $u(t)$ и помеха $N(t)$ некоррелированы, то последнее выражение можно записать в виде:

$$B_u(\tau) = \int_{\gamma} [B_u(\tau - x) + B_N(\tau - x)]h(x)dx, \quad (7.92)$$

где $B_u(\tau)$, $B_N(\tau)$ — корреляционные функции сигнала и шума.

Полученное уравнение называется *уравнением Винера-Хопфа*. Решив это уравнение, можно найти импульсную характеристику $h(t)$, которая минимизирует среднеквадратичную ошибку.

Уравнение Винера-Хопфа легко решается для нереализуемых фильтров, т.е. когда область интегрирования γ равна $[-\infty, \infty]$.

В этом случае применяют преобразование Фурье к обеим частям уравнения (7.92) и получают результат в частотной области:

$$G_u(f) = [G_u(f) + G_N(f)]K(f).$$

Из полученного уравнения находят выражение для коэффициента передачи оптимального линейного фильтра:

$$K(f) = \frac{G_u(f)}{G_u(f) + G_N(f)}. \quad (7.93)$$

Рассмотрим частные случаи.

1. Спектральная плотность помехи $G_N(f) = 0$.

По формуле (7.93) определяем коэффициент передачи оптимального линейного фильтра

$$K(f) = \frac{G_u(f)}{G_u(f)} = 1.$$

Если помеха отсутствует, то можно использовать фильтр, коэффициент передачи которого равен 1 на любой частоте.

2. Интенсивность помехи превышает сигнал так, что $G_N(f) \gg G_u(f)$.

Для этого случая

$$K(f) = G_u(f) / G_N(f), \quad (7.94)$$

т.е. коэффициент передачи оптимального линейного фильтра определяется отношением спектральной плотности сигнала к спектральной плотности помехи. На тех частотах, где интенсивность спектральной плотности помехи имеет наибольшее значение, коэффициент передачи должен иметь минимальное значение.

Линейные фильтры, характеристики которых приближаются к оптимальным, находят достаточно широкое применение в системах электросвязи. Такими являются: фильтры нижних частот, полосовые фильтры и др., имеющие равномерный коэффициент передачи в заданной полосе частот. При малом уровне помехи эти фильтры передают сигнал без существенных искажений.

Результаты теории линейной фильтрации используются и при построении широкополосного канала связи для защиты от сосредоточенных помех.

Коэффициенты передачи этих каналов выбираются так, чтобы удовлетворялось условие (7.94). Например, при частотно-разнесенном приеме канал связи состоит из нескольких узкополосных каналов, по которым передается одна и та же информация. В случае поражения узкополосных каналов сосредоточенными помехами они отключаются от схемы выделения полезной информации. В результате коэффициент передачи канала на пораженных помехами участках становится равным нулю.

Методы борьбы с импульсными и сосредоточенными помехами изложены выше, в разд. 7.7.

Ошибка в воспроизведении сообщения определяется видом коэффициента передачи линейного фильтра. Доказывается [29], что для оптимального фильтра с коэффициентом передачи (7.93)

$$M\{\varepsilon^2(t)\} = \int_{-\infty}^{\infty} \frac{G_u(f)G_N(f)}{G_u(f) + G_N(f)} df.$$

Среднеквадратичная ошибка равна нулю, если $G_N(f) = 0$. Если $G_N(f) \gg G_u(f)$, то

$$M\{\varepsilon^2(t)\} = \int_{-\infty}^{\infty} G_u(f) df.$$

т.е. определяется спектральной плотностью сигнала.

7.10.2. Понятие о нелинейной фильтрации

Если сигнал является нелинейной функцией сообщения (например при фазовой или частотной модуляции), то фильтрация является *нелинейной*.

Рассмотрим наиболее простой случай, когда сигнал зависит от времени и от единственного параметра, изменяющегося в соответствии с сообщением.

Принимаемое колебание является аддитивной суммой сигнала, зависящего от одного параметра, и стационарного белого шума $N(t)$:

$$z(t) = u[t, b(t)] + N(t).$$

Среднее значение шума $M\{N(t)\} = 0$, а корреляционная функция

$$M\{N(t_1)N(t_2)\} = \frac{N_0}{2} \delta(t_2 - t_1),$$

где N_0 — односторонняя спектральная плотность мощности этого шума.

Задачей оптимальной нелинейной фильтрации является нахождение устройства, которое бы с наибольшей точностью воспроизводило сообщение $b(t)$ при заданных выше условиях.

Устройство, вычисляющее апостериорную плотность вероятности и принимающее решение на основании критерия максимума этой плотности, называется *оптимальным*.

Математическое выражение для апостериорной плотности можно найти для двух случаев:

- 1) когда в качестве математической модели сообщения $b(t)$ используется *марковский процесс*;
- 2) когда математической моделью сообщения является *нормальный случайный процесс*.

Задача нелинейной фильтрации с использованием математической модели сообщения в виде нормального случайного процесса решена Большаковым И.А. и Репиным В.Г. [31]. Структура оптимального следящего устройства определяется интегральным уравнением, которое моделируется *фильтрами с переменными параметрами*.

Задаче нелинейной фильтрации с использованием математической модели сообщения в виде марковского процесса решена Р.Л.Стратановичем [32]. В этой модели используется случайный

процесс на выходе линейной интегрирующей RC цепи, являющийся откликом на входной стационарный белый шум $n_1(t)$.

Полагают, что сообщение $b(t)$ является одной из реализаций этого процесса и описывается следующим дифференциальным уравнением 1-го порядка:

$$\frac{db(t)}{dt} = -\alpha b(t) + n_1(t), \quad (7.95)$$

где $\alpha = 1/RC$ — постоянная времени интегрирующей цепи.

Дифференциальное уравнение определяет функцию корреляции

$$B_b(\tau) = \frac{N_1}{4\alpha} e^{-\alpha|\tau|}$$

и спектральную плотность мощности

$$G_b(f) = \frac{2\alpha P_b}{\alpha^2 + 4\pi^2 f^2},$$

где N_1 — односторонняя спектральная плотность белого шума на выходе линейной цепи; $P_b = N_1/(4\alpha)$ — мощность сообщения.

Случайный процесс на выходе интегрирующей цепи является гауссовским и в то же время марковским процессом.

Аппроксимация реального сообщения марковским процессом позволяет достаточно точно описать его, сохраняя при этом возможность упростить математический анализ.

Решение задачи оптимальной нелинейной фильтрации является сложным. На первом этапе находят дифференциальное уравнение в частных производных, которое определяет финальную апостериорную плотность вероятности:

$$\frac{\partial \tilde{W}(t, b)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial b^2} [K_2(b) \tilde{W}(t, b)] - \frac{\partial}{\partial b} [K_1(b) \tilde{W}(t, b)] + [F(t, b) - \langle F(t, b) \rangle] \tilde{W}(t, b), \quad (7.96)$$

где $\tilde{W}(t, b)$ — апостериорная плотность вероятности; $K_1(b)$, $K_2(b)$ — коэффициенты, которые находятся из дифференциального уравнения (7.95) по формулам:

$$K_1(b_\tau) = \lim_{\tau \rightarrow 0} \frac{\langle b_\tau - b \rangle}{\tau}, \quad K_2(b_\tau) = \lim_{\tau \rightarrow 0} \frac{\langle (b_\tau - b)^2 \rangle}{\tau}.$$

Знак $\langle \rangle$ определяет статистическое усреднение. Значения сообщения в моменты времени t и $t + \tau$ обозначены соответственно b и b_τ .

Функция

$$F(t, b) = -\frac{1}{N_0} [z(t) - u(t, b)]^2,$$

$$\langle F(t, b) \rangle = \int_{-\infty}^{\infty} F(t, b) \tilde{W}(t, b) db.$$

Применительно к незнергетическим параметрам сигнала (такими параметрами являются фаза и частота):

$$F(t, b) = \frac{2}{N_0} z(t) u(t, b).$$

Апостериорная плотность вероятности $\tilde{W}(t, b)$ изменяется во времени. Если интервал наблюдения $[0, T]$ смеси сигнала и шума $z(t)$ разбить на m сравнительно малых интервалов длиной $\Delta = T/m$ и мысленно наблюдать за изменением $\tilde{W}(t, b)$, то в конце каждого интервала апостериорная плотность будет принимать новую форму. Значение функции $\tilde{W}(t, b)$ в конце последнего интервала, полученное в результате обработки реализации $z(t)$ на всех первых интервалах, называется *финальной апостериорной плотностью вероятности*.

Оптимальное устройство фильтрации сообщения $b(t)$ должно сформировать $\tilde{W}(t, b)$ в соответствии с уравнением (7.96) и определить значение $b^*(t)$, соответствующее максимуму апостериорной плотности вероятности.

Решить уравнение (7.96) удается при допущении, что апостериорная плотность распределения вероятности определяется нормальным законом:

$$\tilde{W}(t, b) = \frac{1}{\sqrt{2\pi\sigma_B^2}} \exp \left\{ -\frac{[b(t) - b^*(t)]^2}{2\sigma_B^2} \right\},$$

где $b^*(t)$ — среднее значение, определяющее оптимальную оценку передаваемого сообщения $b(t)$; σ_B^2 — дисперсия нормального распределения.

Это позволяет свести решение задачи оптимальной нелинейной фильтрации к решению следующей системы уравнений:

$$\begin{aligned} \frac{d}{dt} b^*(t) &= K_1(t, b^*) + \sigma_B^2(t) \frac{\partial F(t, b^*)}{\partial b^*}; \\ \frac{d}{dt} \sigma_B^2(t) &= 2\sigma_B^2(t) \frac{\partial K_1(t, b^*)}{\partial b^*} + \sigma_B^4(t) \frac{\partial^2 F(t, b^*)}{\partial b^{*2}} + K_2(t, b^*). \end{aligned} \quad (7.97)$$

Для марковского гауссовского процесса, описываемого стохастическим дифференциальным уравнением (7.95):

$$\left. \begin{aligned} K_1(t, b^*) &= -\alpha b(t), \\ K_2(t, b^*) &= \frac{1}{2} N_1. \end{aligned} \right\} \quad (7.98)$$

С учетом уравнений (7.98) уравнения (7.97) запишутся в виде:

$$\begin{aligned} \frac{db^*(t)}{dt} &= -\alpha b^*(t) + \sigma_B^2(t) \frac{\partial F(t, b^*)}{\partial b^*}, \\ \frac{d}{dt} \sigma_B^2(t) &= -2\alpha \sigma_B^2(t) + \frac{1}{2} N_1 + \sigma_B^4(t) \frac{\partial^2 F(t, b^*)}{\partial b^{*2}}. \end{aligned}$$

Первое уравнение определяет структуру оптимального устройства, а второе — дисперсию ошибки, которая будет иметь место при оптимальной нелинейной фильтрации. Рассмотрим пример.

Пример 7.1. Требуется синтезировать оптимальный демодулятор для приема ФМ сигнала

$$u(t, b) = U_0 \cos[\omega_0 t + \theta(t)], \quad \theta(t) = \beta b(t) + \varphi(t),$$

где U_0 , ω_0 — известные значения амплитуды и несущей частоты сигнала; β — индекс фазовой модуляции; $\varphi(t)$ — процесс, описывающий флуктуации начальной фазы.

Решение. Полагаем, что сообщение является марковским гауссовским процессом.

Определим функцию

$$F(t, b^*) = \frac{2}{N_0} z(t) u(t, b^*) = -\frac{2U_0}{N_0} z(t) \cos[\omega_0 t + \theta^*(t)].$$

Полагая, что флуктуации начальной фазы отсутствуют, т.е. $\varphi(t) = \varphi_0$, запишем уравнение

$$\frac{db^*}{dt} = -\alpha b^*(t) + \frac{2\sigma_B^2(t)z(t)\beta U_0}{N_0} \sin[\omega_0 t + \theta^*(t)],$$

где

$$\theta^*(t) = \beta b^*(t) + \varphi_0.$$

Обозначим

$$\frac{2\sigma_B^2(t)\beta U_0 z(t)}{N_0} \sin[\omega_0 t + \theta^*(t)] = y(t).$$

Тогда

$$y(t) = b^*(t) + \frac{1}{\alpha} \frac{db^*}{dt}. \quad (7.99)$$

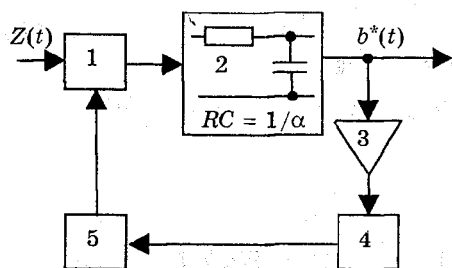


Рис.7.32. Структурная схема оптимального демодулятора ФМ сигнала: 1 — перемножитель; 2 — фильтр нижних частот; 3 — усилитель; 4 — управляемый элемент; 5 — перестраиваемый генератор

Структурная схема, моделирующая уравнение (7.99), приведена на рис.7.32.

Устройство оптимальной оценки $b(t)$ содержит: перестраиваемый генератор, фаза которого модулируется с помощью управляемого элемента; перемножитель и фильтр, выполненный на R, C элементах.

Схема является схемой автоподстройки фазы, в которой реализуется следящий прием сигнала. С изменением сообщения (фазы принимаемого

сигнала) изменяется и его оценка $b^*(t)$, а это в свою очередь приводит и к изменению фазы опорного сигнала. Оценка сообщения $b^*(t)$ формируется следующим образом. Входная смесь $z(t)$ перемножается в перемножителе с опорным подстраиваемым по фазе сигналом, а результат перемножения интегрируется RC цепью.

Операция вычисления взаимно-корреляционной функции с помощью перемножителя и интегратора при следящем приеме является основной операцией, совершаемой над смесью сигнала и шума.

Эта операция присутствует во всех оптимальных корреляционных приемниках, рассмотренных выше.

Принцип следящего приема сохраняется и при приеме ЧМ сигнала. В этом случае оптимальный приемник следит за изменением частоты принимаемого сигнала в соответствии с передаваемым случайным сообщением $b(t)$ с наименьшей ошибкой.

Структура приемника сохраняется такой же, как и при приеме ФМ сигнала, но фаза, определяющая сигнал,

$$\theta(t) = \beta \int_0^t b(t) dt.$$

Теория позволяет решать также и задачи совместной оптимальной фильтрации непрерывных и дискретных параметров сигнала.

7.11. Цифровые методы передачи непрерывных сообщений

7.11.1. Общие сведения о цифровых методах передачи.

Помехоустойчивость цифровых методов передачи непрерывных сообщений

Цифровые методы передачи непрерывных сообщений включают: методы формирования и обработки сигналов на основе их отображения совокупностью чисел; методы передачи этой совокупности по каналу связи.

Теоретические основы цифровых методов формирования сигналов рассмотрены в гл.3. Важное место в этой теории занимает метод представления сигналов обобщенным рядом Фурье, который устанавливает связь между сигналом, базисными функциями и спектральными коэффициентами. Один и тот же сигнал, представленный разными базисными функциями, может иметь разные спектральные коэффициенты. В частном случае, когда в качестве базисных функций используются функции типа $\sin x/x$, спектральными коэффициентами являются выборочные значения сигнала, а сигнал записывается в виде ряда Котельникова. Таким образом, первичный сигнал, непосредственно связанный с непрерывным сообщением, может быть отображен множеством чисел. Обычно первичный сигнал преобразуется в двоичные числа, над которыми производятся дальнейшие преобразования. Примерами таких преобразований являются: а) рассмотренные в гл.3 преобразования (для синтеза АМ, ФМ или ЧМ сигналов); б) преобразования с целью эффективного кодирования (гл.6); в) преобразования с целью помехоустойчивого кодирования (гл.6).

В приемнике производятся обратные преобразования.

Формирование и обработка сигналов на основе их отображения совокупностью чисел называется соответственно *цифровой генерацией и обработкой*.

Передача двоичных чисел по каналу связи может осуществляться или непосредственно без каких либо существенных преобразований или — посредством дискретной модуляции.

Цифровые методы передачи, кроме перечисленных в гл.3 преимуществ, обладают, по сравнению с аналоговыми (непрерывными) методами передачи, более высокой помехоустойчивостью. Это наиболее полно проявляется при *многократной ретрансляции сигналов*.

Принцип ретрансляции сигналов, нашедший применение в кабельных и радиорелейных линиях, поясняется рис.7.33 (на примере радиорелейной линии).

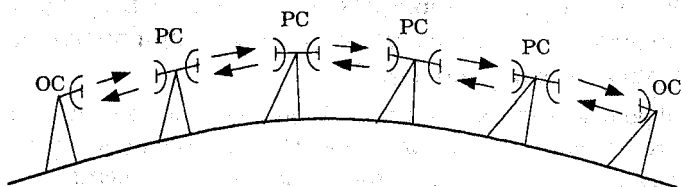


Рис.7.33. Ретрансляция сигналов в радиорелейной линии связи: РС — ретрансляционная станция; ОС — оконечная станция

В аналоговых линиях передачи сигналы передаются по цепи ретрансляторов, содержащих приемо-передающее оборудование. С помощью этого оборудования ослабленные по мощности сигналы принимаются (усиливаются), а затем передаются к следующему (соседнему) ретранслятору.

При передаче возникают искажения, обусловленные неидеальностью канала связи, а на сигнал накладываются флуктуационные помехи. Мощность этих помех на входе последнего звена линии равна сумме мощностей помех всех звеньев.

При передаче возникают искажения, обусловленные неидеальностью канала связи, а на сигнал накладываются флуктуационные помехи. Мощность этих помех на входе последнего звена линии равна сумме мощностей помех всех звеньев. Пусть число звеньев линии передачи, определяемое из условия обеспечения надежной связи, равно m . Тогда для обеспечения заданной верности передачи сообщения отношение сигнал-помеха на входе каждого ретранслятора должно быть в m раз больше, чем при передаче без ретрансляций. Число m в реальных системах связи достигает нескольких десятков, а иногда и сотен, что ведет к существенному ухудшению верности передачи.

В цифровых системах передачи (при том же числе ретрансляторов) верность передачи обеспечивается более высокой за счет регенерации импульсов. При этом форма импульсов, искаженных средой распространения и аппаратурой канала связи, восстанавливается, а помеха накапливается в гораздо меньшей степени. Полностью устранить влияние помех при регенерации сигналов в принципе невозможно. Это объясняется тем, что даже при весьма малом уровне помех существует не равная нулю вероятность ошибочного приема символов. При наличии ошибки вместо символа "0" на выходе регенератора будет формироваться "1", а вместо символа "1" — "0".

Для уменьшения этой вероятности необходимо увеличивать мощность передаваемого сигнала. С ростом числа ретрансляторов

требуемая мощность также увеличивается. Однако по сравнению с переприемом сигналов в аналоговой линии верность передачи в цифровой линии во много раз выше.

Проиллюстрируем это на примере.

Пример 7.2. Для канала с аддитивным гауссовским шумом и некогерентном приеме ортогональных сигналов требуется обеспечить вероятность ошибки $p < 10^{-5}$:

- 1) в цифровой линии без переприема;
- 2) в цифровой линии с переприемом сигналов;
- 3) в аналоговой линии с переприемом сигналов.

Все линии имеют одинаковую протяженность, а линии с переприемом содержат $m = 1000$ ретрансляторов.

Решение. В цифровой системе без переприема при заданной вероятности $p < 10^{-5}$ требуемое отношение h^2 определяется формулой (7.48):

$$h^2 = E / N_0 = -2 \ln(2p) = 21,64.$$

В цифровой линии передачи с переприемом и использовании m регенераторов на каждом из них нужно обеспечить вероятность ошибки в $m = 1000$ раз меньшую, равную 10^{-8} . Предполагается, что события, заключающиеся в появлении ошибок при первой, второй, и т.д. m -й регенерации сигнала, независимы. Поэтому общая вероятность ошибки равна сумме вероятностей ошибок, возникающих в первом, втором и т.д. регенераторах.

Вероятности ошибки $p = 10^{-8}$ соответствует отношение $h \cong 35,45$. Следовательно, мощность сигнала в цифровой линии с регенерацией сигнала в $35,45: 21,64 = 1,64$ раза больше той мощности, которая необходима в цифровой системе без регенерации.

В аналоговой линии передачи с переприемом сигнала на выходе первого ретранслятора сумма мощностей сигнала и помехи равна $P_c + P_{ш}$. Эта суммарная мощность при прохождении смеси сигнала и помехи по линии связи между первым и вторым ретранслятором уменьшается в " k " раз, где k — постоянное число.

Для компенсации потерь сигнала его необходимо усилить в это же число раз. При усилении к имеющейся помехе с мощностью $P_{ш}$ будет добавляться помеха с мощностью $P_{ш}$.

Поэтому сумма мощностей сигнала и помехи на выходе второго ретранслятора будет равна $P_c + 2P_{ш}$. Продолжив анализ, убедимся в том, что на выходе m -го ретранслятора сумма мощностей сигнала и помехи равна $P_c + mP_{ш}$. В аналоговой системе

с переприемом сигнала отношение сигнал-помеха ухудшается в $\rho_{\text{вых}}/\rho_{\text{вх}}$, где $\rho_{\text{вых}} = P_c/mP_{\text{ш}}$, $\rho_{\text{вх}} = P_c/P_{\text{ш}}$.

Для данного примера ухудшение отношения сигнал-помеха в $m = 1000$ раз.

7.11.2. Структурная схема цифровой системы передачи непрерывных сообщений

Структурная схема цифровой системы передачи непрерывных сообщений, наиболее часто используемой на практике, приведена на рис.7.34.

Для согласования непрерывного источника сообщения с дискретным каналом связи на передающей стороне введен аналого-цифровой преобразователь (АЦП) сигнала, а для согласования

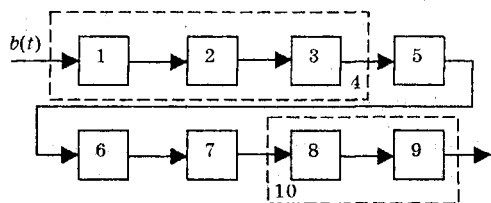


Рис.7.34. Структурная схема цифровой системы передачи непрерывных сообщений: 1 — дискретизатор; 2 — квантователь; 3 — кодер; 4 — АЦП; 5 — передатчик; 6 — линия связи; 7 — приемник; 8 — декодер; 9 — ФНЧ; 10 — ЦАП

дискретного канала связи с получателем непрерывного сообщения на приемной стороне введен цифро-аналоговый преобразователь (ЦАП).

Преобразование непрерывного сообщения с помощью АЦП в двоичный код называется *импульсно-кодовой модуляцией (ИКМ)*. Основы ИКМ рассмотрены в гл.3.

Сигнал ИКМ поступает на передатчик. В зависимости от вида линии связи сигнал ИКМ соответствующим образом преобразуется. Методы преобразований двоичных кодовых комбинаций в дискретные модулированные сигналы (АМ, ФМ, ОФМ, ЧМ) рассмотрены в гл.3.

В приемнике двоичные кодовые комбинации после демодуляции и регенерации поступают на ЦАП, где преобразуются вначале в квантованные отсчеты, которые затем пропускаются через фильтр нижних частот для получения непрерывного сообщения. Теория дискретизации, квантования, а также восстановления первичного сигнала по его выборкам рассматривалась в гл.2 (разд.2.7), а также в гл.3 (разд.3.8.3, 3.9.5).

7.11.3. Мощность шума квантования

Под шумом квантования понимают совокупность разностей

$$\varepsilon = \sum_k \Delta_k = \sum_k b(k\Delta t) - b_d(k\Delta t), \quad k = 0, 1, 2, 3, \dots$$

между отсчетными значениями первичного сигнала (сообщения) и ближайшими к этим значениям уровнями квантования

(см. рис.3.54). При равномерном квантовании, которое наиболее распространено на практике, максимальное значение шума квантования не превосходит половины интервала квантования, т.е.

$$\varepsilon_{\text{макс}} \leq 0,5\Delta b.$$

Если полный размах непрерывного первичного сигнала $b(t)$ равен $2b_{\text{макс}}$, то число уровней квантования

$$L = \frac{2b_{\text{макс}}}{\Delta b} + 1. \quad (7.100)$$

При $L \gg 1$ принято считать, что шум квантования распределен по равномерному закону в интервале от $-0,5\Delta b$ до $+0,5\Delta b$.

Средняя мощность шума квантования

$$P_{\varepsilon} = M \langle \varepsilon^2 \rangle = \frac{1}{\Delta b} \int_{-\Delta b/2}^{\Delta b/2} \varepsilon^2 d\varepsilon = \frac{\Delta b^2}{12}.$$

Как показано выше (разд.7.9.4), средняя мощность ненормированного первичного сигнала (сообщения):

$$P_b = b_{\text{макс}}^2 / \Pi^2,$$

где Π — пик-фактор сообщения.

Отношение сигнал-шум квантования:

$$\frac{P_b}{P_{\varepsilon}} = \frac{b_{\text{макс}}^2 \cdot 12}{\Pi^2 \Delta b^2} \cong \frac{3L^2}{\Pi^2} = \frac{3 \cdot 2^{2n}}{\Pi^2}. \quad (7.101)$$

Это отношение зависит от числа уровней L квантования и связанного с этим уровнем длины n двоичной кодовой комбинации. Выбирая число уровней квантования (число разрядов n двоичного кода) можно снизить влияние шума квантования на верность передачи до минимума. Следует, однако помнить, что с увеличением числа разрядов увеличивается сложность ЦАП и АЦП и ширина спектра двоичной кодовой комбинации.

Воздействие шума квантования на качество передачи можно также уменьшить, применяя вместо равномерного *неравномерное квантование*. Неравномерность квантования обусловлена зависимостью шага квантования Δb от уровня входного сигнала. Этот шаг увеличивается с ростом уровня входного сигнала и малые уровни сигнала имеют меньшую ошибку квантования по сравнению с большими. Если плотность распределения вероятности непрерывного сообщения такова, что большие уровни встречаются реже, чем малые, то при усреднении дисперсия ошибки уменьшается. Как показано в гл.4 (разд.4.7), плотности распределения непрерывных сообщений, например, речи удовлетворяют этому условию, т.е. весьма близки к нормальному закону.

7.11.4. Влияние ошибок приема на точность восстановления сообщения

При приеме двоичные кодовые комбинации преобразуются в квантованные выборочные значения. Если квантование равномерное, то

$$b_d^*(k\Delta t) = \Delta b \sum_{i=1}^n \alpha_i^* 2^{i-1}, \quad (7.102)$$

где $\alpha_i^* = \{0,1\}$ — элементы принимаемой двоичной кодовой комбинации.

Из-за шума некоторые элементы передаваемой двоичной кодовой комбинации передаются с ошибкой, поэтому вместо истинного значения α_i получаются оценки α_i^* .

Разность

$$\gamma_k = b_d(k\Delta t) - b_d^*(k\Delta t)$$

определяет дополнительную шумовую составляющую, обусловленную этими ошибками. С учетом (7.102) эта разность

$$\gamma_k = \Delta b \sum_{i=1}^n (\alpha_i - \alpha_i^*) 2^{i-1} = \Delta b \sum_{i=1}^n \beta_i 2^{i-1},$$

$\beta_i = \alpha_i - \alpha_i^*$, $\alpha = \{0,1\}$ — элементы передаваемой двоичной кодовой комбинации.

Случайная величина β_i принимает значения, равные 1, -1 с вероятностями:

$P(0) P(1/0)$ (при передаче "0");

$P(1) P(0/1)$ (при передаче "1").

Вероятность ошибки при передаче двоичной кодовой комбинации по симметричному дискретному каналу связи

$$p = P(0) P(1/0) + P(1) P(0/1),$$

где $P(0)$, $P(1)$ — априорные вероятности передачи символов "0" и "1" соответственно; $P(1/0)$, $P(0/1)$ — переходные вероятности.

Если считать, что $P(0) = P(1) = 0,5$, то для симметричного дискретного канала

$$P(0/1) = P(1/0) = p.$$

Математическое ожидание и дисперсия дискретной случайной величины β_i

$$M\{\beta_i\} = 1P(1/0)P(0) + (-1)P(0/1)P(1) = 1 \cdot p \cdot 0,5 - 1 \cdot p \cdot 0,5 = 0,$$

$$M\{\beta_i^2\} = (1)^2 P(1/0)P(0) + (-1)^2 P(0/1)P(1) = p \cdot 0,5 + p \cdot 0,5 = p.$$

Среднее значение шумовой составляющей

$$M\{\gamma_k\} = \Delta b \sum_{i=1}^n M\{\beta_i\} 2^{i-1} = 0.$$

Дисперсия шумовой составляющей

$$M\{\gamma_k^2\} = \Delta b^2 M\left\{\sum_{i=1}^n \beta_i 2^{i-1}\right\}^2 = p \Delta b^2 \sum_{i=1}^n 2^{2(i-1)}.$$

Последняя формула вычислена при условии, что ошибки приема различных символов независимы.

Дисперсия определяется вероятностью p ошибки при приеме символов двоичной кодовой последовательности и тем больше, чем больше эта вероятность.

7.11.5. Дифференциальная импульсно-кодовая модуляция

Этот вид модуляции является разновидностью ИКМ.

Если спектр сообщения в занимаемой им полосе частот равномерен, то отсчеты, взятые в соответствии с теоремой Котельникова через интервал времени $t = 1/2F_v$, где F_v — верхняя граничная частота спектра, являются взаимно-некоррелированными.

Спектры реальных сообщений таких, как речь, телевизионные изображения или данные телеметрических измерений, не являются равномерными, а корреляция между отсчетами, взятыми через интервал $\Delta t \leq 1/2F_v$, не равна нулю.

Передача взаимно-коррелированных отсчетов, как это следует из рассмотренной в гл.6 теории, является неэффективной, так как канал будет загружен, наряду с полезной, также и избыточной, известной на приемном конце, информацией. Очевидно, что устранение этой избыточности создает условия для передачи большего количества информации.

При дифференциальной импульсно-кодовой модуляции (ДИКМ) по каналу связи передаются не значения отсчетов сообщения, а разность между отсчетами $b(k\Delta t)$ и значением сигнала предсказания $\tilde{b}(k)$, сформированного из предыдущих отсчетов. При сильных корреляционных связях значение $\tilde{b}(k)$ формируется по правилу

$$\tilde{b}(k) = b(k-1). \quad (7.103)$$

Разность отсчетов, передаваемая по каналу связи,

$$\varepsilon(k) = b(k) - b(k-1).$$

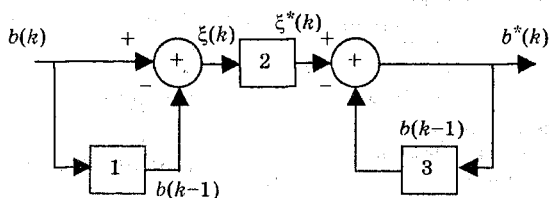


Рис.7.35. Структурная схема системы связи с ДИКМ: 1,3 — предсказатели; 2 — линия связи

Если $b(k) \cong b(k-1)$, то передаваемая разность $\varepsilon(k) = 0$.

На рис.7.35 приведена упрощенная структурная схема системы с ДИКМ, в которой реализован способ устранения избыточности из передаваемого сообщения.

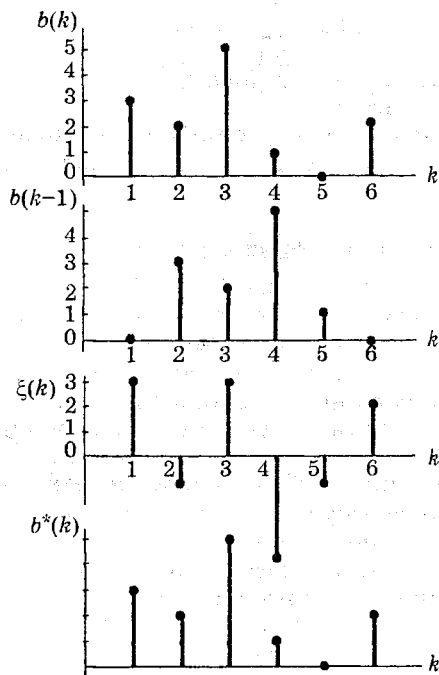


Рис.7.36. Диаграммы, поясняющие работу схемы на рис.7.35

Выборки сообщения подаются на вычитающее устройство, на второй вход которого поступают значения сигнала предсказания. При реализации правила (7.103) предсказатель является идеальной линией задержки, которая задерживает без искажений выборки сообщения на время, равное Δt .

В приемнике содержится точно такой же предсказатель, а вместо вычитающего устройства включается сумматор. На его выходе формируется оценка:

$$b^*(k) = \varepsilon^*(k) + b^*(k-1). \quad (7.104)$$

Диаграммы, приведенные на рис.7.36, поясняют работу схемы при отсутствии помех в линии связи.

Применяя правило формирования оценки (7.104), последовательно находим значения оценок:

$$b^*(0) = \varepsilon^*(0) + b^*(-1) = \varepsilon^*(0),$$

так как $b^*(-1) = 0$.

$$b^*(1) = \varepsilon^*(1) + b^*(0) = \varepsilon^*(1) + \varepsilon^*(0).$$

$$b^*(2) = \varepsilon^*(2) + b^*(1) \text{ и т.д.}$$

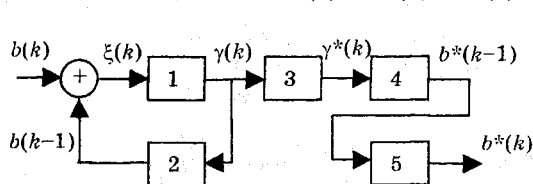


Рис.7.37. Структурная схема системы связи с δ -модуляцией: 1 — квантователь; 2,4 — интеграторы; 3 — линия связи; 5 — фильтр нижних частот

Дельта-модуляция (ДМ), отличается от ДИКМ тем, что по линии передается только знак сигнала ошибки. Квантованный сигнал ошибки

$$\varepsilon_{\text{кв}}(k) = \gamma(k)\Delta b,$$

где

$$\gamma(k) = \begin{cases} +1, & \text{если } \varepsilon(k) \geq 0, \\ -1, & \text{если } \varepsilon(k) < 0. \end{cases}$$

В передатчике (рис.7.37) квантованный сигнал ошибки подается на интегратор, на выходе которого вычисляется сумма всех предыдущих квантованных сигналов ошибки:

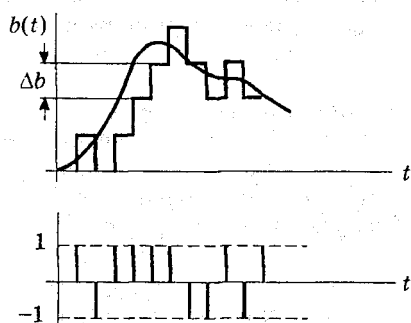


Рис. 7.38. Диаграммы, поясняющие принцип формирования сигналов δ -модуляции

$$b(k-1) = \sum_{i=0}^{n-1} \varepsilon_{\text{КВ}}(i) = \Delta b \sum_{i=0}^{n-1} \gamma(k). \quad (7.105)$$

В вычитающем устройстве входной первичный сигнал $b(k)$ (сообщение) сравнивается с выходным напряжением интегратора.

Если $b(k) > b(k-1)$, то на выходе квантователя формируется "1" а, если $b(k) < b(k-1)$, то — формируется "–1" (рис. 7.38).

В приемнике содержится точно такой же интегратор, как и в передатчике. Разнополярные импульсы с выхода линии

связи суммируются в этом интеграторе по правилу, аналогичному (7.105). Полученное на его выходе ступенчатое напряжение сглаживается фильтром нижних частот. Оценка передаваемого сообщения $b^*(t)$ на выходе ФНЧ зависит от уровня шума в линии связи и от шага квантования Δb . Размер шага Δb определяет величину шума квантования. Чем меньше шаг и чем выше частота дискретизации, тем меньше и шум квантования. К сожалению, уменьшение шага квантования может привести к искажениям из-за быстрого изменения сообщения. В этом случае возникают "перегрузки по наклону", когда оценка $b^*(t)$ запаздывает по отношению к сообщению $b(t)$.

Для того, чтобы устранить это нежелательное явление, нужно, чтобы интервал квантования удовлетворял неравенству:

$$\Delta b \geq |b'(t)|_{\text{макс}} \Delta t, \quad (7.106)$$

где $|b'(t)|_{\text{макс}}$ — максимальное значение крутизны сообщения.

Учитывая, что

$$\Delta b = \frac{|b(t)|_{\text{макс}}}{L},$$

где $|b(t)|_{\text{макс}}$ — максимальное значение сообщения; L — число уровней квантования, из (7.106) найдем выражение для интервала дискретизации

$$\Delta t \leq \frac{|b(t)|_{\text{макс}}}{L |b'(t)|_{\text{макс}}}.$$

С ростом Δb уменьшается "перегрузка по наклону", но растет шум квантования. Поэтому окончательный выбор постоянного интервала квантования производится, исходя из компромиссного удовлетворения этим двум требованиям.

Для передачи телефонных сообщений требуемая частота дискретизации не менее 150...200 кГц. При восьмиразрядной ИКМ тактовая частота составляет 64 кГц, т.е. меньше, чем при ДМ. Это означает, что ширина полосы линейного тракта для передачи сигнала при ДМ больше. Требуемая частота дискретизации существенно меньше при использовании ДМ с переменным шагом квантования. Системы связи, в которых используется такой вид модуляции, являются *адаптивными*. В этих системах удастся снизить частоту дискретизации до значений, которые даже ниже, чем в системах с ИКМ.

Контрольные вопросы

1. Какое приемное устройство называют оптимальным?
2. Каким алгоритмом определяется структура оптимального приемника сигналов, известных точно?
3. Из каких операций состоит алгоритм корреляционного приема?
4. Какую импульсную характеристику и комплексный коэффициент передачи имеет линейный фильтр, согласованный при белом шуме с сигналом $u(t)$?
5. Как можно записать алгоритм оптимального некогерентного приема дискретных сообщений при неопределенной фазе сигнала?
6. Что понимают под потенциальной помехоустойчивостью системы связи? Какими формулами выражается потенциальная помехоустойчивость при приеме АМ, ФМ и ЧМ сигналов?
7. Какие способы борьбы с импульсными, сосредоточенными и межсимвольными помехами Вы знаете?
8. Как количественно оценивается помехоустойчивость при приеме непрерывных сообщений?
9. Что такое выигрыш и обобщенный выигрыш? Найдите их для АМ, АМ-ПН, АМ-ОБП, ФМ, ЧМ.
10. Чем объяснить пороговый эффект в широкополосных системах с ФМ и ЧМ? Почему пороговый эффект отсутствует в системах с АМ, АМ-ПН?
11. В чем отличие между оптимальной линейной и нелинейной фильтрацией?
12. Какие достоинства и недостатки цифровых методов передачи непрерывных сообщений Вы знаете?
13. Как оценивается помехоустойчивость ИКМ?
14. Поясните принцип действия систем с дельта-модуляцией. Какие специфические ошибки появляются при ДМ?

Глава 8. ПРИНЦИПЫ МНОГОКАНАЛЬНОЙ СВЯЗИ И РАСПРЕДЕЛЕНИЯ ИНФОРМАЦИИ

8.1. Основы теории разделения сигналов

Система связи является многоканальной, если она обеспечивает передачу нескольких сообщений по одной общей линии связи. Многоканальная передача сообщений позволяет приблизить скорость передачи информации к пропускной способности линий связи (кабельных, волоконно-оптических, радио и т.д.), которая намного больше производительности источника сообщений. Очевидно, что суммарная производительность нескольких независимых источников должна быть меньше пропускной способности линии C , т.е.

$$\sum_{k=1}^N H'_k \leq C,$$

где H'_k — производительность k -го источника; N — число источников.

Улучшается также и такой важный показатель, как стоимость W строительства и эксплуатации одного канала-километра линии связи:

$$W = V/NL,$$

где V — стоимость строительства и эксплуатации системы в целом; L — длина линии в км; N — число каналов в системе.

Стоимость строительства и эксплуатации $V = V_{\text{л}} + V_{\text{а}}$, где $V_{\text{л}}$, $V_{\text{а}}$ — части стоимости, приходящиеся соответственно на линейные сооружения и аппаратуру. В последнее время стоимость аппаратуры возрастает. Если положить, что $V_{\text{л}} = V_{\text{а}}$ то $V = 2V_{\text{л}}$ и

$$W = 2V_{\text{л}}/NL.$$

Если в одноканальной системе стоимость аппаратуры $V_{\text{а}} \ll V_{\text{л}}$, то

$$W = \frac{2}{N} W_0,$$

где $W_0 = V_{\text{л}}/L$ стоимость одного канала-километра линии в одноканальной системе. В этом случае многоканальная система выигрывает по параметру W в $N/2$ раз.

На рис.8.1 приведена укрупненная структурная схема многоканальной системы связи. Первичные сигналы $a_1(t)$, $a_2(t)$, ..., $a_N(t)$, обычно имеющие одинаковые спектральные плотности, преобразуются в формирователе 1 в групповой сигнал:

$$u(t) = \sum_{i=1}^N u_i(t),$$

где $u_i(t)$ — канальные сигналы, однозначно связанные с первичными сигналами $a_i(t)$; N — число первичных сигналов.

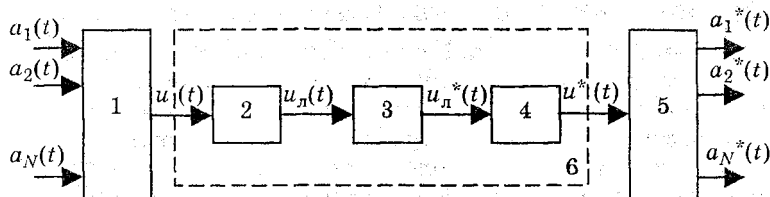


Рис.8.1. Укрупненная структурная схема многоканальной системы связи: 1 — формирователь группового сигнала; 2 — передатчик; 3 — линия связи; 4 — приемник; 5 — селектор канальных сигналов; 6 — групповой канал

Ширина спектра группового сигнала по сравнению с шириной спектра первичных сигналов увеличивается не менее чем в N раз. Расширение ширины спектра нежелательно главным образом по двум причинам:

- 1) из-за ограниченности частотного диапазона линий связи;
- 2) из-за возрастания сложности аппаратуры, следовательно, и ее стоимости.

Однако расширение ширины спектра является неизбежной платой за возможность разделения канальных сигналов.

Передатчик служит для согласования группового сигнала с параметрами линии связи. С этой целью групповой сигнал преобразуется в линейный $u_n(t)$. В приемнике линейный сигнал $u_n^*(t)$ преобразуется в групповой $u^*(t)$. В селекторе канальных сигналов, который является многоканальным приемником, из группового сигнала выделяются канальные сигналы $u_i^*(t)$, $i = 1, 2, \dots, N$. Эти сигналы затем преобразуются в первичные сигналы $a_i^*(t)$, $i = 1, 2, \dots, N$. Если работу k -го канального приемника описать линейным оператором L_k , то сигнал на его выходе при отсутствии помех в линии связи

$$L_k[u(t)] = L_k\left[\sum_{i=1}^N u_i(t)\right] = \sum_{i=1}^N L_k[u_i(t)], \quad (8.1)$$

Для разделения сигналов нужно выполнить условия:

$$L_k[u_i(t)] = \begin{cases} u_k(t), & i = k, \\ 0, & i \neq k. \end{cases} \quad (8.2)$$

Если записать сигнал k -го канала в виде

$$u_k(t) = c_k \varphi_k(t),$$

где $\varphi_k(t)$ — функция переносчика; c_k — некоторый коэффициент, отображающий передаваемое сообщение, то выражение (8.2) можно записать также в виде:

$$L_k[u_i(t)] = \begin{cases} c_k \varphi_k(t), & i = k, \\ 0, & i \neq k. \end{cases} \quad (8.3)$$

В частном случае откликом на сигнал $u_k(t)$ может быть некоторое число α_k , однозначно связанное с c_k . При выполнении условий (8.2) идеальное k -е приемное устройство реагирует только на сигнал $u_k(t)$ и не откликается на остальные, т.е. обладает свойством избирательности.

8.2. Методы частотного, временного и фазового разделения сигналов

8.2.1. Частотный метод разделения сигналов

В системе связи с частотным разделением сигналов (рис.8.2) первичные сигналы $a_1(t)$, $a_2(t)$, ..., $a_N(t)$ помощью модуляторов (рис.8.2) первичные сигналы $a_1(t)$, $a_2(t)$, ..., $a_N(t)$ помощью модуляторов и полосовых фильтров преобразуются в каналные сигналы $u_1(t)$, $u_2(t)$, ..., $u_N(t)$ с неперекрывающимися спектрами (рис.8.3, б).

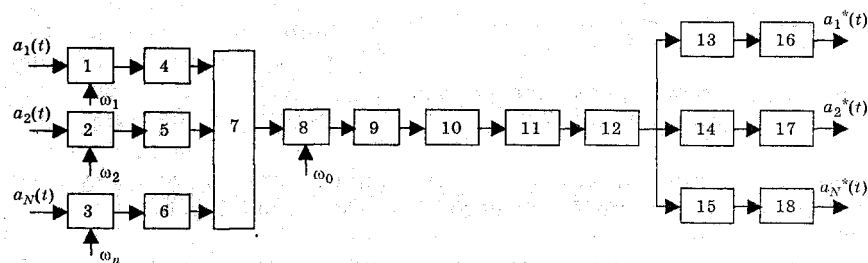


Рис.8.2. Структурная схема многоканальной системы связи с частотным разделением сигналов: 1,2,3 — модуляторы; 4,5,6 — полосовые фильтры; 7 — сумматор; 8 — групповой модулятор; 9 — усилитель; 10 — линия связи; 11 — усилитель; 12 — детектор; 13,14,15 — полосовые фильтры; 16,17,18 — каналные детекторы

Это преобразование обеспечивается выбором:

1) вида модуляции, при которой ширина спектра канального сигнала минимальна (например, такой модуляцией является АМ-ОБП);

2) поднесущих частот ω_i , $i = 1, 2, \dots, N$, значения которых не равны, причем разность между частотами в соседних участках диапазона выбирается не менее разности $\omega_i - \omega_{i-1} = \Omega_B$ где Ω_B — верхняя граничная частота спектра первичного сигнала $a_i(t)$ (рис.8.3, а); 3) коэффициента передачи полосового фильтра, который при минимальной полосе пропускания передает спектральные компоненты собственного канального сигнала без искажений и подавляет спектральные компоненты соседних канальных сигналов.

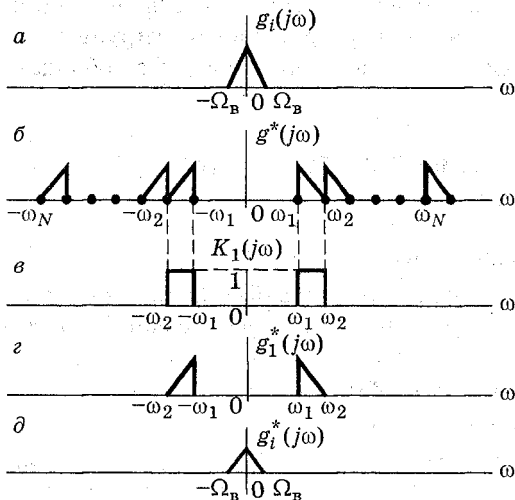


Рис.8.3. Диаграммы, поясняющие работу схемы на рис.7.2

Групповой сигнал с выхода сумматора канальных сигналов преобразуется в линейный при помощи группового модулятора и усилителя. В групповом модуляторе групповой сигнал модулирует один из параметров (например частоту) гармонического сигнала переносчика. В результате передача полезной информации происходит в диапазоне частот, который отводится для передачи линейного сигнала. Область этого диапазона определяется средним значением несущей частоты ω_0 сигнала переносчика. Групповой усилитель линейного сигнала доводит этот сигнал до уровня, необходимого для нормальной работы приемника.

На приемном конце линейный сигнал детектируется. Полученный на выходе детектора групповой сигнал фильтруется полосовыми фильтрами (рис.8.3, б). Каждый из них пропускает лишь те частотные компоненты, которые попадают в полосу пропускания данного фильтра. Спектральные компоненты вне полосы пропускания фильтр подавляет.

Модулированные сигналы с выхода фильтров детектируются. Способ детектирования определяется видом модуляции. Для АМ-ОБП сигналу используются синхронные детекторы. Рис.8.3, в, г, д поясняют работу 1-го канального приемника.

Групповой сигнал с выхода сумматора канальных сигналов преобразуется в линейный при помощи группового модулятора и усилителя. В групповом модуляторе групповой сигнал модулирует один из параметров (например частоту) гармонического сигнала переносчика. В результате передача полезной информации происходит в диапазоне частот, который отводится для передачи линейного сигнала. Область этого диапазона определяется средним значением несущей частоты ω_0 сигнала переносчика.

Определим оператор k -го канального приемника в схеме рис.8.2, используя равенства (8.2):

$$L_k[u^*(t)] = L_k \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} g^*(j\omega) e^{j\omega t} d\omega \right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} L_k g^*(j\omega) e^{j\omega t} d\omega = \begin{cases} u_k^*(t), & k = i, \\ 0, & k \neq 0. \end{cases} \quad (8.4)$$

Из (8.4) следует:

$$L_k g^*(j\omega) = g^*(j\omega) K_k(j\omega) = g_k^*(j\omega). \quad (8.5)$$

Следовательно, для разделения сигналов необходимо, чтобы коэффициент передачи $K_k(j\omega)$ имел П-образный вид в диапазоне частот от ω_k до $\omega_k + \Omega_{\text{в}}$. При перемножении спектральной плотности группового сигнала на этот коэффициент передачи получим требуемую спектральную плотность k -го канального сигнала.

Для идеального разделения сигналов, кроме того, нужно выполнить дополнительное условие: спектральные составляющие сигналов $u_k(t)$ должны полностью находиться в пределах полосы, отведенной для этих сигналов. В этом случае они разделяются без взаимных помех. К сожалению, упомянутые выше условия принципиально невыполнимы. Амплитудно-частотные характеристики фильтров имеют плавные переходы от зоны "прозрачности" к зоне "непрозрачности", а спектральные составляющие сигналов содержатся также и вне заданной полосы частот. В результате возникают взаимные помехи между соседними каналами. Для ослабления взаимного влияния каналов и уменьшения помех вводят защитные частотные интервалы, что снижает эффективность использования полосы пропускания линии связи. Так, например, в стандартных телефонных многоканальных системах связи каждому телефонному каналу отводится линейная полоса частот 4 кГц, хотя частотный спектр передаваемых тональных звуковых сигналов лежит в пределах от 300 до 3400 Гц.

8.2.2. Временное и фазовое разделение сигналов

Принцип передачи при *временном способе разделения* канальных сигналов заключается в их последовательной передаче по групповому каналу.

Структурная схема многоканальной системы связи приведена на рис.8.4, а временные диаграммы — на рис.8.5 и рис.8.6.

Электронные ключи, на первые входы которых подаются первичные сигналы (рис.8.5, *в*, *е*, *и*), управляются импульсами каналов (рис.8.5, *б*, *д*, *з*). Эти импульсы поочередно выдаются распределителем импульсов каналов (РИК), запуск которого производится тактовыми импульсами с выхода генератора импульсов (ГИ). Период следования тактовых импульсов обычно выбирается

равным $T = 1 / 2F_c$, где F_c — верхняя граничная частота спектра первичных сигналов.

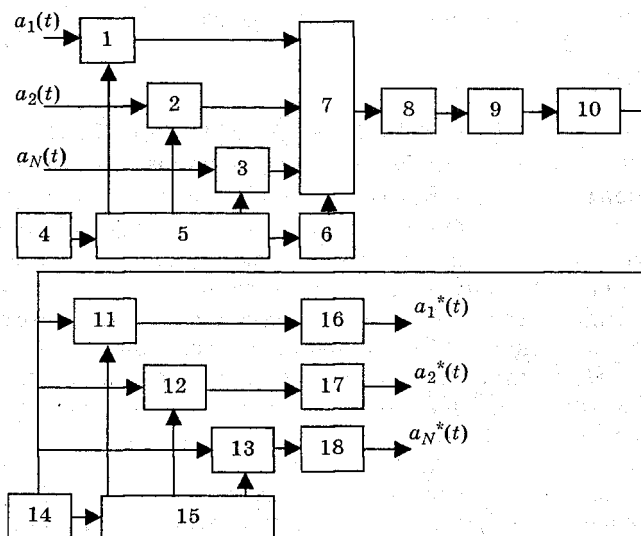


Рис.8.4. Схема многоканальной системы связи с временным разделением сигналов: 1,2,3,11,12,13 — электронные ключи; 4 — генератор импульсов; 5,15 — распределитель импульсов каналов; 6 — генератор синхроимпульсов; 7 — сумматор; 8 — передатчик; 9 — линия связи; 10 — приемник; 14 — схема синхронизации; 16,17,18 — фильтры нижних частот

При поступлении импульсов каналов выходы электронных ключей соединяются с их входами что позволяет подать выборки первичных сигналов на входы сумматора. На сумматор поступает также синхронизирующий импульс (рис.8.5, *a*) с выхода генератора синхроимпульса, который вместе с канальными выборками (рис.8.5, *г, ж, к*) образуют групповой сигнал $u(t)$ (рис.8.5, *л*). Передатчик служит для согласования параметров сигнала $u(t)$ с параметрами линии связи.

Приемник преобразует линейный $u_{л}^*(t)$ в групповой сигнал $u^*(t)$.

Схема синхронизации, откликающаяся только на синхронизирующие импульсы (рис.8.5, *a*), формирует на своем выходе импульсы синхронизации, которые запускают РИК приемника. Его схема не отличается от схемы РИК передатчика. На выходе РИК приемника формируются импульсы каналов, показанные на рис.8.6, *б, д, з*. Этими импульсами управляют электронные

ключи. При их поочередном замыкании каналные сигналы $u_i^*(t)$, $i = \overline{1, N}$ (рис.8.6, в, е, и) поступают на входы канальных демодуляторов (фильтров нижних частот).

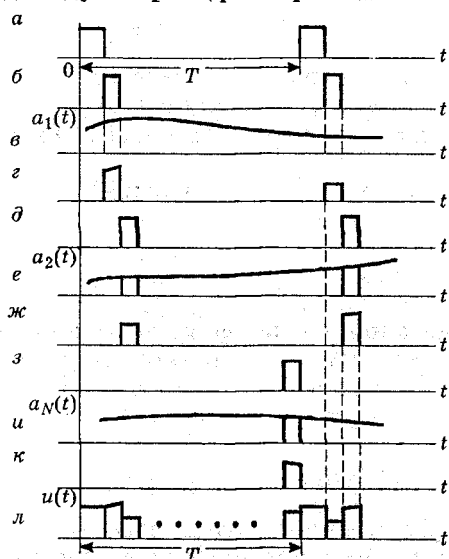


Рис.8.5. Временные диаграммы, поясняющие работу передающей части системы, показанной на рис.8.4

т.е. восстановление непрерывных первичных сигналов по их дискретным значениям. Восстановленные непрерывные первичные сигналы $a_1^*(t)$, $a_2^*(t)$, ..., $a_N^*(t)$ (рис.8.6, г, ж, к) далее поступают к получателям сообщений.

Работу каждого k -го электронного ключа, можно описать следующими уравнениями:

$$Ku^*(t) = \sum_{i=1}^N Ku_i^*(t) = \begin{cases} u_k^*(t), & i = k; \\ 0, & i \neq k, \end{cases} \quad (8.6)$$

где

$$K = \begin{cases} 1, & (k-1)\tau_n < t \leq k\tau_n, \\ 0, & k\tau_n < t < (k-1)\tau_n, \end{cases}$$

τ_n — длительность импульса канала.

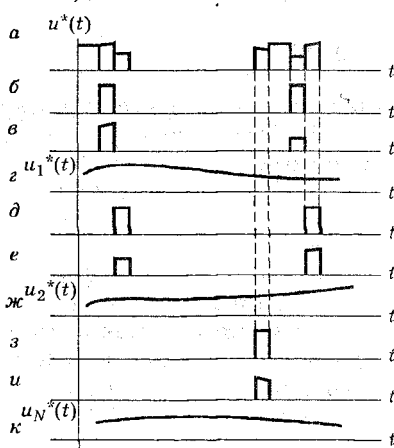


Рис.8.6. Временные диаграммы, поясняющие работу приемной части системы, показанной на рис.8.4

С подачей канальных сигналов (рис.8.6, в, е, и) на фильтры нижних частот производится их демодуляция,

Сравнивая (8.6) и (8.2), находим, что оператор L_K в схеме рис.8.4 описывает параметрическую линейную систему (электронный ключ).

Как следует из приведенного выше описания, работа многоканальной системы с временным разделением сигналов включает следующие операции.

1. Дискретизацию первичных сигналов в неперекрывающиеся моменты времени с интервалом дискретизации $1/2F_c$.

2. Взятие выборок из первичных сигналов в определенные выше моменты времени.

3. Суммирование этих выборок с сигналом синхронизации.

4. Выделение сигнала синхронизации в приемной части системы.

5. Разделение канальных сигналов.

6. Демодуляцию канальных сигналов.

Принцип временного разделения сигналов используется в *цифровых системах передачи*. Эти системы отличаются от рассмотренной тем, что вместо импульсно-модулированных сигналов в них используются ИКМ сигналы. Для этого выборки (канальные сигналы) квантуются по уровню, а затем производится кодирование ближайшего к выборке уровня, т.е. представление его двоичной кодовой комбинацией.

В системах с временным разделением сигналов возникают взаимные помехи между каналами в основном по двум причинам.

1. Из-за ограниченности полосы частот импульсы кодовых последовательностей "расплываются" во времени. Это приводит к межсимвольной интерференции, когда импульсы одного канала накладываются на импульсы других каналов.

2. Из-за присутствия помех в линии связи моменты появления синхроимпульсов и импульсов каналов в приемнике становятся случайными (флуктуирующими).

Для уменьшения уровня взаимных помех вводят "защитные" временные интервалы, что при сохранении скорости передачи приводит к увеличению ширины спектра группового сигнала.

Фазовое разделение сигналов строится с использованием различия сигналов по фазе.

Пусть информация в N каналах передается изменением амплитуды непрерывных косинусоидальных сигналов с одинаковой несущей частотой ω_0 . Требуется разделить эти сигналы с использованием только различия в их начальных фазах.

Сигналы равны:

$$\left. \begin{aligned} u_1(t) &= A_1(t) \cos(\omega_0 t + \varphi_1); \\ u_2(t) &= A_2(t) \cos(\omega_0 t + \varphi_2); \\ &\dots\dots\dots \\ u_N(t) &= A_N(t) \cos(\omega_0 t + \varphi_N). \end{aligned} \right\} \quad (8.7)$$

Как показывает анализ, различение сигналов возможно, если система содержит только два канала, по которым передаются косинусная и синусная составляющие:

$$u_1(t) = A_1(t) \cos(\omega_0 t + \varphi_1), \quad u_2(t) = A_2(t) \sin(\omega_0 t + \varphi_1),$$

а выделение первичных сигналов производится с использованием синхронного детектирования.

8.3. Другие виды разделения сигналов

8.3.1. Разделение сигналов по форме

Передаваемые каналные сигналы многоканальных систем с частотным разделением не перекрываются в частотной, а каналные сигналы многоканальных систем с временным разделением — во временной области.

Эти условия записываются соответственно в виде следующих равенств:

$$\int_{\Delta\omega} g_i(j\omega) g_j(j\omega) d\omega = \begin{cases} E_i, & i = j, \\ 0, & i \neq j, \end{cases} \quad (8.8)$$

где $\Delta\omega$ — частотная область, в которой располагаются спектральные составляющие группового сигнала, $g_i(j\omega)$, $g_j(j\omega)$ — спектральные плотности каналных сигналов системы с частотным разделением;

$$\int_0^T u_i(t) u_j(t) dt = \begin{cases} E_i, & i = j, \\ 0, & i \neq j, \end{cases} \quad (8.9)$$

где $u_i(t)$, $u_j(t)$ — сигналы i -го и j -го каналов системы с временным разделением, T — интервал времени, в котором располагаются каналные сигналы, причем $T = 1/2F_v$, где F_v — верхняя граничная частота спектра передаваемого сообщения.

Для уверенного разделения каналных сигналов необходимо:

а) в системе с частотным разделением соответствующим образом выбрать АЧХ полосовых фильтров и настроить их на определенные участки спектра передаваемого каналного сигнала;

б) в системе с временным разделением обеспечить поочередное подключение каналных приемников к тракту передачи группового сигнала в нужные моменты времени и на заданное время.

Кроме сигналов с неперекрывающимися спектрами и сигналов, неперекрывающихся во времени существует класс сигналов, которые могут передаваться одновременно и иметь перекрывающиеся частотные спектры.

Разделение этих сигналов принято называть *разделением по форме*.

К числу таких сигналов относятся последовательности Уолша, Радемахера и разнообразные шумоподобные последовательности.

Последовательности Уолша и Радемарха строятся на базе кодового алфавита 1, -1, а любые пары этих последовательностей удовлетворяют условию 8.9. Применение кодов Уолша и Радемахера связано с передачей по каналу связи специальных синхросигналов для поддержания определенных временных соотношений между принимаемыми и опорными кодовыми словами.

В случае использования шумоподобных последовательностей необходимости в передаче специальных синхросигналов нет, так как эту роль могут выполнять последовательности-переносчики информации.

Шумоподобные сигналы должны удовлетворять следующим условиям:

$$\int_0^T u_i(t)u_i(t+\tau)dt = \begin{cases} E, \tau = 0, \\ 0, -\tau_i \geq \tau \geq -T, \\ T \geq \tau \geq \tau_i, \end{cases} \quad (8.10)$$

$$\int_0^T u_i(t)u_j(t+\tau)dt = 0, i \neq j, \quad (8.11)$$

для $-T \leq \tau \leq T$, $i, j = 1, 2, 3, \dots, N$, где T — длительность шумоподобного сигнала; E — энергия сигнала; τ_i — длительность единичного интервала шумоподобного сигнала.

При выполнении условий (8.10) обеспечивается работа системы синхронизации без передачи специального синхросигнала, так как автокорреляционная функция любого канального сигнала имеет ярко выраженный пик при $\tau = 0$ и нулевые значения при сдвиге $-\tau_i \geq \tau > \tau_i$. При выполнении условий (8.11) обеспечивается разделение канальных сигналов, так как взаимокорреляционная функция для любой пары сигналов равняется нулю.

К сожалению, скалярные произведения (8.10) для $\tau \neq 0$ и (8.11) для реальных сигналов не равны нулю. Это приводит к снижению достоверности разделения канальных сигналов.

Структурная схема многоканальной системы связи с разделением сигналов по форме приведена на рис.8.7. Передающая часть системы содержит N идентичных модуляторов, сумматор и передатчик. В модуляторах в качестве несущих колебаний ис-

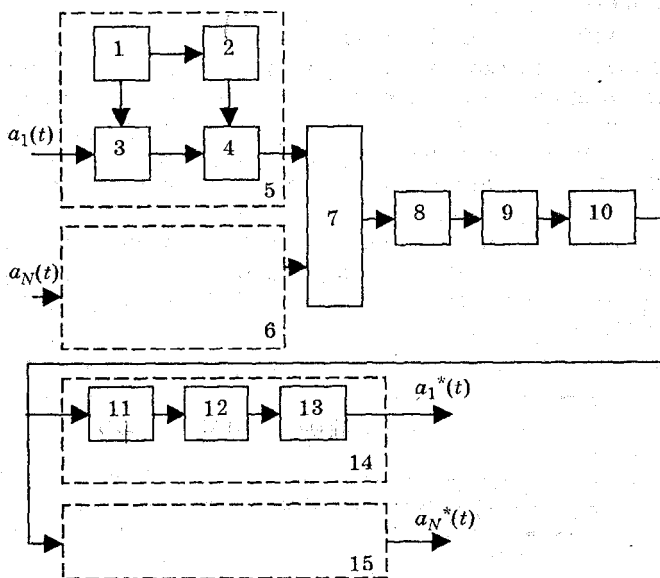


Рис.8.7. Структурная схема многоканальной системы связи с разделением сигналов по форме: 1 — генератор тактовых импульсов; 2 — генератор шумоподобного сигнала; 3 — АЦП; 4 — перемножитель; 5,6 — модуляторы; 7 — сумматор; 8 — передатчик; 9 — линия связи; 10 — приемник; 11 — согласованный фильтр; 12 — решающее устройство; 13 — ЦАП; 14,15 — демодуляторы

пользуются шумоподобные сигналы, а в качестве модулирующих — сфазированные с этими сигналами двоичные кодовые последовательности с выхода АЦП. Период шумоподобных сигналов выбирается равным длительности единичного элемента кодового слова с выхода АЦП. В процессе модуляции символу "1" двоичного кодового слова (диаграмма а на рис.8.8) соответствует полный период шумоподобного сигнала $T = N\tau_0$ (диаграмма б), а символу "0" — отсутствие этого сигнала. Если F_c — верхняя граничная частота спектра первичного сигнала, а L — число уровней квантования, то ширина спектра сигнала на выходе перемножителя (см. схему на рис.8.7)

$$\Delta f \geq 2F_c N \log_2 L, \quad (8.12)$$

где N — длина (период) шумоподобной последовательности.

Сравнивая формулу (8.12) с формулой (3.62) для ИКМ сигнала, находим, что ширина спектра каждого канального сигнала в N раз больше ширины ИКМ сигнала.

Отметим, что каждый канальный сигнал имеет свою форму, а временные процессы, протекающие в каналах, могут быть независимы. Групповой сигнал на выходе сумматора, равный сумме канальных сигналов, представляет собой случайный процесс, среднее значение и дисперсия которого зависит от загрузки отдельных каналов.

Приемная часть системы содержит приемник и N идентичных канальных приемников (демодуляторов). В структуру каждого демодулятора входит согласованный фильтр, решающее устройство и ЦАП.

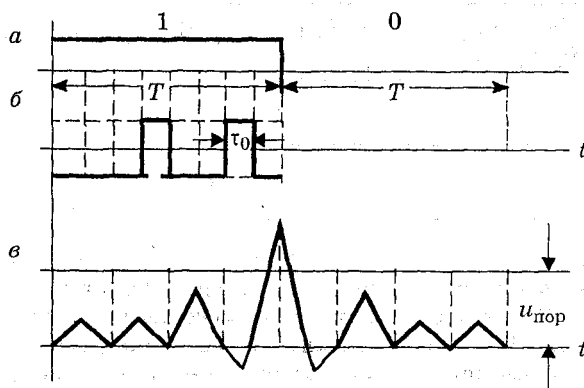


Рис.8.8. Временные диаграммы, поясняющие работу схемы, приведенной на рис.8.7

Каждый из согласованных фильтров откликается только на тот сигнал, с которым он согласован. Например, согласованный фильтр 11 первого канала откликается на сигнал, который формируется в первом модуляторе (рис.8.8, б). Отклик фильтра показан на рис.8.8, в. Сигналы других каналов и их отклики на рис.8.8 для простоты не показаны. В решающем устройстве отклик согласованного фильтра 11 огибающая радиосигнала сравнивается с заданным пороговым уровнем $U_{\text{пор}}$. Если происходит пересечение порога, то формируется оценка передаваемого символа, равная 1, а если пересечения не происходит, то формируется оценка, равная нулевому символу. Кодовые слова с выхода решающего устройства 12 поступают на ЦАП 13 и преобразуются в сообщение $a_1^*(t)$.

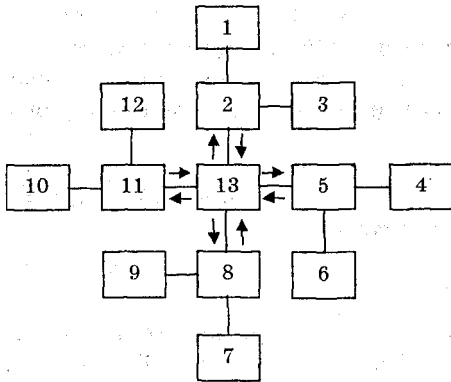


Рис.8.9. Структурная схема многоканальной асинхронной адресной системы связи:

1,4,7,10 — абоненты 1, i,k,N ;
2,5,8,11 — передатчики;
3,6,9,12 — генераторы адресного сигнала; 13 — линия связи

Демодуляция сигнала происходит в присутствии помехи, которая состоит из двух составляющих. Первая является известной по предыдущим главам суммой внутренней и внешней флуктуационных помех, а вторая — специфичной для систем с шумоподобными сигналами помехой. Эта помеха является суммой шумоподобных сигналов других каналов и называется структурной или взаимной помехой. Структурная помеха обусловлена тем, что системы используемых реальных сигналов являются "почти ортогональными", т.е. для них не выполняется условие (8.11). Ее уровень определяет-

ся значениями взаимокорреляционных функций между опорным канальным шумоподобным сигналом и присутствующими шумоподобными сигналами других каналов. С целью обеспечения заданного качества передаваемой информации, должны предусматриваться меры по уменьшению уровня этой структурной помехи. Рассмотренные принципы разделения сигналов по форме и построения многоканальной системы связи используется в *многоканальных асинхронных адресных системах связи (ААСС)*. В ААСС (рис.8.9) каждому абоненту присваивается один из "почти ортогональных" шумоподобных сигналов, который является адресом канала. Пусть, например абоненту 1 нужно связаться с абонентом " k ". С этой целью набирается номер абонента " k " и таким образом в генераторе адресного сигнала 1 устанавливается форма шумоподобного сигнала с номером " k ". Если число абонентов равно N , то и число набираемых форм также равно N . Шумоподобный сигнал с номером " k " посылается в линию связи и таким образом действует на входах приемников всех остальных абонентов. На шумоподобный сигнал " k " настроена приемная аппаратура только абонента " k ", поэтому связь устанавливается между абонентами 1 и " k ". Приемники других абонентов на этот шумоподобный сигнал не откликаются. Ответная информация от абонента " k " передается с использованием шумоподобного сигнала

с номером " k ". Важной особенностью ААСС является отсутствие центральной коммутационной станции. Все абоненты имеют прямой доступ к друг другу, а если используется радиолиния, то частотная перестройка приемо-передатчиков для вхождения в связь не производится.

В заключении отметим, что в технической литературе имеется описание ААСС, в которых используется от 1000 до 1500 каналов с 50...100 активными абонентами.

8.3.2. Комбинационное разделение

Комбинационное разделение сигналов применяется в мало-канальных системах с числом каналов, равным 2-4. Примерами систем, в которых используется комбинационные разделения являются двухканальная система частотного телеграфирования и двухканальная система фазового телеграфирования.

Рассмотрим двухканальную систему, по каналам которой передаются двоичные кодовые последовательности с элементами 0 и 1. Пусть значащие моменты в каналах совпадают. Выпишем в табл.8.1 возможные комбинации элементов в одинаковых единичных интервалах.

Таблица 8.1

Номер комбинации	Канал		Частота сигнала
	N1	N2	
1	0	0	f_1
2	1	0	f_2
3	0	1	f_3
4	1	1	f_4

Поставим в однозначное соответствие каждой кодовой комбинации сигнал заданной длительности и частоты. В табл.8.1 двоичной кодовой комбинации 00 соответствует сигнал заданной длительности с частотой f_1 ; двоичной кодовой комбинации 10 соответствует сигнал той же длительности с частотой f_2 и т.д.

Преобразование двоичных комбинаций в сигналы производится в преобразователе, выход которого соединен со входом линии связи (рис.8.10). Если число каналов равно n , а основание кодовых комбинаций равно m , то количество кодовых комбинаций

$$N = m^n.$$

В рассматриваемом примере $n = 2$, $m = 2$, $N = 2^2 = 4$.

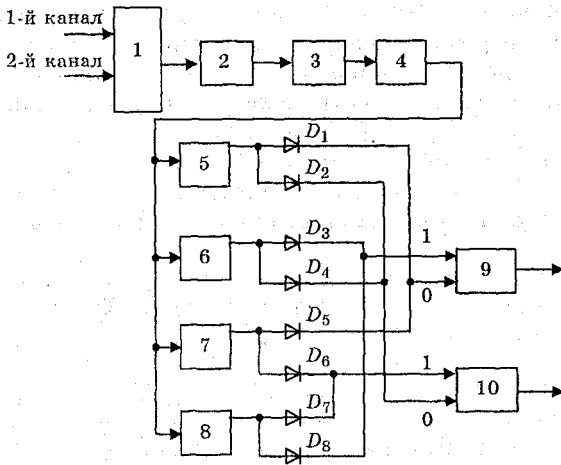


Рис.8.10. Структурная схема системы связи с комбинационным разделением сигналов: 1 — преобразователь двоичной комбинации в сигнал; 2 — передатчик; 3 — линия связи; 4 — приемник; 5,6,7,8 — фильтры; 9,10 — 1-й и 2-й формирователи

Линейный сигнал, отображающий комбинацию канальных сигналов, передается по линии связи. На приемном конце принятый сигнал разделяется полосовыми фильтрами Φ_1 , Φ_2 , Φ_3 и Φ_4 и детектируется. Если линейный сигнал имеет частоту f_1 , то на выходах детекторов D_1 и D_2 появятся видеопульсы, которые будут поданы на входы "0" формирователей 1 и 2. При наличии этих импульсов на их выходах формируются символы "0". Если линейный сигнал

имеет частоту f_2 , то видеопульсы появятся на выходах детекторов D_3 и D_4 . Первый из них появится на входе "1" преобразователя 1, а второй — на входе "0" преобразователя 2. В результате на выходах этих формирователей сформируются в соответствии с табл.8.1 символы "1" и "0". Аналогично декодируются сигналы с частотами f_3 и f_4 . Рассмотренный способ разделения сигналов называется комбинационным разделением, а система связи — двухканальной системой частотного телеграфирования.

8.3.3. Пространственное разделение

Наряду с временным, частотным разделением и разделением сигналов по форме в сетях электросвязи применяется *пространственное разделение*.

На выходах пространственно разнесенных источников сообщений образуются первичные сигналы, которые в системах проводной связи поступают на входы абонентских линий.

Объединение первичных сигналов в групповой при пространственном разделении связано с формированием пучков проводов и размещением их в одном кабеле.

Обычно в многоканальных системах связи каналы передачи образуются с помощью различных методов разделения сигналов.

Например, после пространственного часто используют временное или частотное разделение и реже — разделение сигналов по форме. Пространственное разделение сигналов в многоканальных системах передачи отличается от пространственного разделения сигналов в системах коммутации. Если в многоканальных системах разделяются сигналы заранее закрепленных друг с другом пар (источник сообщений — получатель сообщений), то в системах коммутации они разделяются в соответствии с адресами вызываемых абонентов.

Коммутация, при которой используется пространственное разделение сигналов, называется *пространственной*. Коммутация, при которой используется временное разделение сигналов, называется *временной*.

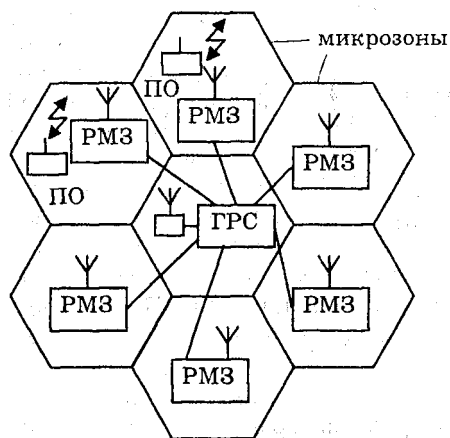


Рис.8.11. Структурная схема сотовой системы связи: ГРС — главная радиостанция; РМЗ — радиостанция микрозоны; ПО — подвижный объект

Пространственное разделение сигналов в радиосвязи определяется свойствами распространения радиоволн. Известно, что мощность сигнала на входе приемника прямо пропорциональна мощности передатчика и обратно пропорциональна квадрату расстояния между ними. Поэтому на некотором расстоянии от передатчика мощность на входе приемника становится недостаточной для уверенного приема. Это свойство используется в *сотовых системах связи*. Принцип пространственного разделения сигналов в такой системе поясняется рис.8.11.

Обслуживаемая территория разбивается на микрозоны без перекрытия и пропус-

ка участков. Наиболее подходящей фигурой, используемой при этом разбиении, является шестиугольник. В центре каждого шестиугольника устанавливается приемо-передающее оборудование с антенной, диаграмма направленности которой равномерна во всех направлениях, т.е. круговая. Мощность передатчиков выбирается сравнительно небольшой, чтобы обеспечить уверенный прием только в своей микрозоне. Из-за ослабления сигнала в других микрозонах в большинстве микрозон передатчики могут работать на одних и тех же частотах, так как мощность сигнала успевает

в достаточной степени уменьшиться примерно через две ячейки. Поэтому считается допустимым использовать одни и те же частоты через две ячейки.

В центре зоны устанавливается главная радиостанция, которая соединительными линиями связывается со всеми остальными, расположенными в других микрizonaх. Главная радиостанция соединяется, кроме того, и с телефонной сетью. Все радиостанции, включая и главную, обеспечивают связь абонентов, имеющих приемо-передающее оборудование и находящихся в одной из микрizonaх. Часть приемо-передающего оборудования устанавливается на подвижных объектах. При связи абонента автоматической телефонной станции с абонентом, например, подвижного объекта сигнал вызова из телефонной сети попадает на главную радиостанцию, от нее по соединительным линиям к одной из радиостанций микрizona, а затем по радиоканалу к абоненту подвижного объекта. Сотовые системы связи охватывают большие территории и большое число абонентов. Например, система AMPS (США) состоит из 10 ячеек, обслуживающих территорию 5400 км². Емкость системы может достигать до 100 тыс. абонентов.

8.4. Пропускная способность многоканальных систем

При анализе различных методов разделения сигналов было установлено, что в реальных многоканальных системах связи, наряду с аддитивными, всегда имеются *взаимные помехи от соседних каналов*. Средняя мощность переходных помех пропорциональна средней мощности сигналов, передаваемых по соседним каналам, т.е. равна μP_c , где μ — коэффициент пропорциональности, определяющий мешающее действие соседних каналов. Общая мощность помех в канале равна сумме мощностей аддитивной и переходной помехи:

$$P_{\Pi} = P_{\text{ш}} + \mu P_c. \quad (8.13)$$

Эта мощность определит пропускную способность канала:

$$C_K = F_K \log \left(1 + \frac{P_c}{P_{\text{ш}} + \mu P_c} \right). \quad (8.14)$$

Последняя формула получена из формулы (6.46) заменой (8.13). Пропускная способность многоканальной системы

$$C = \sum_{K=1}^N C_K = \sum_{K=1}^N F_K \log \left(1 + \frac{P_c}{P_{\text{ш}} + \mu P_c} \right). \quad (8.15)$$

Пропускная способность растет с уменьшением уровня помех. Однако при $P_{\text{ш}} \rightarrow 0$ этот рост ограничивается уровнем взаимных помех и определяется выражением:

$$C = \sum_{k=1}^N F_k \log \left(1 + \frac{1}{\mu} \right). \quad (8.16)$$

В многоканальной системе связи взаимные помехи не приводят к существенному снижению пропускной способности, если $P_{\text{ш}} \gg \mu P_c$ или $\mu \ll P_{\text{ш}}/P_c$.

Например, если отношение $P_c/P_{\text{ш}} = 10^4$, то $\mu \ll 10^{-4}$. Последнему условию отвечает число, равное 10^{-6} или меньше.

Отметим, что уменьшая μ путем введения защитных интервалов при временном и частотном методах разделения сигналов, тем самым увеличиваем ширину полосы каналов связи, что в свою очередь приводит к увеличению мощности шума $P_{\text{ш}}$.

В системах связи с разделением сигналов по форме μ можно уменьшить, увеличивая базу сигнала (произведение длительности сигнала на ширину его спектра). Но, с ростом базы увеличивается ширина полосы канала связи, следовательно и мощность шума $P_{\text{ш}}$. Эти замечания следует учитывать при проектировании каналов многоканальных систем.

8.5. Сеть связи и ее элементы. Основные понятия теории распределения информации

Сети распределения информации (телефона и телеграфа) являются системами распределенной многосторонней связи. Особенностью этих сетей (рис.8.12) является то, что в них обязательно содержатся средства автоматизированного распределения информации — сетевые коммутационные узлы (КУ). В этих узлах

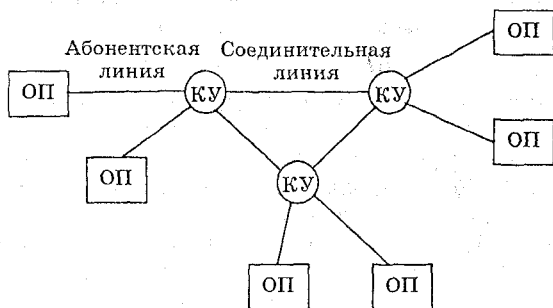


Рис.8.12. Упрощенная структурная схема сети связи: ОП — оконечный пункт; КУ — коммутационный узел

(рис.8.13) могут быть два типа коммутационных систем — *полнодоступные* и *неполнодоступные*. Первые позволяют устанавливать соединения между каждой входящей и каждой исходящей линиями узла. В неполнодоступных схемах к отдельным входящим линиям можно подключить только часть исходящих.

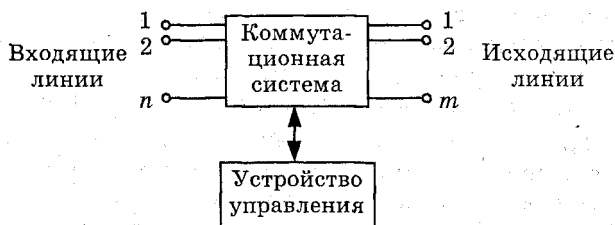


Рис.8.13. Упрощенная структурная схема коммутационного узла

Кроме сетевых коммутационных узлов, сеть содержит оконечные пункты и линии связи. Линию связи между оконечным пунктом и коммутационным узлом называют *абонентской*, а между коммутационными узлами — *соединительной*.

Различают два вида коммутации: *сообщений* и *каналов*. При коммутации сообщений записанные заранее сообщения передают таким образом, чтобы минимизировать время занятия приборов, обслуживающих входные линии. При этом сообщение может быть передано целиком или по частям.

При коммутации каналов соединительный путь между входом и выходом коммутационной системы предоставляется на время, достаточное для передачи всего сообщения.

Предметом теории распределения информации является изучение процессов обслуживания системами распределения информации (автоматическими телефонными станциями, коммутационными узлами и т.п.) поступающих потоков сообщений и их количественных характеристик.

Изучаемые в этой теории математические модели включают: модели потоков сообщений, модели систем обслуживания, характеристики качества и дисциплины обслуживания. Под *дисциплиной обслуживания поступающих сообщений* понимают способ и порядок обслуживания, а также режим исчисления выходов коммутационной системы.

Основные понятия, с которыми оперируют в теории распределения информации:

- 1) *вызов* — событие, заключающееся в поступлении в сеть сообщения или требования на обслуживание;
- 2) *поток сообщений* — множество последовательных или одновременно происходящих событий, связанных с поступлением большого числа вызовов, запросов, требований;
- 3) *длительность занятия* — время, в течение которого происходит установление соединения, передача сообщения и разъединение соединительных путей.

Потоки могут быть детерминированными и случайными. Детерминированным потоком называется *последовательность моментов поступления вызовов* (запросов, требований):

$$0 \leq t_1 \leq t_2 \leq \dots \leq t_n \dots, \quad (8.17)$$

где t_n ($n = 1, 2, \dots$) момент поступления n -го вызова. Значения t_n являются неотрицательными вещественными числами.

В сетях связи потоки, как правило, являются случайными, так как моменты поступления вызовов в этих потоках являются случайными величинами. Изучение характеристик случайных потоков являются частью теории распределения информации. К числу решаемых в рамках этой теории задач относятся задачи на определение вероятностных характеристик потоков по результатам экспериментальных наблюдений или испытаний. Другие задачи связаны с оценкой законов распределения потоков и сравнением их с пуассоновскими законами распределения.

Простейший стационарный пуассоновский поток определяется распределением вероятности:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (8.18)$$

где $P_k(t)$ вероятность того, что за время t произойдет k событий, связанных с поступлением " k " требований на установление соединения; λ — параметр распределения, называемый интенсивностью.

Простейший поток называется *стационарным*, если его вероятностные характеристики не меняются с течением времени. Вероятность $P_k(t)$ для стационарного потока зависит только от величины промежутка времени t и не зависит от его расположения на временной оси.

Параметр распределения λ определяется выражением:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{k \geq 1}(t, t + \Delta t)}{\Delta t}, \quad (8.19)$$

где $P_{k \geq 1}$ — вероятность поступления не менее одного вызова в промежутке времени $[t, t + \Delta t]$.

Простейший поток называется *ординарным*, если одновременно не происходят групповые вызовы, т.е. вероятность поступления двух и более вызовов за любой бесконечно малый промежуток времени Δt есть величина бесконечно малая. Стационарный поток без последствия называется *неординарным пуассоновским потоком*.

В теории распределения информации также исследуются характеристики следующих потоков: нестационарного пуассоновского; примитивного (пуассоновский поток второго рода); сглаженного пуассоновского; с повторными вызовами; Пальма и др.

Подробно с описанием этих потоков можно ознакомиться в [16]. В этой теории решаются также задачи анализа, связанные с определением качества обслуживания. При их решении учитываются параметры и свойства входящего потока сообщений, параметры и структура системы обслуживания и дисциплина обслуживания.

Одной из основных характеристик обслуживания вызовов коммутационной системой является *нагрузка*. Под нагрузкой понимают суммарное время обслуживания вызовов потока, поступающих в единицу времени. Различают поступающую нагрузку, пропущенную нагрузку и потерянную нагрузку. Единица измерения — *часо-занятие (ч-зан)*. Скорость приращения нагрузки называется *интенсивностью* нагрузки и измеряется в часо-занятиях, отнесенных к любому часу. Единица измерения интенсивности нагрузки — *эрланг*. Интенсивность нагрузки равна математическому ожиданию

$$Y' = \sum_{i=1}^9 iP_i(t), \quad (8.20)$$

где $P_i(t)$ — вероятность занятия i ($i = 0, 9$) линий в момент времени t в пучке емкостью 9.

При постоянной вероятности P_i в течение некоторого промежутка времени интенсивность нагрузки также постоянна:

$$Y' = \sum_{i=1}^9 iP_i.$$

Она определяет среднее число занятых линий.

В течение суток интенсивность нагрузки неравномерна. При расчете оборудования берут не среднее значение нагрузки, а тот отрезок времени, когда она максимальна. Это время, равное любому 60 минутам, в течение которого в среднем наблюдается наибольшая величина нагрузки, называется *час наибольшей нагрузки (ЧНН)*.

В теории распределения информации изучаются следующие системы: полnodоступная система с явными потерями; полnodоступная система с ожиданием; полnodоступная система с повторением; неполnodоступная система с явными потерями; звеньевые системы с явными потерями.

Поступающие потоки сообщений обслуживаются без потерь и с потерями. Потери отсутствуют, если для передачи каждого сообщения предоставляется требуемое соединение. При отказах или задержках соединений возникают потери. Они делятся на явные и условные.

Обслуживание с явными потерями означает, что при получении отказа в немедленном соединении сообщение

и соответствующий ему вызов полностью теряются и в дальнейшем на данную коммутационную систему не поступают. Обслуживание с задержкой сообщения сверх допустимого времени приводит к условным потерям. Оно организуется по системе с ожиданием соединения и с повторением вызова. В первом случае организуется очередь, которая обслуживается по мере освобождения соединительных устройств.

При обслуживании с повторением вызовов необслуженные источники образуют как бы активную очередь и повторяют вызовы до установления соединения.

В каждой конкретной системе характеристика качества обслуживания определяется спецификой решаемой задачи. Например, в системах с ожиданием качество обслуживания задают вероятностью ожидания и функцией распределения длительности ожидания. Пусть на полнодоступный пучок из 9 линий поступает простой пуассоновский поток с параметром λ . Если есть свободные линии, то при поступлении вызова занимает одна из них. При отсутствии свободных линий вызов становится в очередь. Обслуживание вызовов из очереди производится в порядке их поступления. Длительности занятий независимы друг от друга и от потока вызовов. Распределение длительности занятий подчиняется экспоненциальному закону: $H(x) = 1 - e^{-\mu x}$, где μ^{-1} — средняя длительность занятия. Число вызовов, поступающих за среднюю длительность занятия (поступающая нагрузка):

$$y = \lambda / \mu.$$

Решение задачи дает следующий результат. Вероятность ожидания

$$P\{\gamma > 0\} = \frac{E_{9,9}(y)}{1 - \frac{y}{9}(1 - E_{9,9}(y))}, \quad (8.21)$$

где γ — время ожидания; $E_{9,9}(y)$ — потери, вычисляемые по формуле Эрланга:

$$P_i = \frac{\lambda i / i!}{\sum_{k=1}^9 \lambda^k} = E_{i,9}(\lambda), i = 0, 1, \dots, 9. \quad (8.22)$$

Закон распределения времени ожидания определяется формулой

$$P(\gamma > t) = P\{\gamma > 0\} e^{-\mu(9-y)t}. \quad (8.23)$$

Здесь $P\{\gamma > t\}$ — вероятность того, что для поступившего в произвольный момент вызова время ожидания будет больше t .

Задачи анализа неполнодоступных схем являются более сложными.

В результате их решения определяются качественные показатели этих схем, включая и потери вызовов. Использование результатов анализа позволяет выбрать оптимальный вариант неполнодоступного включения приборов.

Кроме задач анализа, в теории распределения информации решаются также и задачи синтеза. К их числу относятся: а) задача синтеза структуры коммутационной системы квазиэлектронных и электронных АТС; б) задача оптимизации параметров коммутационной системы.

Схема коммутационного поля строится так, чтобы при заданной дисциплине обслуживания, заданных характеристиках потока сообщений и качестве обслуживания стоимость коммутационной системы была минимальной.

Следующий круг задач связан с нормированием и оптимальным распределением по участкам сети показателей качества обслуживания. От общей нормы качества обслуживания зависит стоимость оборудования коммутационных узлов и всей сети и время передачи сообщений через сеть связи. Очевидно, что необходимо стремиться к снижению как этой стоимости, так и времени передачи.

Наконец, важной и еще не в полной мере решенной задачей является разработка методов анализа, синтеза и оптимизации систем распределения информации в целом. Ее решение диктуется интеграцией систем передачи, систем распределения информации и систем управления, широким использованием в сетях связи ЭВМ.

Контрольные вопросы

1. Какие преимущества имеет многоканальная передача сообщений?
2. Какие требования предъявляются к канальным сигналам при формировании группового сигнала системы многоканальной передачи?
3. В чем сходство и различие методов частотного и временного разделения сигналов?
4. В чем заключается принцип фазового разделения сигналов?
5. Каким условиям должны удовлетворять шумоподобные сигналы при их разделении?
6. В чем заключается способ разделения сигналов в асинхронно-адресных системах связи?
7. Какие способы пространственного разделения сигналов Вам известны? На каких принципах они основаны?
8. Какими параметрами системы связи определяется пропускная способность многоканальных систем?
9. Что является содержанием теории распределения информации?

Глава 9. МЕТОДЫ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ СИСТЕМ СВЯЗИ

9.1. Оценка эффективности систем связи

Под *эффективностью* Q понимают некоторую функцию показателей качества, которая характеризует систему связи с технической точки зрения:

$$Q = f(k_1, k_2, \dots, k_{m-1}). \quad (9.1)$$

Показателями качества k_i , $i = 1, 2, \dots, m - 1$ называют такие параметры, которые удовлетворяют условию: чем больше (меньше) k_i тем лучше (хуже) система. Например, показателями качества являются: точность (верность) воспроизведения сообщения, скорость передачи сообщения, пропускная способность, помехозащищенность, надежность и т.д. Примерами параметров являются: длина рабочей волны, на которой работает радиосистема, ширина полосы канала связи и др. В отношении параметров нельзя утверждать, что чем он больше (меньше), тем лучше (хуже) система. Для более полного описания вводят m -й показатель системы — стоимость.

Функцию эффективности Q и стоимость C применяют для оценки оптимальности системы связи. Критерий оптимальности записывают в виде:

$$Q = f(k_1, k_2, \dots, k_{m-1}) = \max(\min); \quad (9.2)$$

$$C \leq C_m,$$

где C_m — максимально допустимое значение стоимости.

Задача отыскания оптимальной системы связи таким образом связана с отысканием функции эффективности Q , которая называется также *целевой функцией*. К этой функции предъявляются следующие требования.

1. Она должна наиболее полно характеризовать систему.
2. Функция должна быть простой для вычислений и удобной для практического использования.

Эти требования являются противоречивыми. Чтобы удовлетворить, например, первому требованию надо использовать большее число показателей качества, но при этом целевая функция становится сложной и неудобной для расчетов.

Часто целевую функцию берут равной весовой сумме показателей качества:

$$Q = \sum_{i=1}^n \lambda_i k_i, \quad (9.3)$$

где λ_i — постоянные весовые коэффициенты.

Выбор коэффициентов и количества показателей качества в определенной мере является произвольным и основывается в том числе на накопленном опыте проектирования и эксплуатации системы. Одной из методик выбора является метод экспертных оценок, при котором учитывается относительная важность показателей качества.

Часто оценку оптимальности системы ведут по одному (наиболее важному) показателю качества, переводя остальные в разряд ограничений. Условие оптимальности записывают в виде:

$$k_1 = \max(\min); \quad k_2 \leq k_{2m}, \dots, k_m \leq k_{mm}, \quad (9.4)$$

где k_1 — наиболее важный показатель; $k_{2m}, \dots, k_{mm}$ — максимально допустимые значения других показателей качества.

Применяют также поэтапную процедуру оптимизации. Вначале производят оптимизацию по одной группе параметров, а затем — по другой.

9.2. Принципы системного подхода к исследованию и разработке систем связи

Условия оптимальности, приведенные выше, используются при синтезе системы или ее отдельных частей. *Синтез системы* — это определение ее структуры или алгоритма функционирования, связанного с этой структурой. В процессе синтеза определяется состав системы, вид ее отдельных частей (например, передатчика, приемника), характер их взаимосвязи и взаимодействия, форма несущих колебаний, способ модуляции и демодуляции, вид кодирования.

Различают *математический* и *инженерный* синтез.

Математический синтез это отыскание чисто математическим путем структуры системы на основе математических выражений и критерия оптимальности. Из-за сложности решаемой задачи математический синтез больших систем, к которым относятся и системы связи, пока еще не представляется возможным. Поэтому его используют только для отыскания более мелких частей систем, таких, например, как демодулятор, линейный фильтр и т.д. Примеры математического синтеза приведены в гл. 7.

Под инженерным синтезом понимают отыскание структуры системы с использованием:

1) математических методов исследования, включающих синтез отдельных частей системы, расчеты на ЭВМ, математическое моделирование и автоматизированное проектирование;

2) экспериментальных методов исследования на макетах и образцах системы;

3) результатов лабораторных, заводских и других испытаний;

4) эвристических методов, использующих предыдущий опыт и знания по созданию и эксплуатации аналогичных систем, интуицию и творческие способности членов коллектива, участвующих в проектировании.

Название "*эвристический*" происходит от слова "эврика" — находить. При инженерном синтезе условия оптимальности системы используют не для отыскания (чисто математическим путем) ее структуры, а чаще всего для выбора структуры системы из некоторого числа заранее найденных структур. Такой выбор называется дискретным.

Если структура системы связи, полученная в результате проведения инженерного синтеза, удовлетворяет заданному критерию оптимальности, то такая система называется *оптимальной*.

В противном случае синтезированная структура является *неоптимальной*. Критерий оптимальности (кроме математических выражений) может содержать и описательную часть, в которой содержатся не поддающиеся математическому описанию условия и ограничения, накладываемые на систему.

В отличие от инженерного синтеза, структура устройства, полученная в результате математического синтеза, всегда является оптимальной. Однако оптимизация отдельных частей системы не гарантирует построение оптимальной системы в целом.

Инженерный синтез системы является составной частью системного подхода к проектированию.

Сущность системного подхода [24] заключается в целостном, всестороннем рассмотрении проектируемой системы с учетом:

1) этапов ее "жизненного цикла" (проектирование, производство, эксплуатация и утилизация);

2) истории и перспектив ее развития (эволюции);

3) взаимодействия с природой и обществом в целом;

4) взаимного влияния элементной базы и системотехники;

5) основных видов взаимодействия внутри системы (функционального, конструктивного, динамического, информационного, энергетического);

6) изменения исходных данных и даже решаемой задачи в процессе проектирования, производства и эксплуатации с целью получения более надежных результатов проектирования и обеспечения универсальности применения системы, чтобы при изъятии или добавлении некоторых блоков система была пригодной для решения новых задач;

7) выделения главных показателей качества, подлежащих улучшению в первую очередь;

8) сочетания принципов композиции, декомпозиции и иерархичности;

9) вскрытия основных технических противоречий, препятствующих улучшению качества системы и ускорению процесса ее разработки;

10) правильного сочетания различных методов проектирования, в первую очередь математических, эвристических и экспериментальных, а в рамках математических методов — аналитических и проводимых с помощью ЭВМ;

11) обеспечения должного взаимодействия в процессе проектирования специалистов различных уровней и профилей.

9.3. Анализ систем связи по их эффективности

9.3.1. Энергетическая и частотная эффективность систем связи

Предварительный анализ систем можно вести по небольшому числу показателей качества. Обычно в их качестве берут *скорость передачи* R (бит/с) и *верность передачи*, определяемую вероятностью ошибки p при передаче дискретных сообщений или отношением сигнал/шум на выходе демодулятора при передаче непрерывных сообщений. Полагают, что сигналы передаются в канале с аддитивным гауссовским белым шумом. Полосу канала задают равной F , отношение сигнал/шум — $\rho_0 = P_c/N_0$, где P_c — мощность сигнала, N_0 — спектральная плотность мощности шума.

Разделим скорость передачи на пропускную способность канала и таким образом вычислим коэффициент использования пропускной способности:

$$\eta = R/C. \quad (9.5)$$

С учетом формулы Шеннона перепишем выражение (9.5) в виде:

$$\eta = \frac{R}{F \log(1 + \rho)}, \quad (9.6)$$

где $\rho = P_c/P_{\text{ш}}$ — отношение мощности сигнала и шума в полосе частот F .

Введем коэффициенты использования канала по мощности β и по полосе частот γ . Эти коэффициенты определяют соответственно эффективность использования энергии сигнала (энергетическая эффективность) и эффективность использования полосы канала (частотная эффективность):

$$\beta = R/\rho_0, \quad \gamma = R/F. \quad (9.7)$$

Подставляя выражения (9.7) в (9.6), получим

$$\eta = \gamma / [\log(\gamma / \beta + 1)], \quad \gamma = \rho\beta. \quad (9.8)$$

Определим максимальное значение полученной целевой функции $Q(\beta, \gamma) = \eta(\beta, \gamma)$.

В идеальной системе связи η сколь угодно близка к единице при сколь угодно малой вероятности ошибки. Подставляя $\eta = 1$ в (9.8), получим зависимость между β и γ для идеальной системы связи:

$$\gamma = \log(\gamma / \beta + 1) \quad \text{или} \quad \beta = \gamma / (2^\gamma - 1). \quad (9.9)$$

Зависимость $\beta = f(\gamma)$ изобразим кривой на плоскости с осями $x = \gamma$, $y = \beta$ (рис.9.1). Частотная эффективность изменяется в пределах от 0 до ∞ , а энергетическая эффективность от $-\infty$ до $1/\ln 2$.

Максимальное значение энергетической эффективности найдем из (9.9):

$$\beta_{\max} = \lim_{\gamma \rightarrow 0} \lim_{\gamma \rightarrow 0} \frac{\gamma}{2^\gamma - 1} = \frac{1}{\ln 2} \approx 1,443(1,59 \text{ дБ}).$$

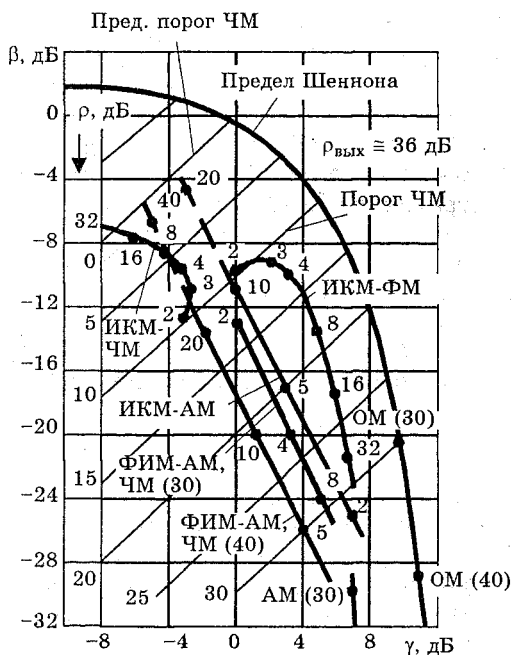


Рис.9.1. Кривые энергетической и частотной эффективности аналоговых и цифровых систем связи

Найденная кривая определяет зависимость между энергетической и частотной эффективностями идеальной системы связи, в которой $\eta = 1$, а вероятность ошибки воспроизведения сообщения сколь угодно мала.

Для сравнительной оценки эффективности исследуемой системы рассчитывают аналогичную кривую $\beta = f(\gamma)$, подставляя в формулы (9.7) вместо скорости R выражение для пропускной способности канала этой системы. Каждой точке полученной кривой будет соответствовать вариант исследуемой системы.

В реальной системе связи коэффициент использования канала $\eta < 1$, а вероятность

ошибки p имеет конечное значение. Поэтому следует ожидать, что полученная кривая для реальной системы будет располагаться ниже кривой, соответствующей идеальной системе.

9.3.2. Выбор способов модуляции и помехоустойчивого кодирования

Используя методику, рассмотренную выше, можно построить кривые зависимостей для различных способов приема, модуляции и кодирования. На рис.9.1, 9.2 приведены кривые, которые позволяют провести сравнительный анализ для: 1) аналоговых систем; 2) цифровых систем; 3) систем, в которых используются многопозиционные сигналы; 4) систем, в которых используются корректирующие коды.

Информационная эффективность непрерывных систем

$$\eta = \frac{R}{C} = \frac{H'_\epsilon}{C} = \frac{F_c \log \rho_{\text{вых}}}{F \log(1 + \rho_{\text{вых}}/g)} = \frac{\log \rho_{\text{вых}}}{\alpha \log(1 + \rho_{\text{вых}}/g)}, \quad (9.10)$$

где H'_ϵ — энтальпия-производительности непрерывного источника; F_c — ширина спектра сообщения; F — ширина полосы частот, занимаемой сигналом; $\alpha = F/F_c$; $\rho_{\text{вых}}$ — отношение мощностей сигнала и шума на выходе приемника; $g = \rho_{\text{вых}}/\rho$; ρ — отношение мощностей сигнала и шума на входе приемника в полосе F .

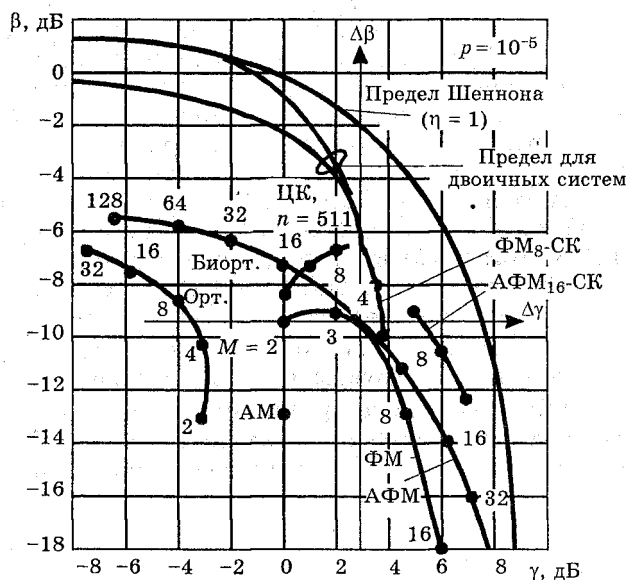


Рис.9.2. Кривые энергетической и частотной эффективности систем с многопозиционными сигналами и корректирующими кодами

По формуле (9.10) рассчитаны значения выигрыша g и информационной эффективности η для различных аналоговых и цифровых систем связи. Расчетные данные для $\rho_{\text{вых}} = 40$ дБ приведены в табл.9.1 [3]. Расчеты выполнены для оптимального приема сигналов при условии передачи по каналу связи сообщений с наивысшей частотой F_c и пик-фактором $\Pi = 3$. Все параметры сигнала, кроме модулируемого, полностью известны.

Таблица 9.1

N п/п	Система модуляции	$\alpha = F/F_c$	$g = \rho_{\text{вых}}/\rho$	$\rho' = \rho/\alpha$	$\eta = R/C$
1	АМ	2	0,2	0,1	0,42
2	БМ (АМ-ПН)	2	2	1	0,50
3	ОМ (АМ-ОБП)	1	1	1	1
4	ФМ	20	222	11,1	0,12
5	ЧМ	20	666	33,3	0,17
6	ФИМ-АМ	20	666	33,3	0,17
7	ИКМ-АМ	20	250	12,5	0,23
8	ИКМ-ЧМ	20	500	25	0,32
9	ИКМ-ФМ	20	1000	50	0,47
10	Идеальная система	20	6310	315	1

В системе ФИМ-АМ на первой ступени модуляции используется фазово-импульсная модуляция, а на второй — амплитудная модуляция (АМ) перемножением ФИМ сигнала на непрерывный косинусоидальный сигнал. ИКМ-АМ сигнал получают перемножением соответствующего двоичного кода (с выхода АЦП) на непрерывный косинусоидальный сигнал. При ИКМ-ЧМ и ИКМ-ФМ модуляции двоичная кодовая последовательность с выхода АЦП управляет или несущей частотой или начальной фазой непрерывного косинусоидального колебания. Число возможных значений частот и фаз выбрано равным четырем.

Наибольшая информационная эффективность среди сравниваемых непрерывных систем достигается в системе с АМ-ОБП и наименьшая — в системе с ФМ. Но системы с АМ, АМ-ПН и АМ-ОБП имеют низкую помехоустойчивость, определяемую небольшими значениями выигрыша, равными соответственно $g = 0,2$; 2; 1. Это означает, что верность передачи в этих системах может быть повышена только за счет увеличения мощности передаваемого сигнала.

Кривые $\beta = f(\gamma)$ для аналоговых и цифровых систем рассчитаны по формулам:

$$\beta = \frac{g \log \rho_{\text{вых}}}{\alpha \rho_{\text{вых}}}; \quad \gamma = \frac{\log \rho_{\text{вых}}}{\alpha}. \quad (9.11)$$

Кривые построены в координатах β , γ [3]. Кроме прямоугольной сетки, на рис.9.1. имеются также прямые линии под углом в 45° , соответствующие входному отношению сигнал/шум в децибеллах. Цифры на кривых, построенных для аналоговых систем, указывают значения α , а цифры в скобках — значения $R_{\text{вых}}$ в децибеллах. В пороговой области кривые для ЧМ показаны штриховыми линиями. Расчеты для цифровых систем выполнены при условии отсутствия межсимвольных помех и минимально необходимой ширины полосы канала F . Цифры на кривых указывают основание кода. Кривые построены для ИКМ-АМ, ИКМ-ФМ, ИКМ-ЧМ. На том же рисунке построены кривые и для систем с ФИМ.

Анализ кривых показывает, что реальные системы проигрывают идеальной по Шеннону системе связи.

Кривые позволяют оценить изменение энергетической и частотной эффективностей в зависимости от α для аналоговых систем и в зависимости от основания кода для цифровых. Например, в системе с ФИМ-АМ с увеличением α энергетическая эффективность увеличивается, а частотная падает.

В цифровых системах (например с ИКМ-ФМ) с увеличением основания кода энергетическая эффективность падает, а частотная увеличивается.

Кривые для систем, в которых применяются многопозиционные сигналы: ортогональные (ОС), биортогональные (БС) и фазоманипулированные (ФМ), приведены на рис.9.2.

Количество сигналов равно M , вероятность ошибки во всех системах $p = 10^{-5}$. Цифры рядом с точками указывают значение M [3].

В качестве ортогональных сигналов выбраны дискретные ЧМ сигналы. Анализ кривых показывает, что в системах с ЧМ с ростом M энергетическая эффективность β увеличивается, а в системах с ФМ — уменьшается. Частотная эффективность в системах с ЧМ с ростом M уменьшается, а в системах с ФМ — увеличивается.

Пусть заданы скорость передачи R^* полоса частот F^* и отношение сигнал/шум ρ . Системы, для которых $\beta > \beta^*$, $\gamma > \gamma^*$ обладают высокой энергетической и частотной эффективностью. К числу этих систем относятся системы с ФМ сигналами ($M = 2; 4$) и системы с биортогональными сигналами ($M = 4, 8$ и 16). Соответствующие этим системам точки расположены в 1-м квадранте.

Системы с биортогональными сигналами ($M = 32, 64, 128$) и системы с ЧМ сигналами ($M = 4, 8, 16, 32$) имеют высокую энергетическую эффективность и не удовлетворяют требованию по

частотной эффективности. Этим системам соответствуют точки, расположенные во 2-м квадранте.

Системы с ФМ сигналами ($M = 8, 16$) имеют высокую частотную и низкую энергетическую эффективность. Этим системам соответствуют точки, расположенные в 4-м квадранте.

Наконец из анализа рис.9.1 следует, что системы связи с ЧМ ($M = 2$) имеют низкие показатели по γ и β .

Для сравнительного анализа систем начало координат можно перенести в точку, соответствующую цифровой системе с фазовой модуляцией и $M = 4$. В новой системе координат $\Delta\beta$, $\Delta\gamma$ сравнение эффективности систем будет производиться по отношению к эффективности цифровой системы с фазовой модуляцией и $M = 4$. В спутниковых системах фазовая модуляция с $M = 4$ наиболее распространена и принята в качестве стандарта. В новой системе координат высокоэффективными системами по показателям $\Delta\beta$, $\Delta\gamma$ являются цифровые системы с фазовой модуляцией сигнала и $M = 8$ и КАМ с $M = 16$.

Применение циклического кода позволяет повысить энергетическую эффективность в обмен на снижение частотной на 2...3 дБ. Частотная эффективность при этом падает примерно на 3 дБ по сравнению с характеристиками системы связи с ФМ и $M = 4$. Из рис.9.2 следует, что циклическое кодирование в канале с ФМ ($M = 8$) позволяет получить одновременно выигрыш как по энергетической, так и частотной эффективностям. Реализация подобных высокоэффективных систем приводит к увеличению их сложности. Однако, если затраты на усложнение конструкции кодирующих и декодирующих устройств меньше затрат на увеличение мощности сигналов в канале связи, то это усложнение окажется оправданным. Поэтому совершенствование элементной базы и технологии производства кодирующих и декодирующих устройств с целью их удешевления открывает перспективу для их широкого внедрения. Таким образом, построение и анализ кривых $\beta = f(\gamma)$ позволяет из ряда вариантов всевозможных систем отбросить заведомо негодные варианты.

Если, например, необходимо решить задачу оптимизации системы проводной связи, то нужно отбросить все варианты, находящиеся в первом втором и третьем квадрантах (рис.9.2). Это объясняется тем, что в системах проводной связи γ — эффективность является важнейшим качественным показателем.

При решении задач оптимизации космических и спутниковых систем связи из ряда систем следует выделить системы с высокой энергетической эффективностью. Они находятся во втором квадранте.

Если же нужно обеспечить компромиссное решение, то выбор системы с достаточно высокими показателями β и γ следует вести из числа систем, расположенных в первом квадранте.

Однако выбор системы по показателям эффективности не может быть окончательным, так как при этом не учитываются другие важные показатели качества, например, стоимость.

9.3.3. Совместная оптимизация модуляции и кодирования

Высокие показатели качества можно обеспечить при комплексном подходе к проектированию модулятора, кодирующего устройства и демодулятора, декодирующего устройства с учетом условий и ограничений, накладываемых на вид модуляции и кодирования, структуру и интенсивность помех, вид канала связи. При приеме в целом, хотя и обеспечивается наибольшая верность принятого сообщения, но оптимальный приемник из-за большого числа каналов очень сложен, следовательно имеет большую стоимость. Поэтому, как правило, используется поэлементный прием с последующим декодированием принятой кодовой комбинации. Некоторое ухудшение в верности принятого сообщения здесь компенсируется существенным упрощением приемника.

Очевидно, что и при поэлементном приеме стремятся обеспечить максимально возможную верность принятого сообщения. Эта цель достигается: 1) при наиболее полном использовании информации, содержащейся в апостериорных плотностях распределения вероятности, полученных после приема каждого элемента кодовой комбинации; 2) при согласовании канала связи со структурой используемого кода.

Уменьшить потери информации при обработке сигнала можно различными способами, в том числе за счет более позднего принятия решения. Такой вид решения называется "мягким". В отличие от "жесткого" решения, когда напряжения с выходов согласованных фильтров (или корреляторов) сравниваются с пороговым уровнем и на основании этого сравнения принимается окончательное решение о принимаемых символах, при "мягком" решении напряжения с выходов согласованных фильтров (или корреляторов) хранятся в запоминающем устройстве и используются при декодировании принимаемой кодовой комбинации.

Примеры "жесткого" решения рассмотрены в гл.7. Прием сигнала с использованием вариантов приема сигнала с "мягким" решением рассматривается в специальных дисциплинах. [12].

Повышение верности принятого сообщения достигается также согласованием кодирующего и декодирующего (кодека) устройства с каналом связи.

Например, дискретный симметричный канал (ДСК) с независимыми ошибками оказывается согласованным с различными корректирующими кодами, предназначенными для исправления этих независимых ошибок. Однако при появлении в линии связи медленных замираний происходит рассогласование кодека и дискретного канала из-за изменений его характеристик. ДСК трансформируется в дискретный канал с памятью.

При сильном сигнале характеристики дискретного канала с памятью почти не отличаются от характеристик ДСК, а корректирующий код успешно справляется с возникающими при приеме кодовой комбинации ошибками. При слабом сигнале характеристики канала отличаются от ДСК, а вероятность возникновения ошибок существенно возрастает. Корректирующий код в этом случае с ошибками не справляется. Прием сообщения становится невозможным.

Разработаны два способа согласования кодека с каналом. Первый из них связан с подбором кода. Если ошибки в каналах с памятью располагаются в пределах относительно небольшого числа смежных разрядов принимаемой кодовой комбинации, то для их обнаружения и исправления используются специально разработанные для этих целей коды для исправления "пачек" ошибок. Второй способ связан с преобразованием исходных каналов к "стандартному" дискретному каналу. Эти преобразования происходят в модуляторе.

Примером такого преобразования является декорреляция ошибок. При этом способе символы, входящие в кодовую комбинацию, передаются не друг за другом, а перемежаются символами других кодовых комбинаций. Интервал между символами, входящими в одну комбинацию, выбирается больше интервала корреляции ("памяти") канала. В этом случае ошибки, возникающие в символах, становятся независимыми, и их можно исправлять также, как и в канале без памяти.

Для борьбы с замираниями применяют сигналы сложной структуры: многочастотные, частотно-временные, шумоподобные.

Рассмотрим пример применения частотно-временного сигнала (ЧВС). ЧВС формируется следующим образом. Каждой двоичной комбинации первичного кода ставится в однозначное соответствие комбинация импульсов, имеющих разные частоты (рис.9.3). Число комбинаций импульсов выбирается равным числу комбинаций первичного кода. Максимальное число комбинаций определяется их различием в частотах и временным положением на оси времени. Достаточно простыми средствами можно получить несколько тысяч комбинаций. Каждая комбинация им-

пульсов должна иметь автокорреляционную функцию с малыми боковыми лепестками, а взаимно-корреляционные функции между всеми комбинациями импульсов также должны иметь небольшие остатки. Это позволяет разделять ЧВС согласованными фильтрами и корреляторами. Максимальное число комбинаций импульсов с уровнем боковых лепестков корреляционных функций не превышающих $1/\sqrt{FT}$ приближается к числу FT , где F — ширина спектра сигнала; T — его длительность.

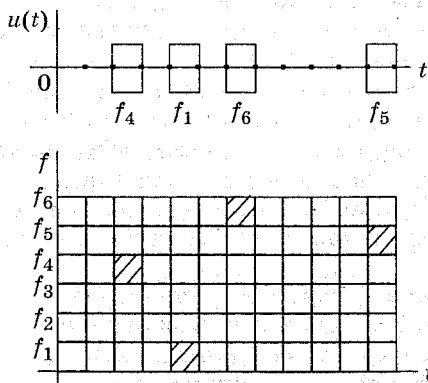


Рис.9.3. Частотно-временной сигнал

Выбирая комбинации импульсов, можно перераспределять память в дискретном канале, уменьшить корреляцию по времени между символами передаваемой кодовой комбинации и тем самым ослабить влияние замираний. ЧВС можно использовать и для борьбы с сосредоточенными и импульсными помехами.

В случае применения ЧВС в ААСС каждому абоненту сети присваивается двоичная комбинация, следовательно, и комбинация импульсов ЧВС. Передача "1" информационно-

го кодового слова в направлении этого абонента производится в течении заданного времени посылкой соответствующей этому абоненту комбинации импульсов ЧВС. При передаче "0" ЧВС не передается в течение того же времени.

9.3.4. Использование обратного канала для повышения эффективности передачи дискретных сообщений

Верность передаваемого сообщения можно существенно повысить, если в состав системы связи ввести дополнительный канал и по нему передавать (ретранслировать) кодовые комбинации или специальные сигналы. При "дуплексной связи" для передачи дополнительной информации можно использовать имеющиеся прямой и обратный каналы. Если система с ретрансляцией на передающее устройство принятых кодовых комбинаций использует обратный канал полностью (при ретрансляции полезная информация по обратному каналу не передается), то такая система называется *системой с полной информационной обратной связью*.

Если по обратному каналу передаются, наряду с полезной информацией, только некоторые признаки принятых информационных кодовых комбинаций, то такая система называется *системой с укороченной информационной обратной связью*. Передатчик на основе принятой по обратному каналу информации принимает решение о стирании неправильно принятой приемником информации, посылает на приемник сигнал стирания, а затем повторяет нужную комбинацию.

В системах с информационной обратной связью сведения о поступающих на приемник кодовых комбинациях передаются на передатчик до их окончательной обработки (декодирования) и принятия какого-либо заключительного решения.

В другой разновидности систем — *системах с решающей обратной связью* приемник производит прием кодовой комбинации, анализирует ее на наличие ошибок и принимает окончательное решение о выдаче комбинации потребителю информации или ее стирании. В последнем случае по обратному каналу на передатчик передается сигнал о повторной передаче этой кодовой комбинации.

Система, в которой повторение кодовых комбинаций происходит при обнаружении в них ошибок, является разновидностью систем с *адаптивным кодированием и декодированием*. При адаптивном кодировании в соответствии с состоянием канала связи или меняется структура кода и связанное с этой структурой соотношение между числом информационных и проверочных символов или производится повторение кодовых комбинаций.

Структурная схема системы с решающей обратной связью приведена на рис.9.4.

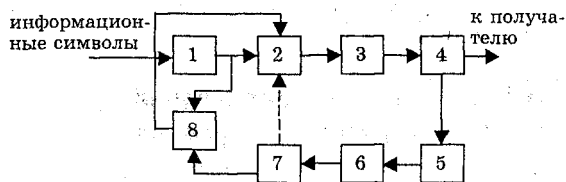


Рис.9.4. Структурная схема системы связи с решающей обратной связью: 1 — кодер; 2 — блок управления; 3 — прямой канал; 4 — декодер; 5 — датчик сигнала переспроса; 6 — обратный канал; 7 — приемник сигнала переспроса; 8 — накопитель-повторитель

Кодовая комбинация с выхода кодера поступает в накопитель-повторитель и прямой канал. В приемнике (декодере) производится обнаружение ошибок. Если ошибок нет, принятая кодовая комбинация декодируется, а информационные символы поступают для дальнейшей обработки.

При обнаружении ошибок датчик сигнала переспроса вырабатывает сигнал, который через обратный канал поступает на

приемник сигнала переспроса. Сигнал управления с выхода этого приемника блокирует выход кодера и соединяет на необходимое время выход накопителя-повторителя со входом прямого канала. По прямому каналу передается (повторяется) ранее переданная кодовая комбинация.

Система с решающей обратной связью оказывается достаточно эффективной в каналах с замираниями. С ростом вероятности ошибки число переспросов увеличивается, а скорость передачи информации уменьшается, и наоборот. При этом верность передачи остается удовлетворительной, если вероятность ошибки не становится равной $1/2$. Система связи как бы приспособляется (адаптируется) к состоянию канала. Рассогласования кодека с каналом связи в этом случае не происходит и необходимость в дополнительном согласовании этих устройств отпадает.

9.3.5. Устранение избыточности

Устранение избыточности реальных источников сообщений в ряде случаев диктуется необходимостью повышения эффективности систем связи. Эта задача решается в рамках системного подхода к проектированию с помощью инженерного синтеза структуры системы. Поэтому в ряде систем избыточность устраняется, а в ряде других систем — нет. Например, в системах телеграфной связи методы устранения избыточности с применением эффективного кодирования не нашли применения из-за их относительной сложности и высокой стоимости.

В то же время эффективное кодирование оказывается экономически и технически целесообразным в системах передачи данных. Не вызывает сомнений и устранение избыточности при передаче непрерывных сообщений (речи, телевизионных изображений). Системы синтетической телефонии позволяют сократить полосу частот примерно в 10 раз при достаточно хорошей разборчивости речи [3]. Принцип устранения избыточности заключается в кодированной передаче значений амплитуд составляющих речевого сигнала в отдельных полосах частот, определяющих разборчивость речи. Другой принцип устранения избыточности связан с передачей в кодированной форме первых коэффициентов ряда Фурье спектральной плотности речевого сигнала.

В приемнике производится декодирование принятых кодовых комбинаций и синтез речи, близкой к передаваемой, но с потерей естественности звучания.

Устройства, реализующие эти принципы устранения избыточности речевого сигнала, называются соответственно *формантным* и *гармоническим вокодерами*. Решить задачу устранения избыточности при цифровом методе передачи речи позволяет также использование методов кодирования с предсказанием и методов

декорреляции. Методы кодирования с предсказанием рассмотрены в гл.7. В системе ДИКМ при передаче речи объем передаваемых данных можно уменьшить в 2-3 раза при достаточном качестве воспроизведения речи. Методы декорреляции основаны на аппроксимации непрерывных сигналов с помощью различных базисных функций. В качестве таких функций используются степенные полиномы нулевого и первого порядка. Применяется также кусочно-полиномиальная интерполяция на основе сплайн-функций. Более подробно с этими методами можно познакомиться в специальной литературе [23].

Другая возможность повышения эффективности телефонной связи реализуется при использовании пауз, суммарная продолжительность которых сравнима со временем звучания. Известен ряд совмещенных систем, которые наряду с передачей речи в паузах осуществляют передачу и дискретной информации. Это позволяет в 2-3 раза повысить эффективность использования каналов тональной частоты.

Устранение избыточности источников при цифровой передаче непрерывных сообщений как правило сопровождается согласованием кодека источника с цифровым каналом.

Контрольные вопросы

1. Что называют эффективностью системы связи и как она определяется количественно?
2. Сформулируйте задачу отыскания оптимальной системы связи. Как записываются условия оптимальности?
3. Что Вы понимаете под синтезом системы? Что такое математический и инженерный синтез системы?
4. В чем сущность системного подхода к исследованию и разработке систем связи?
5. Какие математические выражения определяют энергетическую и частотную эффективность систем связи? Как эти характеристики связаны между собой в идеальной системе связи?
6. Чем отличаются характеристики эффективности реальных систем связи от идеальной по Шеннону системы связи?
7. Какими методами можно осуществить обмен показателей эффективности β на γ и γ на β в системе с ИКМ-АМ и в системе ИКМ-ФМ?
8. Почему код, оптимальный для ДСК без памяти, не является оптимальным для канала с медленными замираниями?
9. Можно ли преобразовать канал с замираниями в канал, близкий к ДСК? Если такое преобразование возможно, то в каком из блоков системы оно осуществляется?
10. Какие методы устранения избыточности используются в телефонии?

ЛИТЕРАТУРА

1. Котельников В.А. Теория потенциальной помехоустойчивости. М.: Госэнергоиздат, 1956.

2. Шеннон К. Математическая теория связи /Работы по теории информации и кибернетике: Пер. с англ. /Под ред. Р.Л.Добрушина и О.В.Лупанова. — М.: ИЛ, 1963.

3. Теория передачи сигналов. Учебник для вузов /А.Г.Зюко, Д.Д.Кловский, М.В.Захаров, Л.М.Финк. — 2-е изд., перераб. и доп. — М.: Радио и связь, 1986.

4. Кловский Д.Д., Шилкин В.А. Теория электрической связи. Сб. задач и упражнений: Учеб. пособие для вузов. — М.: Радио и связь, 1990.

5. Баскаков С.И. Радиотехнические цепи и сигналы. — М.: Высшая школа, 1988.

6. Радиотехнические системы передачи информации: Учебн. пособие для вузов /В.А.Борисов, В.В.Калмыков, Я.М.Ковальчук и др.; Под ред. В.В.Калмыкова. — М.: Радио и связь, 1990.

7. Дмитриев В.Н. Прикладная теория информации. — М.: Высшая школа, 1989.

8. Лахти Б.П. Системы передачи информации. — М.: Связь, 1971.

9. Системы электросвязи: Учебник для вузов /В.П.Шувалов, Г.П.Катунин и др.; Под ред. В.П.Шувалова. — М.: Радио и связь, 1987.

10. Гауэр Дж. Оптические системы связи. М.: Радио и связь, 1989.

11. Алишев Я.В. Многоканальные системы связи оптического диапазона. Минск, Высшая школа, 1986.

12. Передача дискретных сообщений. Учебник для вузов /В.П.Шувалов, Н.В.Захарченко и др.; Под ред. В.П.Шувалова. М.: Радио и связь, 1990.

13. Зайдлер Е. Системы передачи дискретной информации. — М.: Связь, 1977.

14. Кушнир В.Ф., Ферсман Б.А. Теория нелинейных электрических цепей. — М.: Связь, 1974.

15. Баркун М.А. Цифровые автоматические телефонные станции. — Минск, Высшая школа, 1990.

16. Корнышев Ю.Н., Фань Г.Л. Теория распределения информации. — М.: Радио и связь, 1985.

17. Теория телефонных и телеграфных сообщений. Учебник для вузов /Б.С.Лившиц, Я.В.Фидлин, А.Д.Харкевич. — М.: Связь, 1971.

18. Основы теории информационных сетей. Учебник для вузов /В.К.Морозов, А.В.Долганов. М.: Высшая школа, 1987.

19. Назаров М.В., Кувшинов Б.И., Попов О.В. Теория передачи сигналов. — М.: Связь, 1970.

20. Зюко А.Г. Элементы теории передачи информации. Киев. Техніка, 1969.

21. Апорович А.Ф. Проектирование радиотехнических систем. — Вышэйшая школа, 1988.

22. Системы подвижной связи /И.М.Пышкин, И.И.Дежурный, В.Н.Талызин, Г.Д.Чвилев; Под ред. И.М.Пышкина. — М.: Радио и связь, 1986.

23. Помехоустойчивость и эффективность систем передачи информации /А.Г.Зюко, А.И.Фалько, И.П.Панфилов и др.; Под ред. А.Г.Зюко. — М.: Радио и связь, 1985.

24. Гуткин Л.С. Проектирование радиосистем и радиоустройств. — М.: Радио и связь, 1986.

25. Тихонов В.И., Кульман Н.К. Нелинейная фильтрация и квазикогерентный прием сигналов, — М.: Сов. радио, 1975.

26. Чердынцев В.А. Статистическая теория совмещенных радиотехнических систем. — Минск.: Вышэйшая школа, 1980.

27. Тихонов В.Н. Статистическая радиотехника. — М.: Радио и связь, 1982.

28. Левин Б.Р. Теоретические основы статистической радиотехники. Книга 1. — М.: Сов. радио, 1974; книга 2., 1975; книга 3, 1976.

29. Гуткин Л.С. Теория оптимальных методов радиоприема при флуктуационных помехах. — М.: Сов. радио, 1972.

30. Залманзон Л.А. Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях. — М.: Наука, 1989.

31. Большаков М.А., Репин В.Г. Вопросы нелинейной фильтрации. Автоматика и телемеханика, 1961, №4.

32. Стратонович Р.Л. К теории оптимальной нелинейной фильтрации случайных функций. Теория вероятностей и ее применения, 1959, №2.

33. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки: Пер. с англ. /Под ред. Р.Л.Добрушина и С.Н.Самойленко. М.: Мир, 1976.

34. Кларк Дж, мл. Кейн Дж. Кодирование с исправлением ошибок в системах цифровой связи: Пер. с англ. — М.: Радио и связь, 1987.

35. Многоканальная связь на железнодорожном транспорте /В.Л.Тюрин, Д.В.Дьяков и др.; под ред. В.Л.Тюрина. — М.: Транспорт, 1992,

36. Красная книга, том VI, вып. VI.1. Общие рекомендации по телефонной коммутации и сигнализации. Стыки с морскими и сухопутными подвижными службами. Рекомендации Q.1-Q.118 bis. Международный союз электросвязи. Международный консультативный комитет по телефонии и телеграфии. VIII пленарная ассамблея. Малага - Торремолинос, 8 - 19 октября 1984.

37. Теория передачи информации. Терминология. Сб. рекомендуемых терминов, вып.101, — М.: Наука, 1984.

38. Беларуская ССР. Кароткая энцыклапедыя, т.3.

39. БРЭТОН Ф., ПРУ С. Выбух камунікацыі: пераклад з французскай. — Мінск.: Беларускі фонд Сораса, 1995.

40. Чарри К. Человек и информация: Пер. с англ. — М.: Связь, 1972.

41. Шинаков Ю.С., Колодяжный Ю.М. Теория передачи сигналов электросвязи. — М.: Радио и связь, 1988.

42. Давыдов Г.Б., Рогинский В.Н., Толчан А.Я. Сети электросвязи. — М.: Связь, 1977.

43. Толковый словарь по вычислительным системам. /Под ред. В.Иллингхурта, Э.Л.Глейзера, И.К.Пайла. Пер. с англ. под ред. Е.К.Масловского. — М.: Машиностроение, 1989.

44. Термины и определения в курсах МСП, МЭС и РРЛ, МЭС (методическая разработка). — Ленинград, ЛЭИС им. проф. М.А.Бонч-Бруевича, 1985.

СОДЕРЖАНИЕ

Предисловие	3
Введение. Краткая история развития теории и техники электрической связи	4
Глава 1. Основные понятия теории связи	6
1.1. Общие сведения о системах связи. Информация, сообщение, сигнал	6
1.2. Связь, сеть связи, система связи	8
1.3. Кодирование и модуляция	12
1.4. Демодуляция и декодирование	15
1.5. Основные характеристики систем связи	17
Контрольные вопросы	18
Глава 2. Детерминированные сигналы и их математические модели	19
2.1. Разложение детерминированных функций в ортогональные ряды	19
2.2. Спектральная плотность сигнала	23
2.3. Спектральная плотность мощности детерминированного сигнала	27
2.4. Автокорреляционная функция	28
2.5. Комплексное и квазигармоническое представление узкополосных сигналов	31
2.6. Геометрические представления сигналов	33
2.7. Теорема Котельникова	36
Контрольные вопросы	40
Глава 3. Методы формирования и преобразования детерминированных сигналов в системах связи	41
3.1. Определение линейной, нелинейной и параметрической систем	41
3.2. Преобразования спектров сигналов в параметрических и нелинейных цепях	42
3.2.1. Линейное преобразование дискретного спектра	42
3.2.2. Линейное преобразование непрерывной спектральной плотности	44
3.2.3. Преобразование дискретного спектра параметрической цепью	45
3.2.4. Преобразование параметрической цепью спектральной плотности	46
3.2.5. Безынерционное преобразование дискретного спектра	48
3.2.6. Безынерционное преобразование спектральной плотности	51
3.3. Преобразование частоты. Модуляция и детектирование	53
3.3.1. Преобразование частоты	53
3.3.2. Модуляция	56
3.3.3. Детектирование	59

3.4. Амплитудная модуляция гармонического переносчика	60
3.4.1. Модулирующие сигналы	60
3.4.2. Амплитудная непрерывная модуляция с подавленной несущей	61
3.4.3. Амплитудная непрерывная модуляция с большим уровнем несущей	63
3.4.4. Дискретная двоичная амплитудная модуляция	65
3.4.5. Амплитудная модуляция с одной боковой полосой	67
3.4.6. Векторное представление непрерывных АМ сигналов	69
3.5. Частотная непрерывная модуляция гармонического переносчика	70
3.5.1. Непрерывный частотно-модулированный сигнал с малым индексом модуляции	70
3.5.2. Непрерывный ЧМ сигнал с большим индексом модуляции	72
3.5.3. Однополосный непрерывный сигнал с угловой модуляцией	74
3.5.4. Методы формирования непрерывного ЧМ сигнала	75
3.5.5. Дискретная частотная модуляция	77
3.6. Фазовая модуляция гармонического переносчика	79
3.6.1. Непрерывная фазовая модуляция	79
3.6.2. Дискретная двоичная фазовая модуляция	79
3.6.3. Многократная фазовая модуляция	81
3.6.4. Квадратурная амплитудная модуляция	83
3.6.5. Дискретная фазовая модуляция с одной боковой полосой	84
3.7. Дискретная модуляция шумоподобного сигнала	84
3.7.1. Дискретная модуляция постоянного тока	84
3.7.2. Генерирование шумоподобных сигналов	85
3.7.3. Модуляция шумоподобного сигнала по форме	87
3.8. Модуляция импульсного переносчика	91
3.8.1. Амплитудно-импульсная модуляция	91
3.8.2. Широтно-импульсная модуляция. фазоимпульсная модуляция	93
3.8.3. Импульсно-кодовая модуляция	94
3.9. Некогерентное детектирование	96
3.9.1. Некогерентное детектирование непрерывных АМ сигналов	96
3.9.2. Некогерентное детектирование дискретных сигналов	99
3.9.3. Детектирование непрерывных ЧМ сигналов	100
3.9.4. Детектирование сигналов с амплитудно-импульсной модуляцией	102
3.9.5. Декодирование сигналов с импульсно-кодовой модуляцией	102
3.9.6. Детектирование сигналов с широтно-импульсной и фазо-импульсной модуляцией	103
3.10. Когерентное детектирование	103
3.10.1. Математическая модель когерентного детектора	103
3.10.2. Когерентное детектирование АМ-ОВП сигнала	105
3.10.3. Когерентное детектирование двоичных ФМ сигналов	108
3.10.4. Когерентное детектирование сигналов с ОФМ	109

3.10.5. Детектирование шумоподобных сигналов	112
3.10.6. Особенности реализации операции умножения в устройствах формирования и преобразования сигналов.....	114
3.11. Принципы цифровой обработки сигналов.....	116
3.11.1. Дискретные преобразования Фурье (прямое и обратное)	116
3.11.2. Квантование	120
3.11.3. Прямое и обратное z -преобразование	121
3.11.4. Трансверсальные цифровые фильтры	123
3.11.5. Рекурсивный цифровой фильтр	124
3.11.6. Пример синтеза линейных цифровых фильтров.....	124
3.11.7. Цифровые модуляторы и детекторы	125
Контрольные вопросы	127
Глава 4. Случайные сигналы и их математические модели	128
4.1. Математические модели непрерывных случайных процессов	128
4.2. Математические модели случайных дискретных процессов	133
4.3. Функция корреляции и энергетический спектр.....	134
4.4. Нормальный случайный процесс	137
4.5. Узкополосные случайные процессы.....	140
4.6. Геометрические представления случайных сигналов и помех	143
4.7. Математические модели сообщений	145
Контрольные вопросы	149
Глава 5. Каналы связи	150
5.1. Общие сведения о каналах передачи информации	150
5.1.1. Классификация каналов	150
5.1.2. Краткие сведения о искажениях сигналов.....	150
5.1.3. Помехи в каналах связи	154
5.2. Математические модели каналов связи	155
5.2.1. Модель непрерывного канала без шума	155
5.2.2. Модель непрерывного канала с аддитивным гауссовым шумом	156
5.2.3. Канал с неопределенной фазой сигнала	157
5.2.4. Канал с межсимвольной интерференцией и аддитивным шумом.....	158
5.2.5. Математические модели дискретных каналов связи	159
5.2.6. Математическая модель дискретно-непрерывного канала связи	161
5.2.7. Особенности оптических каналов связи.....	161
5.3. Воздействие сигналов и помех на линейные и нелинейный звенья каналов связи.....	163
5.3.1. Постановка задачи	163
5.3.2. Прохождение случайных сигналов и помех через линейные системы.....	163
5.3.3. Прохождение случайных сигналов и помех через нелинейные системы	166
Контрольные вопросы	168
Глава 6. Теория передачи и кодирования сообщений.....	169

6.1. Общие сведения	169
6.1.1. Количество информации в дискретных сообщениях. Понятие энтропии	169
6.1.2. Производительность источника дискретных сообщений. Скорость передачи информации	173
6.2. Эффективное кодирование источников дискретных сообщений	174
6.3. Информационные характеристики источников непрерывных сообщений	176
6.3.1. Дифференциальная энтропия	176
6.3.2. ϵ -энтропия	179
6.4. Пропускная способность канала связи	181
6.4.1. Пропускная способность дискретного канала связи	181
6.4.2. Пропускная способность непрерывного канала связи	183
6.5. Основы теории помехоустойчивого кодирования	185
6.5.1. Теорема кодирования для канала с помехами	185
6.5.2. Принципы построения помехоустойчивых кодов	186
6.5.3. Классификация помехоустойчивых кодов	193
6.5.4. Линейные двоичные блочные коды	193
6.5.5. Понятие о циклических, сверточных и каскадных кодах	200
Контрольные вопросы	212
Глава 7. Теория помехоустойчивости систем связи	213
7.1. Оптимальный прием элементов дискретных сообщений. Критерии оптимального приема	213
7.2. Синтез оптимальных приемников при приеме элементов дискретных сигналов, известных точно	217
7.2.1. Реализация оптимального приема на базе коррелятора	217
7.2.2. Бинарное обнаружение сигналов	220
7.2.3. Бинарное распознавание сигналов	222
7.2.4. Реализация алгоритма оптимального приема на основе согласованных фильтров	224
7.3. Оптимальный прием при неопределенной фазе и амплитуде элементов дискретного сигнала	228
7.3.1. Оптимальный прием при неопределенной фазе дискретного сигнала	228
7.3.2. Оптимальный прием при неопределенной фазе и амплитуде элементов сигнала	234
7.4. Неоптимальные методы приема дискретных сообщений	235
7.5. Цифровые методы обработки сигналов в приемнике	238
7.6. Методы приема дискретных сообщений в каналах с рассеянием	239
7.7. Способы борьбы с импульсными, сосредоточенными и межсимвольными помехами	242
7.8. Анализ помехоустойчивости систем связи с различными методами приема сигналов	244
7.8.1. Потенциальная помехоустойчивость двоичной системы при приеме элементов сигналов, известных точно	244

7.8.2. Помехоустойчивость приема m сигналов, известных точно	246
7.8.3. Потенциальная помехоустойчивость оптимального приемника ЧМ сигналов с неизвестной начальной фазой	249
7.8.4. Потенциальная помехоустойчивость оптимального приемника АМ сигнала с неизвестной начальной фазой ..	250
7.8.5. Потенциальная помехоустойчивость приема дискретных сообщений при замираниях сигнала	251
7.8.6. Помехоустойчивость приема оптических сигналов	252
7.9. Оптимальный прием непрерывных сообщений	256
7.9.1. Показатели качества передачи	256
7.9.2. Прием отдельных значений непрерывных сообщений. Критерии оптимальности	257
7.9.3. Прием сигналов, модулированных сообщением	261
7.9.4. Сравнительная оценка помехоустойчивости оптимальных систем с различными видами модуляции ..	264
7.9.5. Порог помехоустойчивости в системах с широкополосными видами модуляции	267
7.10. Оптимальная линейная фильтрация. Понятие о нелинейной фильтрации	270
7.10.1. Линейная фильтрация сигнала	270
7.10.2. Понятие о нелинейной фильтрации	274
7.11. Цифровые методы передачи непрерывных сообщений	279
7.11.1. Общие сведения о цифровых методах передачи. Помехоустойчивость цифровых методов передачи непрерывных сообщений	279
7.11.2. Структурная схема цифровой системы передачи непрерывных сообщений	282
7.11.3. Мощность шума квантования	282
7.11.4. Влияние ошибок приема на точность восстановления сообщения	284
7.11.5. Дифференциальная импульсно-кодовая модуляция	285
Контрольные вопросы	288

Глава 8. Принципы многоканальной связи и распределения информации

и распределения информации	289
8.1. Основы теории разделения сигналов	289
8.2. Методы частотного, временного и фазового разделения сигналов	291
8.2.1. Частотный метод разделения сигналов	291
8.2.2. Временное и фазовое разделение сигналов	293
8.3. Другие виды разделения сигналов	297
8.3.1. Разделение сигналов по форме	297
8.3.2. Комбинационное разделение	302
8.3.3. Пространственное разделение	303
8.4. Пропускная способность многоканальных систем	305
8.5. Сеть электросвязи и ее элементы. Основные понятия теории распределения информации	306
Контрольные вопросы	311

Глава 9. Методы повышения эффективности систем связи.....	312
9.1. Оценка эффективности систем связи.....	312
9.2. Принципы системного подхода к исследованию и разработке систем связи	313
9.3. Анализ систем связи по их эффективности	315
9.3.1. Энергетическая и частотная эффективность систем связи	315
9.3.2. Выбор способов модуляции и помехоустойчивого кодирования	317
9.3.3. Совместная оптимизация модуляции и кодирования	321
9.3.4. Использование обратного канала для повышения эффективности передачи дискретных сообщений	323
9.3.5. Устранение избыточности	325
Контрольные вопросы	326
Литература	327

Учебное издание

Леонид Леонидович Клюев

Теория электрической связи

Редактор: *д.т.н., профессор Фельдштейн Е.Э.*

Технический редактор: *Еремина Е.В.*

Подписано в печать с оригинал-макета 15.06.98. Формат 60×90 1/16. Бумага офсетная.

Печать офсетная. Усл. печ. л. 21. Уч. изд. л. 21. Тираж 1000 экз. Заказ 5380.

Лицензия ЛПЗ № 93 от 2.12.1997.

Издательство "Дизайн ПРО", 220027, Минск, пр. Ф.Скорины, 65

Отпечатано в типографии "Победа", 222310, г.Молодечно, ул.Тавлая, 11