

Министерство образования и науки Российской Федерации

Федеральное агентство по образованию

Государственное образовательное учреждение

высшего профессионального образования

Российский государственный гидрометеорологический университет

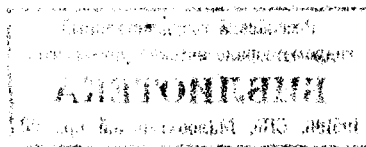
Е.А.Чернецова

Теория передачи дискретных сообщений

Конспект лекций

Санкт-Петербург

2007



УДК 681.3.07

Чернецова Е.А. Теория передачи дискретных сообщений: конспект лекций. – СПб.: Изд-во РГГМУ, 2007. – 166 с.

Рецензент И.В. Алешин, проф. Санкт-Петербургского государственного морского технического университета.

Конспект лекций предназначен для студентов специальности «Информационная безопасность телекоммуникационных систем» и всех желающих познакомиться с основными принципами, на которых основана работа систем связи. Конспект лекций содержит одиннадцать лекций, являющихся введением в теорию передачи дискретных сообщений. В лекциях достаточно подробно описаны основные положения цифровой связи, относящиеся к сигналам, поведению канала и модуляции. Предполагается, что студенты имеют начальные знания по теории вероятности. Все математические выкладки, изложенные в тексте, поданы в максимально упрощенном виде. Там, где необходимо, формулы сопровождаются соответствующими пояснениями. Также для пояснения излагаемого материала приведены примеры и расчеты в виде задач и их решений. Данный конспект лекций позволит читателю самостоятельно продолжить изучение систем связи.

© Чернецова Е.А., 2007

© Российский государственный
гидрометеорологический университет (РГГМУ), 2007

Российский государственный
гидрометеорологический университет

БИБЛИОТЕКА

195196, СПб, Малоохтинский пр., 98

92. К. 1317

ЛЕКЦИЯ 1

Обработка сигналов в цифровой связи

1.1. Причина замены аналоговой связи на цифровую

Основным преимуществом цифровой связи является легкость восстановления цифровых сигналов по сравнению с аналоговыми. Рассмотрим рис. 1.1, на котором представлен идеальный двоичный цифровой импульс, распространяющийся по каналу передачи данных. На форму сигнала влияют два основных механизма:

1. Поскольку все каналы и линии передачи имеют неидеальную частотную характеристику, идеальный импульс искажается.

2. Нежелательные электрические шумы или другое воздействие со стороны еще больше искажает форму импульса. Чем протяженнее канал, тем существеннее эти механизмы искажают импульс (рис. 1.1). В тот момент, когда переданный импульс все еще может быть достоверно определен (прежде чем он ухудшится до неоднозначного состояния), импульс усиливается цифровым усилителем, восстанавливающим его первоначальную идеальную форму. Импульс “возрождается” или восстанавливается. За восстановление сигнала отвечают регенеративные ретрансляторы, расположенные в канале связи на определенном расстоянии друг от друга.

Цифровые каналы менее подвержены искажению и интерференции, чем аналоговые. Поскольку двоичные цифровые каналы дают значимый сигнал только при работе в одном из двух состояний — включенном или выключенном — возмущение должно быть достаточно большим, чтобы перевести операционную точку канала из одного состояния в другое. Наличие всего двух состояний облегчает восстановление сигнала и, следовательно, предотвращает накопление в процессе передачи шумов или других возмущений. Аналоговые сигналы, наоборот, не являются сигналами с двумя состояниями; они могут принимать бесконечное множество форм. В аналоговых каналах даже небольшое возмущение может неузнаваемо исказить сигнал. После искажения аналогового сигнала возмущение нельзя убрать путем усиления. Поскольку накопление шума неразрывно связано с аналоговыми сигналами, как следствие, они не могут воспроизводиться идеально. При использовании цифровых технологий очень низкая частота возникновения ошибок плюс применение процедур выявления и коррекции ошибок делают возможным высокую точность сигнала. Остается только отметить, что с аналоговыми технологиями подобные процедуры недоступны.



Рис. 1.1. Искажение и восстановление формы импульса.

Существуют и другие преимущества цифровой связи. Цифровые каналы надежнее и могут производиться по более низким ценам, чем аналоговые. Кроме того, цифровое программное обеспечение позволяет более гибкую реализацию, чем аналоговое (например, микропроцессоры, цифровая коммутация и большие интегральные схемы). Использование цифровых сигналов и уплотнения с временным разделением (time-division multiplexing – TDM) проще применения аналоговых сигналов и уплотнения с частотным разделением (frequency-division multiplexing – FDM). При передаче и коммутации различные типы цифровых сигналов (данные, телеграф, телефон, телевидение) могут рассматриваться как идентичные: ведь бит — это и есть бит. Кроме того, для удобства коммутации и обработки цифровые сообщения могут группироваться в автономные единицы, называемые пакетами. В цифровые технологии естественным образом внедряются функции, защищающие от интерференции и подавления сигнала либо обеспечивающие шифрование или секретность. Кроме того, обмен данными в основном производится между двумя компьютерами или между компьютером и цифровыми устройствами или терминалом. Подобные цифровые оконечные устройства лучше (и естественнее!) обслуживаются цифровыми каналами связи.

Чем же мы платим за преимущества систем цифровой связи? Цифровые системы требуют более интенсивной обработки, чем аналоговые. Кроме того, для цифровых необходимо выделение значительной части ресурсов для синхронизации на личных уровнях. Аналоговые системы, наоборот, легче синхронизировать. Еще одним недостатком систем цифровой связи является то, что ухудшение качества носит пороговый характер. Если отношение сигнал/шум падает ниже некоторого порога, качество обслуживания может внезапно измениться от очень хорошего до очень плохого. В

аналоговых же системах ухудшение качества происходит более плавно.

Главным же преимуществом цифровой связи по сравнению с аналоговой является то, что цифровая связь позволяет реализовать любое, априорно заданное качество связи, выражающееся в вероятности битовой ошибки.

1.2. Основные понятия и определения цифровой связи

Информация – это сведения, являющиеся объектом передачи, распределения, преобразования, хранения или непосредственного использования.

Сообщение является формой представления информации.

Одно и то же сведение может быть представлено в различной форме. Например, сведение о моменте начала наступления может быть передано по телефону или телеграфом или тремя зелеными ракетами. В первом случае мы имеем дело с информацией, представленной в непрерывном виде (непрерывное сообщение). Будем считать, что это сообщение вырабатывается источником непрерывных сообщений. Во втором и третьем случаях – с информацией, представленной в дискретном виде (дискретное сообщение). Это сообщение вырабатывается источником дискретных сообщений.

Основное отличие дискретного и непрерывного источников состоит в следующем. Множество всех различных сообщений, вырабатываемых дискретным источником, всегда конечно. Поэтому на конечном отрезке времени количество символов дискретного источника также является конечным. В то же время число возможных различных значений звукового давления (или напряжения в телефонной линии), измеренное при разговоре, даже на конечном отрезке времени будет бесконечным.

Информация, содержащаяся в сообщении, передается от источника сообщений к получателю по каналу передачи дискретных сообщений (ПДС) (рис. 1.2).

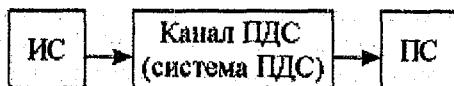


Рис. 1.2. Тракт передачи дискретных сообщений.

Характеристики источника дискретных сообщений.

Сообщение поступает от источника дискретных сообщений, который характеризуется алфавитом передаваемых сообщений $A = \{a_1, a_2, \dots, a_n\}$.

Алфавит – есть совокупность всех возможных (различных) сообщений (знаков) данного источника.

Объем алфавита – число различных символов алфавита K .

Принципиальное отличие систем аналоговой и цифровой связи связано со способом оценки их производительности. Критерием производительности систем связи является точность воспроизведения на приемном конце переданного сигнала. В качестве оценки производительности аналоговых систем может выступать процент искажения или ожидаемая среднеквадратическая ошибка между переданным и принятым сигналами. Критерием качества связи в цифровых системах является вероятность неверного обнаружения цифры или вероятность ошибки.

В системе цифровой связи каждое сообщение алфавита a_i появляется с некоторой вероятностью.

Вероятность выдачи символа (сообщения) – $p(a_i)$.

Количество информации в сообщении (символе) определяется вероятностью его появления. Чем меньше вероятность появления того или иного сообщения, тем большее количество информации мы извлекаем при его получении. В 1928 г. Хартли предложил определять количество информации, которое приходится на одно сообщение a_i , выражением:

$$I(a_i) = \log_2 \frac{1}{p(a_i)} = -\log_2 p(a_i). \quad (1.1)$$

Энтропия. Среднее количество информации $H(A)$, которое приходится на одно сообщение, поступающее от источника без памяти, получим, применив операцию усреднения по всему объему алфавита

$$H(A) = -\sum_{i=1}^K p(a_i) \log_2 p(a_i), \quad (1.2)$$

Выражение (1.2) известно как формула Шеннона для энтропии источника дискретных сообщений. *Энтропия* – мера неопределенности в поведении источника дискретных сообщений. Энтропия равна нулю, если с вероятностью единица источником выдается всегда одно и то же сообщение (в этом случае неопределенность в поведении источника сообщений отсутствует). Энтропия максимальна, если символы источника появляются независимо и с одинаковой вероятностью.

Один бит – это количество информации, которое переносит один символ источника дискретных сообщений в том случае, когда алфавит источника состоит из двух равновероятных символов.

Среднее количество информации, выдаваемое источником в единицу времени, называют *производительностью источника*:

$$H^*(A) = \frac{H(A)}{t_{cp}} \quad [\text{бит/с}], \quad (1.3)$$

где t_{cp} – среднее время, отводимое на передачу одного символа (сообщения).

Основные характеристики канала ПДС.

Для каналов передачи дискретных сообщений вводят аналогичную характеристику – *скорость передачи информации* по каналу R . Она определяется количеством бит, передаваемых в секунду. Максимально возможное значение скорости передачи информации по каналу называется *пропускной способностью канала* и обозначается C .

Пропускная способность непрерывного канала с белым гауссовским шумом определяется известной формулой Шеннона:

$$C = F \log_2 \left(1 + \frac{P_c}{P_{ш}} \right). \quad (1.4)$$

Как видно из выражения (1.4), данная величина определяется шириной полосы пропускания и соотношением сигнал-шум. Формула (1.4) определяет теоретическую верхнюю границу безошибочной пропускной способности канала связи. Если действительная пропускная способность канала меньше, то при использовании соответствующих кодов коррекции ошибок можно приблизиться к величине, определяемой формулой (1.4).

Сигналы – форма сообщения для передачи по каналу связи. Любая система связи обеспечивает передачу именно сигналов, а не сообщений. Поэтому сообщение, поступающее от источника, предварительно должно быть преобразовано в *сигнал определенной природы* (электрический, оптический ...), который является его переносчиком в данной системе связи.

Сигналы, формируемые на выходе преобразователя дискретного сообщения, как правило, являются по информационному параметру дискретными, т. е. описываются функцией дискретного времени и конечным множеством возможных значений. В технике передачи данных такие сигналы называют *цифровыми сигналами данных* (ЦСД).

Рассмотрим далее *основные определения, относящиеся к ЦСД.*

Представляющий (информационный) параметр сигнала данных – параметр сигнала данных, изменение которого отображает изменение сообщения.

На рис. 1.3 изображен ЦСД, представляющим параметром которого является амплитуда, а множество возможных значений представляющего параметра равно двум ($U=U_1$ и $U=0$).

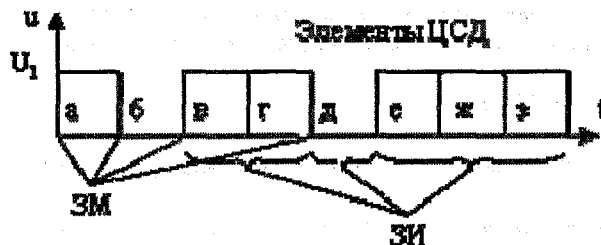


Рис. 1.3. Цифровой сигнал данных.

Элемент ЦСД – часть цифрового сигнала данных, отличающаяся от остальных частей значением одного из своих представляющих параметров.

Значащая позиция – фиксируемое значение состояния представляющего параметра сигнала.

Значащий момент (ЗМ) – момент, в который происходит смена значащей позиции сигнала.

Значащий интервал времени – интервал времени между двумя соседними значащими моментами сигнала.

Единичный интервал – минимальный интервал времени, которому равны значащие интервалы времени сигнала (интервалы а-б, б-в и другие на рис.1.3).

Единичный элемент – элемент сигнала, имеющий длительность, равную единичному интервалу времени.

1.3. Типичная блочная диаграмма и основные преобразования в цифровой системе связи

Функциональная блочная диаграмма, приведенная на рис.1.4, иллюстрирует распространение сигнала и этапы его обработки в типичной системе цифровой связи. Верхние блоки – форматирование, кодирование источника, шифрование, канальное кодирование, уплотнение, импульсная модуляция, полосовая модуляция, расширение спектра и множественный доступ — отражают преобразования сигнала на пути от источника к передатчику. Нижние блоки диаграммы — преобразования сигнала на пути от приемника к получателю информации, и, по сути, они противоположны верхним блокам. Блоки модуляции и демодуляции/обнаружения вместе называются модемом. Термин “модем” часто объединяет несколько этапов обработки сигналов, показанных на рис.1.4; в этом случае модем можно представлять как мозг системы. Передатчик и приемник можно рассматривать как “мускулы” системы. Для

беспроводных приложений передатчик состоит из схемы повышения частоты в область радиочастот, усилителя мощности и антенны, а приемник — из антенны и малошумящего усилителя. Обратное понижение частоты производится на выходе приемника и/или демодулятора. На рис.1.4 иллюстрируется соответствие блоков верхней (передающей) и нижней (принимающей) частей системы. Этапы обработки сигнала, имеющие место в передатчике, являются преимущественно обратными к этапам приемника.

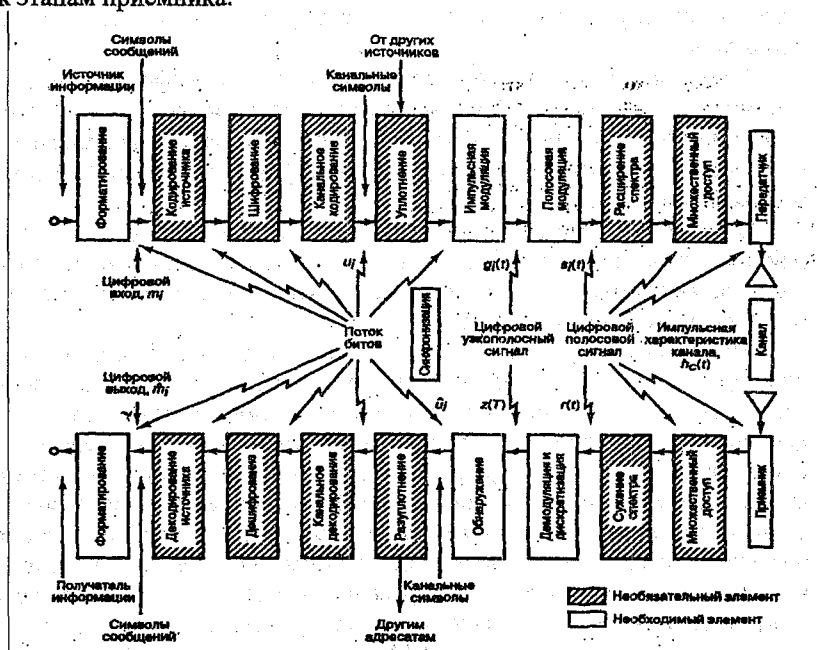


Рис. 1.4. Блочная диаграмма типичной системы цифровой связи.

На рис. 1.4 исходная информация преобразуется в двоичные цифры (биты). После этого биты группируются в цифровые сообщения или символы сообщений. Каждый такой символ (m_i , где $i = 1, \dots, M$) можно рассматривать как элемент конечного алфавита, содержащего M элементов. Следовательно, для $M=2$ символ сообщения m_i является бинарным (т.е. состоит из одного бита). Несмотря на то что бинарные символы можно классифицировать как M -арные (с $M=2$), обычно название “ M -арный” используется для случаев $M > 2$; значит, такие символы состоят из последовательности двух или большего числа битов.

Для систем, использующих канальное кодирование (коды коррекции ошибок), последовательность символов сообщений преобразуется в последовательность канальных символов (кодовых символов), и каждый канальный символ обозначается u_i . Поскольку символы сообщений или канальные символы могут состоять из одного бита или группы битов, последовательность подобных символов называется потоком битов.

Рассмотрим ключевые блоки обработки сигналов, изображенные на рис. 1.4; Необходимыми для цифровых систем связи являются только этапы форматирования, модуляции, демодуляции/обнаружения и синхронизации.

Форматирование преобразовывает исходную информацию в биты, обеспечивая, таким образом, совместимость информации и функций обработки сигналов с системой связи. С этой точки рисунка и вплоть до блока импульсной модуляции информация остается в форме потока битов.

Модуляция — это процесс, посредством которого символы сообщений или канальные символы (если используется канальное кодирование) преобразуются в сигналы, совместимые с требованиями, налагаемыми каналом передачи данных.

Импульсная модуляция — это еще один необходимый этап, поскольку каждый символ, который требуется передать, вначале нужно преобразовать из двоичного представления (уровни напряжений представляют двоичные нули и единицы) в форму узкополосного сигнала. Термин “узкополосный” (baseband) определяет сигнал, спектр которого начинается от (или около) постоянной составляющей и заканчивается некоторым конечным значением (обычно не более нескольких мегагерц). Блок импульсно-кодовой модуляции обычно включает фильтрацию, направленную на минимизацию полосы передачи. При применении импульсной модуляции к двоичным символам результирующий двоичный сигнал называется сигналом в кодировке PCM (pulse-code modulation) (импульсно-кодовая модуляция). Существует несколько типов сигналов PCM; в приложениях телефонной связи эти сигналы часто называются кодами канала. При применении импульсной модуляции к небинарным символам результирующий сигнал именуется M-арным импульсно-модулированным. Существует несколько типов подобных сигналов. После импульсной модуляции каждый символ сообщения или канальный символ принимает форму полосового сигнала $g_i(t)$, где $i = 1, \dots, M$. В любой электронной реализации поток битов, предшествующий импульсной модуляции, представляется уровнями напряжений. Может возникнуть вопрос, почему существует отдельный блок для импульсной модуляции, когда уровни напряжения для двоичных нулей и единиц уже можно рассматривать как

идеальные прямоугольные импульсы, длительность каждого из которых равна времени передачи одного бита. Существует два важных отличия между подобными уровнями напряжения и полосовыми сигналами, используемыми для модуляции. Во-первых, блок импульсной модуляции позволяет использовать бинарные и М-арные сигналы. Во-вторых, фильтрация, производимая в блоке импульсной модуляции, формирует импульсы, длительность которых больше времени передачи одного бита. Фильтрация позволяет использовать импульсы большей длительности, таким образом, импульсы расширяются на соседние временные интервалы передачи битов. Этот процесс иногда называется формированием импульсов; он используется для поддержания полосы передачи в пределах некоторой желаемой области спектра.

Для приложений, включающих передачу в диапазоне радиочастот, следующим важным этапом является *полосовая модуляция* (bandpass modulation). Она необходима всегда, когда среда передачи не поддерживает распространение сигналов, имеющих форму импульсов. В таких случаях среда требует полосового сигнала $s_i(t)$, где $i = 1, \dots, M$. Термин (bandpass) используется для отражения того, что узкополосный сигнал $g_i(t)$ сдвинут несущей волной на частоту, гораздо большую спектральных составляющих $g_i(t)$. По мере распространения сигнала $s_i(t)$ по каналу, на него воздействуют характеристики канала, которые можно выразить через импульсную характеристику $h_c(t)$. Кроме того, в различных точках вдоль маршрута сигнала, дополнительные случайные шумы искажают принятый сигнал $r(t)$, поэтому прием должен выражаться через поврежденную версию сигнала поступающего от передатчика. Принятый сигнал $r(t)$ можно выразить следующим образом:

$$r(t) = s_i(t) * h_c(t) + n(t); i = 1, \dots, M, \quad (1.5)$$

где знак $*$ представляет собой операцию свертки.

В обратном направлении входной каскад приемника и/или демодулятор обеспечивают понижение частоты каждого полосового сигнала $r(t)$. В качестве подготовки к обнаружению демодулятор восстанавливает $r(t)$ в виде оптимального узкополосного сигнала $z(t)$. Обычно с приемником и демодулятором связано несколько фильтров. Фильтрация производится для удаления нежелательных высокочастотных составляющих и формирования импульса.

Если импульсная характеристика канала связи $h_c(t)$ настолько плоха, что принимаемый сигнал сильно искажен, применяется *выравнивание* (equalization), которое можно описать как разновидность фильтрации, используемое для удаления всех эффектов ухудшения качества сигнала, причиной которых может быть канал.

Остальные этапы обработки сигнала в модеме являются необязательными и направлены на удовлетворение специфических системных нужд.

Кодирование источника (source coding) – это преобразование аналогового сигнала в цифровой и удаление избыточной (ненужной) информации. Система может применять либо форматирование, либо кодирование источника.

Шифрование используется для обеспечения конфиденциальности связи.

Канальное кодирование (channel coding) – применяется для обеспечения заданного качества связи (уменьшения вероятности ошибки).

Уплотнение и множественный доступ объединяют сигналы с тем, чтобы они совместно могли использовать ресурс связи (полоса частот и время связи). Отличие терминов заключается в том, что уплотнение подразумевает объединение сигналов пользователей, требования которых к ресурсу связи остаются неизменными во времени, а сами пользователи находятся в одной точке пространства. Множественный доступ подразумевает объединение сигналов пользователей, требования которых к ресурсу связи меняются во времени, а сами пользователи находятся в разных точках пространства.

Расширение частоты может давать сигнал, относительно неуязвимый для помех, и может использоваться для повышения уровня конфиденциальности общающихся сторон. Также оно является ценной технологией для множественного доступа.

ЛЕКЦИЯ 2

2.1. Классификация сигналов

Сигнал можно классифицировать как детерминированный (в отсутствии неопределенности относительно его значения в любой момент времени) или случайный (в противном случае). Детерминированные сигналы моделируются математическим выражением (например, $x(t) = 5 \cos 10t$), для случайного сигнала такое выражение написать невозможно. Впрочем, при наблюдении случайного сигнала в течении длительного времени можно отметить некоторые его закономерности, которые можно описать через вероятности и среднее статистическое.

Сигнал $x(t)$ называется *периодическим во времени*, если существует постоянное $T_0 > 0$, такое, что

$$x(t) = x(t + T_0) \text{ для } -\infty < t < \infty, \quad (2.1)$$

где через t обозначено время.

Наименьшее T_0 , удовлетворяющее это условие, называется периодом сигнала $x(t)$. Период определяет длительность одного полного цикла функции $x(t)$. Сигнал, для которого не существует значения T_0 , удовлетворяющего выражению (2.1), называется *непериодическим*.

Аналоговый сигнал является непрерывной функцией времени. *Дискретный сигнал* $x(kT)$, в котором k -целое число, T - фиксированный промежуток времени, характеризуется последовательностью чисел, определенных для каждого момента времени.

Электрический сигнал, возникший после преобразования некоторым устройством *физического сигнала* (например, речи), можно представить как изменение напряжения $v(t)$ или тока $i(t)$ с мгновенной мощностью $p(t)$, подаваемой на сопротивление R :

$$p(t) = \frac{v^2(t)}{R} \quad (2.2a)$$

или

$$p(t) = i^2(t)R \quad (2.2b)$$

В системах связи мощность часто нормируется (предполагается, что сопротивление R равно 1 Ом, хотя в реальном канале оно может быть любым). В нормированном случае уравнения (2.2a) и (2.2b) имеют одинаковый вид:

$$p(t) = x^2(t) \quad (2.3)$$

где $x(t)$ – это либо напряжение, либо ток.

Рассеяние энергии в течении промежутка времени $(-T/2; T/2)$ реального сигнала с мгновенной мощностью, полученной с помощью уравнения (2.3), можно записать как

$$E = \int_{-T/2}^{T/2} x^2(t) dt \quad (2.4)$$

Средняя мощность, рассеянная сигналом в течение этого интервала, равна:

$$P = \frac{1}{T} E = \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt. \quad (2.5)$$

Производительность системы связи зависит от энергии принятого сигнала. Сигналы с более высокой энергией обнаруживаются более достоверно – работу по обнаружению выполняет принятая энергия. Мощность – это скорость поступления энергии.

Определенный сигнал можно отнести либо к энергетическому, либо к периодическому. Энергетический сигнал имеет конечную энергию, но нулевую среднюю мощность. В реальной ситуации всегда передаются сигналы с конечной энергией. Впрочем, периодические сигналы по определению существуют всегда и, следовательно, имеют бесконечную энергию и ненулевую конечную мощность. Для работы со случайными сигналами также удобно определить класс сигналов, выражаемых через мощность.

Общее правило: периодические и случайные сигналы выражаются через мощность, а сигналы, являющиеся детерминированными и непериодическими, – через энергию.

Полезной функцией в теории связи является *единичный импульс*, или *дельта-функция Дирака* $\delta(t)$. Это – абстракция, импульс с бесконечно большой амплитудой, нулевой шириной и единичным весом (площадью под импульсом), сконцентрированный в точке, в которой значение его аргумента равно нулю. Единичный импульс задается следующими соотношениями:

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (2.6)$$

$$\delta(t) = 0 \quad \text{для } t \neq 0 \quad (2.7)$$

$$\delta(t) \text{ не ограничена в точке } t=0 \quad (2.8)$$

$$\int_{-\infty}^{\infty} x(t) \delta(t - t_0) dt = x(t_0) \quad (2.9)$$

Если $\delta(t)$ входит в какую-либо операцию, его удобно считать импульсом конечной амплитуды, единичной площади и ненулевой длительности, после чего нужно рассмотреть предел при стремлении длительности импульса к нулю. $A\delta(t - t_0)$ с постоянной A представляет собой импульсную функцию, площадь которой (или вес) равна A , а значение везде нулевое, за исключением точки $t=t_0$.

Уравнение (2.9) известно как *просеивающее* (или *квантующее*) свойство $\delta(t)$. Интеграл от единичного импульса и произвольной функции дает выборку функции $x(t)$ в точке $t=t_0$.

2.2. Преобразование Фурье

Наиболее простой периодический сигнал – синусоидальная волна. В общем случае такую волну можно задать тремя параметрами – максимальной амплитудой A , частотой f и фазой ϕ . Максимальной амплитудой называется максимальное значение или интенсивность сигнала во времени (измеряется в вольтах). Частотой называется темп (в периодах за секунду – $\frac{1}{c}$

или герцах – Гц) повторения сигнала. Фаза является мерой относительного сдвига по времени в пределах одного периода сигнала. Ее можно определить как часть дробного отношения $\frac{t}{T}$, где T – период, а t – время, отсчи-

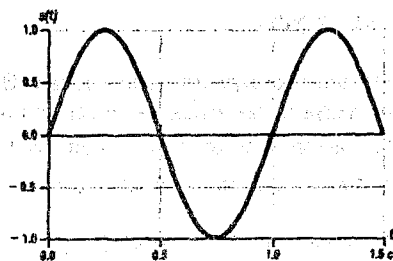
тываемое от произвольного начала отсчета. За начало отсчета обычно принимается последнее предыдущее прохождение интенсивности сигнала через ноль от отрицательного значения к положительному.

В общем случае синусоидальный сигнал можно представить в виде $s(t) = A \sin(2\pi ft + \phi)$. Эффект изменения каждого из трех параметров показан на рис. 2.1. По горизонтальной оси рис. 2.1 откладывается время. На самом же графике показана зависимость от времени величины сигнала в данной точке пространства. Такие же графики, при изменении масштаба, можно получить, если отложить на горизонтальной оси расстояние. В этом случае на графике будет изображена величина сигнала в данный момент времени в зависимости от расстояния. Например, при передаче синусоидальной волны (если рассмотреть электромагнитную радиоволну на некотором расстоянии от радиоантенны или звук на некотором расстоянии от громкоговорителя) в отдельный момент времени интенсивность сигнала меняется по гармоническому закону как функция расстояния от источника.

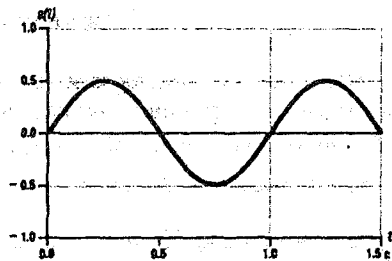
Для синусоидальных сигналов существует два простых соотношения, одно из которых связано со временем, а другое – с пространством. Определим длину волны λ как расстояние, занимаемое одним периодом или, иными словами, как расстояние между двумя точками равных фаз двух последовательных циклов. Предположим, что сигнал распространяется со скоростью v . Тогда длина волны связана с периодом следующим соотношением:

$$\lambda = vT \text{ или равносильно } \lambda f = v. \quad (2.10)$$

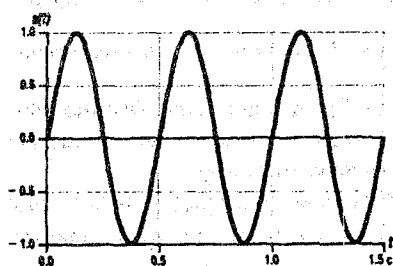
Важным частным случаем этого соотношения является случай $v = c = 3 \cdot 10^8$ м/с, где c – скорость света в вакууме.



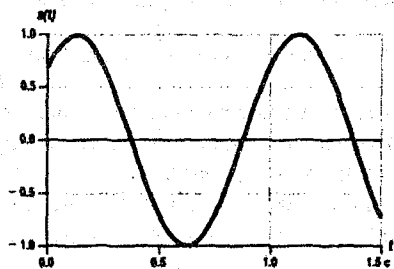
а) $A=1; f=1; \phi=0$



б) $A=0,5; f=1; \phi=0$



в) $A=1; f=2; \phi=0$



г) $A=0,5; f=1; \phi=\pi/4$

Рис. 2.1. Синусоидальный сигнал.

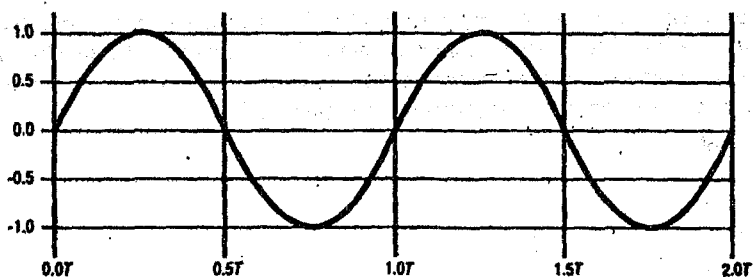
Реальный электромагнитный сигнал состоит из многих частот. Рассмотрим сигнал, представленный на рис. 2.2, аналитически записываемый в виде $s(t) = \frac{4}{\pi} [\sin(2\pi ft) + \frac{1}{3} \sin(2\pi(3f)t)]$. Составляющими

этого сигнала являются простые синусоидальные сигналы с частотами f и $3f$. Отметим несколько особенностей:

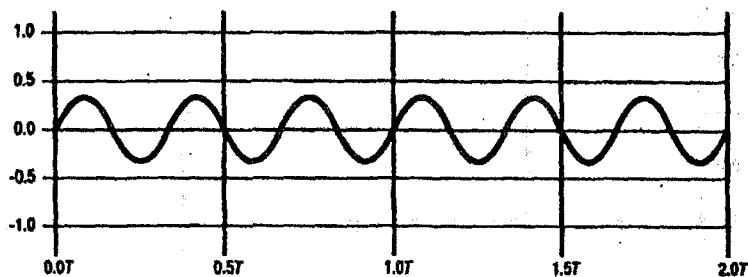
- Вторая частота кратная первой. Если частотные составляющие сигнала кратны одной частоте, то эта частота называется *собственной частотой*.

- Период суммарного сигнала равен периоду сигнала *собственной частоты*.

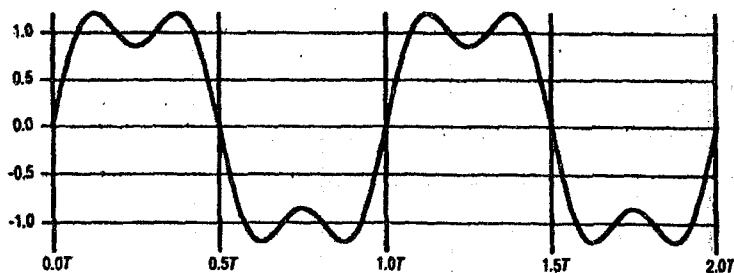
Можно показать, используя для этой цели методику, известную как анализ Фурье, что любой сигнал можно разложить на гармоники различных частот. Этот факт имеет большое значение в теории связи, так как воздействие на сигнал различных передающих сред можно выразить в терминах частот.



а) $\sin(2\pi ft)$



б) $\frac{1}{3} \sin(2\pi(3f)t)$



$$в) s(t) = \frac{4}{\pi} \left(\sin(2\pi ft) + \frac{1}{3} \sin(2\pi(3f)t) \right)$$

Рис. 2.2. Сложение гармоник.

Для каждого сигнала существует функция времени $s(t)$, задающая амплитуду сигнала в каждый момент времени. Подобным образом существует функция частоты $S(f)$, задающая максимальную амплитуду гармоник, составляющих сигнал. На рис. 2.3 показана частотная функция для сигнала, приведенного на рис. 2.2, в. В этом случае функция $S(f)$ дискретна. На рис. 2.4 показана частотная

уч.к. 1317

функция для отдельного прямоугольного импульса, имеющего значение 1 в области от $-X/2$ до $X/2$ и 0 в остальной области. В этом случае функция $S(f)$ непрерывна и принимает неограниченное число ненулевых значений, хотя амплитуда частотных компонент быстро уменьшается с ростом f .

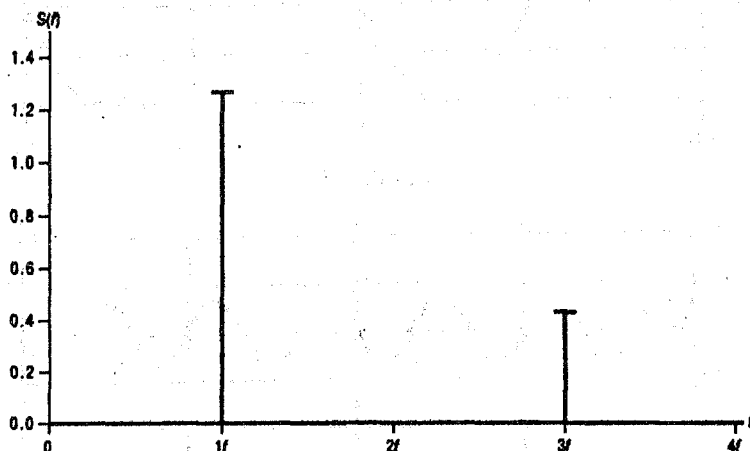


Рис. 2.3. Спектр гармонического сигнала.

$$s(t) = \frac{4}{\pi} \left[\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) \right].$$

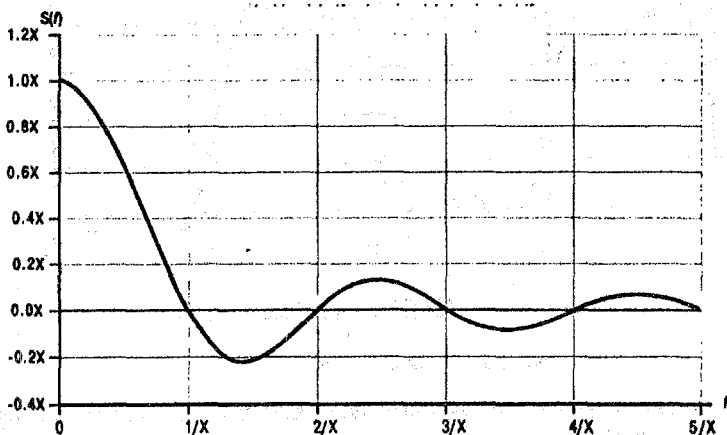
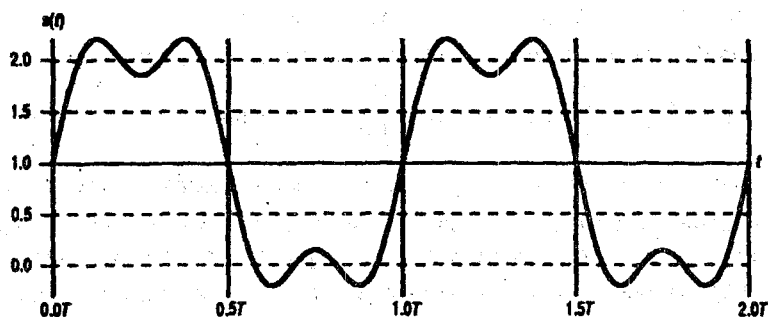
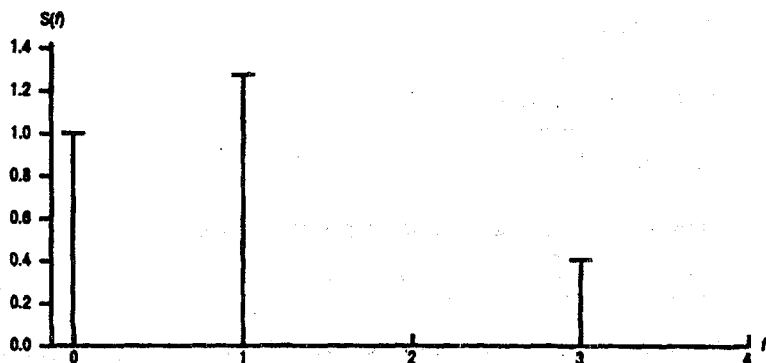


Рис. 2.4. Спектр прямоугольного импульса $s(t) = 1$ $-\frac{X}{2} \leq t \leq \frac{X}{2}$.

Если в сигнале имеется гармоника нулевой частоты, она называется постоянной составляющей. Если постоянной составляющей нет, средняя амплитуда сигнала равна нулю. На рис. 2.5 показан результат добавления постоянной составляющей к сигналу, изображенному на рис. 2.2, в.



$$a) s(t) = 1 + \frac{4}{\pi} (\sin(2\pi ft) + \frac{1}{3} \sin(2\pi(3f)t))$$



б) S(f)

Рис. 2.5. Сигнал с постоянной составляющей.

Спектром сигнала называется диапазон частот, составляющих данный сигнал. Для сигнала, приведенного на рис. 2.2, в спектр лежит в области от f до $3f$.

Периодический сигнал $g(t)$ с периодом T в общем случае имеет бесконечный спектр, который можно выразить в виде ряда Фурье:

$$g(t) = a_0 + a_1 \cos \alpha t + b_1 \sin \alpha t + a_2 \cos 2\alpha t + b_2 \sin 2\alpha t + \dots, (2.11)$$

где $\omega = 2\pi f$ – круговая частота ;

a_0 – постоянная составляющая сигнала, которая может отсутствовать;

$a_1 \cos \omega t$ и $b_1 \sin \omega t$ – гармоники основной частоты;

$a_2 \cos 2\omega t$ и $b_2 \sin 2\omega t$ – вторые гармоники;

Ряд (2.11) можно записать в более компактной форме:

$$g(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t + \sum_{n=1}^{\infty} b_n \sin n\omega t. \quad (2.12)$$

Пары коэффициентов a_n и b_n вместе со своими косинусными и синусными множителями являются ничем иным, как представлением одиночного вращающегося вектора в полярной системе координат (см. рис. 2.6). Амплитуда этого вектора равна c_n , а его угловая скорость ω рад/с. Каждый вектор c_n в момент времени $t=0$ имеет произвольную фазу, которая вычисляется по формуле:

$$\phi_n = \arctg \frac{a_n}{b_n}. \quad (2.13)$$

Амплитуда c_n равна

$$c_n = \sqrt{a_n^2 + b_n^2}. \quad (2.14)$$

Формулу (2.12) можно также переписать в виде:

$$g(t) = c_0 + \sum_{n=1}^{\infty} c_n \sin(n\omega t + f_n). \quad (2.15)$$

Между сигналом $s(t)$ и его спектром $S(\omega)$ существует однозначное соответствие. Для практических приложений важно установить связь между преобразованием сигнала и соответствующим ему преобразованием спектра.

1. Если сигнал запаздывает во времени, амплитуда его частотного спектра не меняется, а фазовый спектр сдвигается по фазе.

2. Умножение произвольного сигнала на синусоиду частоты f_0 приводит к трансляции исходного спектра сигнала на f_0 и $-f_0$ (теорема о модуляции (смещивании))

3. При сжатии сигнала в n раз по временной оси во столько же раз расширяется его спектр на оси частот. Модуль спектральной плотности

при этом уменьшается в n раз. При растягивании сигнала во времени спектр сужается и увеличивается модуль спектральной плотности.

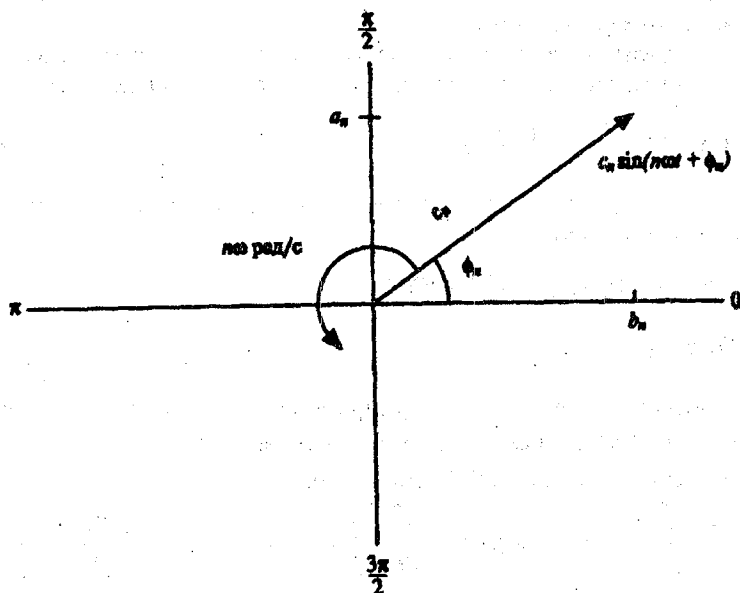


Рис. 2.6. Графическое представление гармоники.

Абсолютной шириной полосы сигнала называется ширина его спектра. На рисунке 2.3 ширина полосы сигнала равна $2f$. Ширина полосы сигнала, спектр которого представлен на рисунке 2.4, бесконечна. Но большая часть энергии сигнала приходится на относительно узкую полосу частот, называемую эффективной полосой.

2.3. Спектральная плотность

Спектральная плотность (spectral density) характеристик сигнала — это распределение энергии или мощности сигнала по диапазону частот. Используя теорему Парсеваля, можно связать энергию сигнала, выраженную во временной области, с энергией, выраженной в частотной области:

$$E = \int_{-T/2}^{T/2} x^2(t) dt = \int_{-T/2}^{T/2} |X(f)|^2 df, \quad (2.16)$$

где $|X(f)|$ — Фурье-образ непериодического сигнала $x(t)$.

Величина $\psi(f) = |X(f)|^2$ является спектральной плотностью энергии сигнала $x(t)$, описывает энергию сигнала на единицу ширины полосы и измеряется в Дж/Гц.

Спектральная плотность мощности периодического сигнала дает распределение мощности сигнала по диапазону частот и определяется как:

$$G(f) = \sum_{n=-\infty}^{\infty} |c_n|^2 \delta(f - nf_0). \quad (2.17)$$

Для непериодических сигналов можно выразить спектральную плотность мощности в пределе:

$$G(f) = \lim_{T \rightarrow \infty} \frac{1}{T} |X(f)|^2. \quad (2.18)$$

Пример

Найдите среднюю нормированную мощность сигнала $x(t) = A \cos \pi f t$, используя усреднение по времени.

Используя уравнение (2.5), получаем:

$$P = \frac{1}{T} \int_{-T/2}^{T/2} (A^2 \cos^2 2\pi f t) dt = \frac{A^2}{2T} \int_{-T/2}^{T/2} (1 + \cos 4\pi f t) dt = \frac{A^2 T}{2T} = \frac{A^2}{2}.$$

2.4. Случайные процессы

Случайный процесс $X(A, t)$ можно рассматривать как функцию двух переменных: события A и времени. На рис. 2.6 представлен пример случайного процесса. Каждую из N выборочных функций времени можно рассматривать как выход отдельного генератора шума. Для каждого события A_j есть единственная функция времени $X(A_j, t) = X_j(t)$, т.е. выборочная функция. Совокупность всех выборочных функций называется ансамблем. Поскольку значение случайного процесса в каждый последующий момент времени неизвестно, случайный процесс, функции распределения которого непрерывны, можно описать статистически через плотность вероятности.

Для нужд систем связи часто достаточно частичного описания, включающего среднее и функцию автокорреляции.

Определим среднее случайного процесса $X(t)$ как

$$E\{X(t_k)\} = \int_{-\infty}^{\infty} xp_{X_k}(x)dx = m_X(t_k), \quad (2.19)$$

где $X(t_k)$ – случайная переменная, полученная при рассмотрении случайного процесса в момент времени t_k ; $p_{X_k}(x)$ – плотность вероятности $X(t_k)$ (плотность по ансамблю событий в момент времени t_k).

Определим корреляционную функцию случайного процесса $X(t)$ как функцию двух переменных t_1 и t_2 :

$$R(t_1, t_2) = E\{X(t_1)X(t_2)\}, \quad (2.20)$$

где $X(t_1); X(t_2)$ – случайные переменные, получаемые при рассмотрении $X(t)$ в моменты времени t_1 и t_2 соответственно. *Автокорреляционная функция* – это мера связи двух временных выборок одного случайного процесса.

Случайный процесс называется *стационарным в строгом смысле*, если ни на одну из его статистик не влияет перенос начала отсчета времени. Случайный процесс называется *стационарным в широком смысле*, если две его статистики, среднее и автокорреляционная функция, не меняются при переносе начала отсчета времени:

$$E\{X(t)\} = m_X = \text{константа},$$

$$R(t_1, t_2) = R(t_1 - t_2) = R(\tau). \quad (2.21)$$

Для вычисления среднего и корреляционной функции путем усреднения по ансамблю нужно иметь полную информацию о взаимном распределении функций плотности вероятности. В общем случае, как правило, такая информация недоступна.

Если случайный процесс принадлежит к особому классу, называемому *классом эргодических процессов*, его среднее по времени равно среднему по ансамблю, и статистические свойства процесса можно определить путем усреднения по времени одной выборочной функции процесса.

Для эргодических процессов фундаментальные электротехнические параметры могут быть связаны с моментами процесса:

1. Величина $E\{X(t)\} = m_X$ равна постоянной составляющей сигнала
2. Величина m_X^2 равна нормированной мощности постоянной составляющей
3. Момент второго порядка $E\{X_2(t)\}$ равен общей средней нормированной мощности

4. Величина $\sqrt{E\{X^2(t)\}}$ равна среднеквадратическому значению сигнала, выраженного через ток или напряжение

5. Дисперсия σ_X^2 равна средней нормированной мощности переменного сигнала

6. Если среднее процесса равно нулю, а дисперсия равна среднеквадратическому значению, то дисперсия представляет общую мощность в нормированной нагрузке.

7. Среднеквадратическое отклонение σ_X^2 является среднеквадратическим значением переменного сигнала

2.5. Связь между скоростью передачи данных и шириной полосы

Хотя данный сигнал может содержать значительный диапазон частот, любая система связи на практике сможет вместить только ограниченную область частот. Рассмотрим прямоугольный сигнал (см. рис. 2.7). Пусть положительный импульс представляет двоичную единицу, а отрицательный – двоичный ноль. Длительность каждого импульса равна $1/2f$, следовательно, скорость передачи данных равна $2f$ бит/с. Из каких же гармоник составлен данный сигнал? Обратимся к рис. 2.2, в. При сложении синусоид с частотами f и $3f$ получается сигнал, форма которого начинает походить на форму прямоугольного сигнала. Если продолжить этот процесс и добавить синусоидальный сигнал с частотой $5f$ (рис. 2.8), а затем – синусоидальный сигнал с частотой $7f$ (рис. 2.9), то результирующий сигнал начнет все больше и больше приближаться к прямоугольной форме.

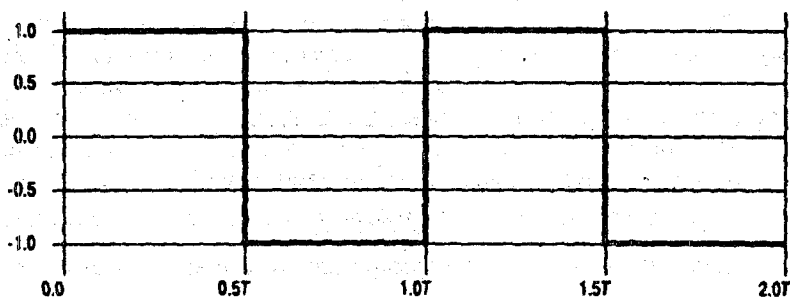


Рис. 2.7 Прямоугольный сигнал. $s(t) = \frac{4}{\pi} \sum_{k=1,3,5} \frac{1}{k} \sin(2\pi(kf)t)$.

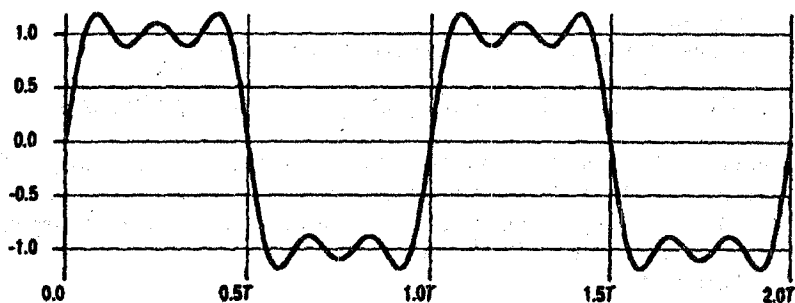


Рис. 2.8 $s(t) = \frac{4}{\pi} (\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) + \frac{1}{5} \sin(2\pi(5f)t))$.

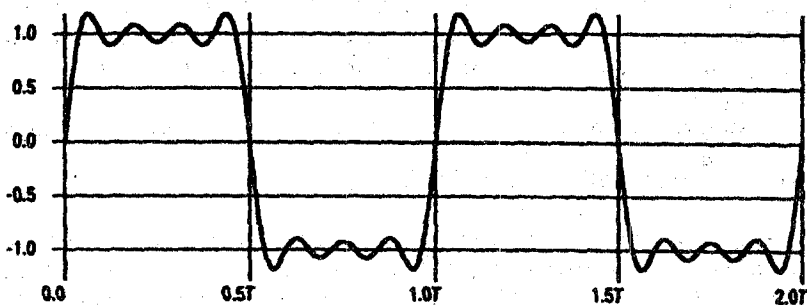


Рис. 2.9.

$$s(t) = \frac{4}{\pi} (\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) + \frac{1}{5} \sin(2\pi(5f)t) + \frac{1}{7} \sin(2\pi(7f)t)).$$

Гармоники прямоугольного сигнала с амплитудами A и $-A$ можно выразить как:

$$s(t) = A \frac{4}{\pi} \sum_{k=1,3,5} \frac{1}{k} \sin(2\pi(kf)t). \quad (2.22)$$

Таким образом, сигнал содержит бесконечное число гармоник и, следовательно, имеет бесконечную ширину полосы. Однако максимальная амплитуда k -той гармоники, kf , равна всего лишь $1/k$, поэтому большая часть энергии приходится на несколько первых гармоник.

Пример

Какой скорости передачи данных можно добиться в цифровой системе?

Случай 1

Аппроксимируем прямоугольный сигнал формой

$$s(t) = \frac{4}{\pi} (\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) + \frac{1}{5} \sin(2\pi(5f)t))$$

(см. рис. 2.8). Если частоту f положить равной 1 МГц, то период собственной частоты $E=1/10^6=1$ мкс, а ширина полосы сигнала равна $5-1=4$ МГц. Один бит будет поступать каждые 0,5 мкс, а скорость передачи данных при этом будет $2 \cdot 10^6=2$ МГц.

Случай 2

Если частоту f положить равной 2 МГц, то период собственной частоты в этом случае будет равен $T=1/2 \cdot 10^6=0,5$ мкс, а ширина полосы сигнала равна $5 \cdot 2 \cdot 10^6 - 1 \cdot 2 \cdot 10^6=8$ МГц. Один бит будет поступать каждые 0,25 мкс, а скорость передачи данных при этом будет $2 \cdot 2 \cdot 10^6=4$ МГц. Следовательно, *при равных других параметрах удвоение ширины полосы приводит к удвоению возможной скорости передачи данных.*

Случай 3

Аппроксимируем прямоугольный сигнал формой

$$s(t) = \frac{4}{\pi} (\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t))$$

(см. рис. 2.2,в). Если частоту f положить равной 2 МГц, то период собственной частоты $T=1/2 \cdot 10^6=0,5$ мкс, а ширина полосы сигнала равна $3 \cdot 2 \cdot 10^6 - 1 \cdot 2 \cdot 10^6=4$ МГц. Один бит будет поступать каждые 0,25 мкс, а скорость передачи данных при этом будет $2 \cdot 2 \cdot 10^6=4$ МГц. Следовательно, *при данной ширине полосы могут поддерживаться различные скорости передачи данных, а зависит эта скорость от способности приемника различать 0 и 1 при наличии помех и искажений сигнала.*

Таким образом, если передавать сигнал через среду, передающая система наложит ограничения на ширину полосы, которую можно передать. Кроме того, чем больше передаваемая полоса, тем выше стоимость передачи. Поэтому, с одной стороны, экономические и практические соображения подсказывают аппроксимировать цифровую информацию сигналом с ограниченной шириной полосы. С другой стороны, ограничение ширины полосы создает искажения, затрудняющие интерпретацию принимаемого сигнала (см. рис. 2.9).

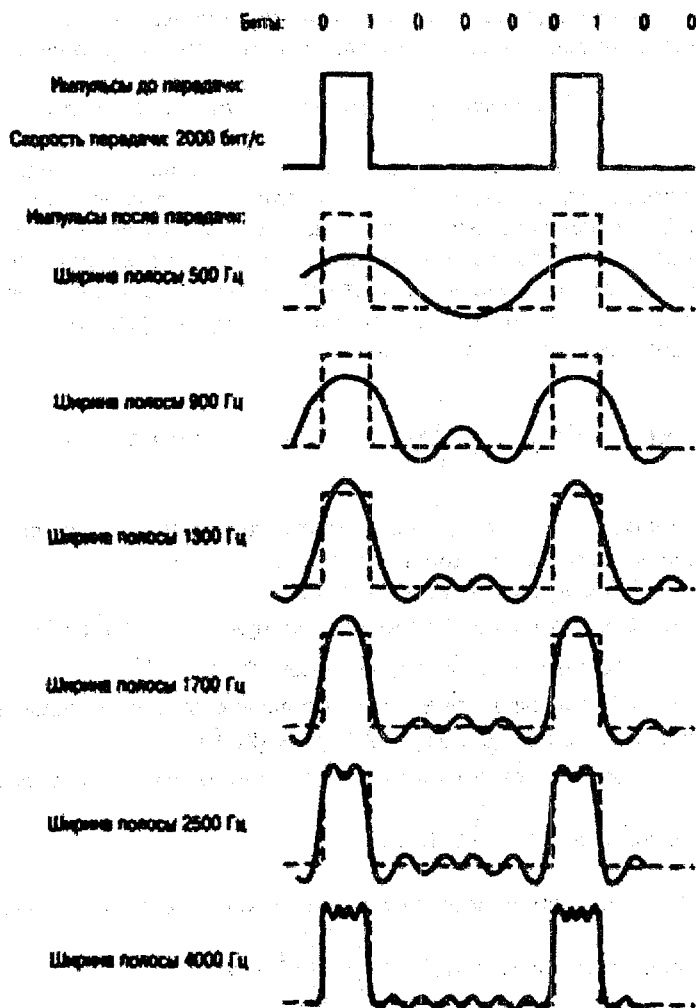


Рис. 2.9. Влияние ширины полосы на цифровой сигнал.

Чем больше ограничена полоса, тем больше потенциальная возможность возникновения ошибок при приеме. Если скорость передачи данных составляет для цифрового сигнала R бит/с, то достаточно хорошее представление сигнала можно получить при ширине полосы $2R$ Гц. Если полоса сигнала сосредоточена около некоторой частоты, называемой

центральной частотой, то чем выше центральная частота, тем шире возможная полоса и выше потенциальная скорость передачи данных.

2.6. Автокорреляционная функция сигнала

С преобразованием Фурье сигнала связана автокорреляционная функция сигнала (АКФ). Когда мы интересуемся корреляцией двух явлений, спрашиваем, насколько близко они соотносятся по поведению или виду и насколько они совпадают. В математике автокорреляционная функция сигнала во временной области описывает соответствие сигнала самому себе, смещенному на некоторый промежуток времени.

Для периодического сигнала

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)x(t+\tau)dt \quad \text{для } -\infty < \tau < \infty. \quad (2.23)$$

Функция автокорреляции – это функция разности времен τ между сигналом и его смещаемой по оси времени копией.

Свойства функции автокорреляции:

1. Симметрия по τ относительно нуля $R(\tau) = R(-\tau)$
2. Максимальное значение в нуле $R(\tau) \leq R(0)$
3. АКФ и энергетическая спектральная плотность сигнала являются Фурье-образами друг друга $R(\tau) \leftrightarrow \psi(f)$

4. Для энергетического сигнала значение АКФ в нуле равно энергии сигнала $R(0) = \int_{-\infty}^{\infty} x^2(t)dt$

5. Для периодического сигнала значение АКФ в нуле равно средней мощности сигнала $R(0) = \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)dt$

Мы последовательно перемещаем копию в положительном направлении временной оси и задаем в каждый момент времени вопрос, насколько исходная версия и копия сигнала соответствуют друг другу. Время τ рассматривается как параметр сканирования. Автокорреляционная функция может дать представление о ширине спектра сигнала. Предположим, что копия сигнала перемещается вдоль оси времени очень медленно. Это значит, что сигнал имеет малую ширину полосы,

тогда автокорреляционная функция будет иметь широкую форму. Если копия сигнала перемещается вдоль оси времени быстро (большая скорость передачи сигнала и соответственно ширина полосы), то автокорреляционная функция будет иметь узкую форму. Коэффициент корреляции ρ – значение функции корреляции в данный момент времени.

Рассмотрим случай $\rho = 1$, соответствующий наилучшей корреляции сигналов $s_1(t)$ и $s_2(t)$ в течение времени передачи символа. Если сигналы рассматривать как векторы, угол между ними будет равен нулю. В реальной системе цифровой связи сигналы (элементы алфавита) должны быть максимально несопоставимы, чтобы их можно было легко различать (обнаруживать). Частный случай $\rho = -1$ соответствует «антикорреляции» $s_1(t)$ и $s_2(t)$ в течение времени передачи символа. Угол между векторами сигналов в этом случае составляет 180° , а сигналы называются *антиподными*. Случай $\rho = 0$, соответствующий нулевой корреляции между сигналами $s_1(t)$ и $s_2(t)$, можно интерпретировать как угол 90° между векторами сигналов. В этом случае сигналы называются *ортогональными*. Чтобы сигналы были ортогональными, они не должны коррелировать в течение времени передачи символа, т.е. должно выполняться условие:

$$\int_0^T s_1(t)s_2(t)dt = 0. \quad (2.24)$$

Задача

Покажите, что три функции, приведенные на рис. 2.10, попарно ортогональны на интервале $(-2; 2)$.

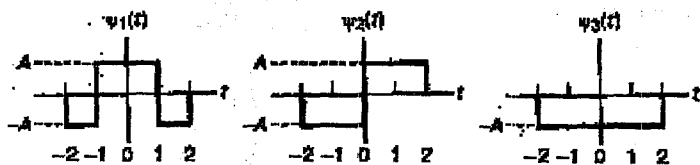


Рис. 2.10.

Решение

Проверим попарную ортогональность сигналов с учетом (2.24).

$$\begin{aligned}
\int_{-2}^2 \psi_1(t) \psi_2(t) dt &= \int_{-2}^{-1} (-A) * (-A) dt + \int_{-1}^0 A * (-A) dt + \int_0^1 A * A dt + \int_1^2 -A * (A) dt = \\
&= A^2 \int_{-2}^{-1} dt - A^2 \int_{-1}^0 dt + A^2 \int_0^1 dt - A^2 \int_1^2 dt = A^2(-1+2) - A^2(0+1) + A^2(1-0) - A^2(2-1) = 0 \\
\int_{-2}^2 \psi_1(t) \psi_3(t) dt &= \int_{-2}^{-1} (-A) * (-A) dt + \int_{-1}^0 A * (-A) dt + \int_0^1 A * (-A) dt + \int_1^2 (-A) * (-A) dt = \\
&= A^2 \int_{-2}^{-1} dt - A^2 \int_{-1}^0 dt - A^2 \int_0^1 dt + A^2 \int_1^2 dt = A^2(-1+2) - A^2(0+1) + A^2(1-0) - A^2(2-1) = 0 \\
\int_{-2}^2 \psi_2(t) \psi_3(t) dt &= \int_{-2}^{-1} (-A) * (-A) dt + \int_{-1}^0 (-A) * (-A) dt + \int_0^1 A * (-A) dt + \int_1^2 A * (-A) dt = \\
&= A^2 \int_{-2}^{-1} dt + A^2 \int_{-1}^0 dt - A^2 \int_0^1 dt - A^2 \int_1^2 dt = A^2(-1+2) - A^2(0+1) + A^2(1-0) - A^2(2-1) = 0
\end{aligned}$$

2.7. Импульсная характеристика

Линейная, инвариантная относительно времени система или сеть, показанная на рис. 2.11, описывается во временной области импульсной характеристикой $h(t)$, представляющей собой реакцию системы при подаче на ее вход единичного импульса $\delta(t)$ (см. рис. 2.12)

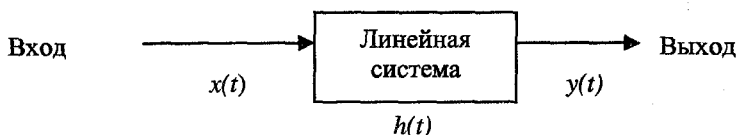


Рис. 2.11. Линейная система и ее ключевые параметры.

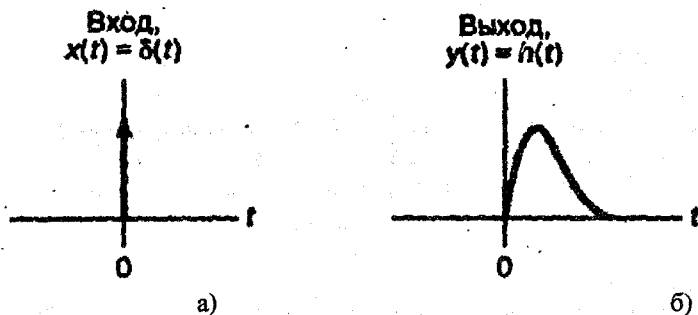


Рис. 2.12. Иллюстрация понятия «Импульсный отклик».

Задача

На рис. 2.13 приведен пример цепи запоминания, широко используемой в импульсных системах. Определите импульсную характеристику этого канала.

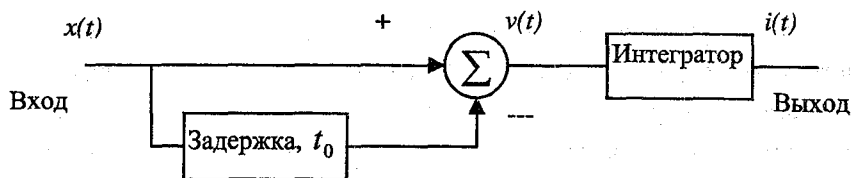


Рис. 2.13.

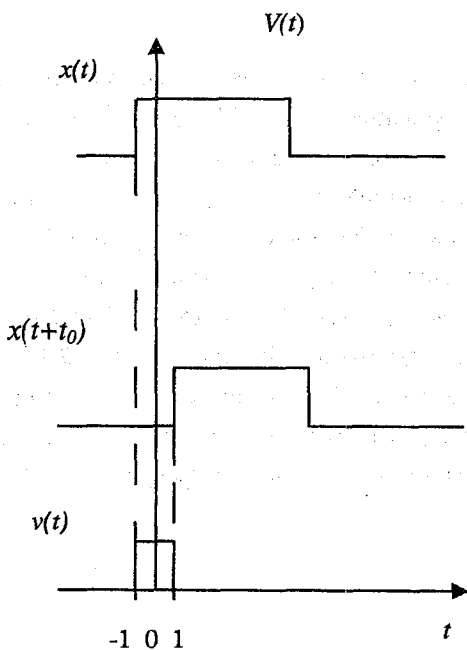


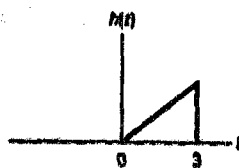
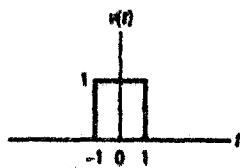
Рис. 2.14.

Решение

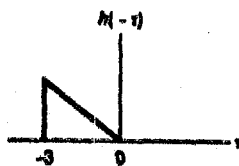
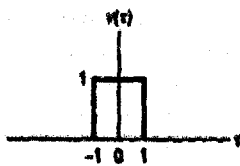
На рис. 2.14 показан входной сигнал $x(t)$, задержанный относительно него второй сигнал $x(t+t_0)$ и сигнал рассогласования $v(t)$. Интегратор представим линейной, инвариантной (неизменной) относительно времени системой, которая характеризуется импульсной характеристикой $h(t)$ (реакция системы при подаче на ее вход единичного импульса), которая изображена на рис. 2.15,а. Отклик системы на произвольный сигнал $v(t)$, поданный на ее вход, является сверткой $v(t)$ с $h(t)$, что записывается следующим образом:

$$i(t) = v(t) * h(t) = \int_0^{\infty} v(\tau) h(t - \tau) d\tau. \quad (2.25)$$

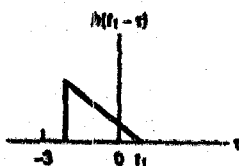
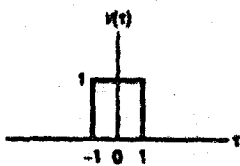
Независимой переменной в интеграле свертки является τ . На рис. 2.15,б показаны функции $v(\tau)$ и $h(-\tau)$. Функция $h(-\tau)$ получается отображением $h(\tau)$ относительно оси $\tau=0$. Член $h(t-\tau)$ представляет функцию $h(-\tau)$, смещенную на t секунд вдоль положительного направления оси τ . На рис. 2.15,в показана функция $h(t_1-\tau)$. Значение интеграла свертки в момент времени $t=t_1$ вычисляем как площадь под кривой произведения $v(\tau)$ на $h(t_1-\tau)$ (см. рис. 2.15,г). Интеграл свертки в момент $t=t_2$ равен заштрихованной области на рис. 2.15,д. На рис. 2.15,е приведен график отклика $i(t)$ на выходе схемы при квадратном импульсе на входе. Каждое вычисление интеграла свертки для некоторого момента времени t_i дает одну точку $i(t_i)$ графика на рис. 2.15,е.



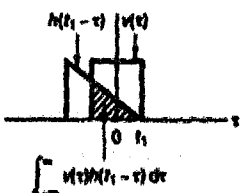
a)



b)

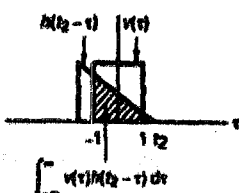


c)



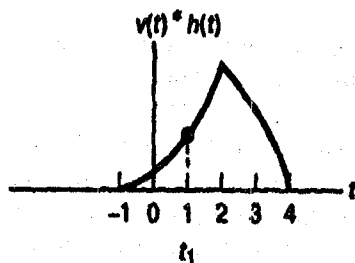
$$\int_{-\infty}^{\infty} v(t)h(t_1 - t) dt$$

r)



$$\int_{-\infty}^{\infty} v(t)h(t_2 - t) dt$$

s)



e)

Puc. 2.15.

ЛЕКЦИЯ 3

3.1. Децибелы

В каждой передающей системе очень важен такой параметр, как интенсивность сигнала. При распространении сигнала по передающей среде может происходить потеря, или затухание, его интенсивности. Помимо этого потери возникают на различных отводах и разветвлениях в среде. Для компенсации таких потерь в различных точках среды могут вводиться усилители, сообщающие сигналу дополнительную интенсивность.

Усиление и снижение интенсивности сигнала, а также его относительные уровни обычно выражаются в децибелах. Ниже перечислены причины такого выбора.

- Интенсивность сигнала часто снижается по логарифмическому закону, поэтому потери проще выражать в децибелах, являющихся логарифмическими единицами.

- Суммарное усиление или ослабление сигнала в каскадном канале передачи можно вычислить с помощью простых операций сложения и вычитания.

Итак, децибелом называется мера отношения между двумя уровнями сигнала:

$$N_{\text{дБ}} = 10 \log_{10} \frac{P_2}{P_1},$$

где $N_{\text{дБ}}$ — число децибел;

P_1 — мощность на входе; P_2 — мощность на выходе;

$\log_{10}$ — логарифм по основанию 10 (вместо записи $\log_{10}$ часто используется форма $\lg$).

Пример

Если на входе линии передачи уровень мощности сигнала составляет 10 мВт, а на некотором расстоянии — 5 мВт, то ослабление сигнала можно выразить следующим образом:

$$N_{\text{дБ}} = 10 \lg(5/10) = 10(-0,3) = -3 \text{ дБ}.$$

Отметим, что в децибелах выражается относительное, а не абсолютное, отличие. Ослабление сигнала с 1000 до 500 мВт также является ослаблением на 3 дБ. Следовательно, ослабление на 3 дБ вдвое снижает уровень мощности; а усиление на 3 дБ удваивает этот уровень.

Децибелы также используются для измерения отношения напряжений, учитывая тот факт, что мощность пропорциональна квадрату напряжения:

$$P = \frac{V^2}{R},$$

где P — мощность, рассеянная на сопротивлении R ;

V — напряжение на сопротивлении R .

Следовательно,

$$N_{\text{дБ}} = 10 \lg \frac{P_2}{P_1} = 10 \lg \frac{V_2^2 / R}{V_1^2 / R} = 20 \lg \frac{V_2}{V_1}.$$

Пример

Использование децибелов полезно при определении усиления или понижения мощности, происходящих на последовательности передающих элементов. Рассмотрим, например, последовательность элементов, на вход которой подается мощность 4 мВт, первый элемент является линией передачи с понижением 12 дБ (—12 дБ), второй элемент — это усилитель с усилением 35 дБ, а третий — линия передачи с затуханием 10 дБ. Суммарное усиление или ослабление равно (—12 + 35 — 10) = 13 дБ. Для вычисления мощности на выходе воспользуемся формулами:

$$13 = 10 \lg (P_2 / 4 \text{ мВт}),$$

$$P_2 = 4 \times 101,3 \text{ мВт} = 79,8 \text{ мВт}.$$

Значения в децибелах относятся к относительным амплитудам или изменениям амплитуд, но никак не к абсолютным уровням. Было бы удобно, если бы мы сумели представить абсолютный уровень мощности также в децибелах, чтобы усиления или понижения мощности по отношению к исходному сигналу могли легко вычисляться. По этой причине получили распространение некоторые производные единицы.

Единица дБВт (децибел-ватт) широко используется в микроволновых приложениях. В качестве эталонного уровня выбрана величина 1 Вт, и ей присвоено значение 0 дБВт. Абсолютный уровень мощности в дБВт определяется следующим образом:

$$\text{мощность, дБВт} = 10 \lg \frac{\text{мощность, Вт}}{1 \text{ Вт}}.$$

Пример

Мощности 1000 Вт соответствует 30 дБВт, а мощности 1 мВт соответствует минус 30 дБВт.

В сферах кабельного телевидения и широковещательных локальных сетях широко используется другая производная единица — дБмВ (децибел-милливольт). Это абсолютная единица, причем 0 дБмВ соответствует 1 мВ. Следовательно,

$$\text{напряжение, дБмВ} = 20 \lg \frac{\text{напряжение, мВ}}{1 \text{ мВ}}.$$

В этом случае предполагается подача напряжения на сопротивление 75 Ом.

Пример

В цифровых системах связи критерием качества связи выступает нормированная к энергии бита версия отношения сигнал/шум $\frac{E_b}{N_0}$,

где E_b – мощность сигнала, умноженная на время передачи бита;

$$N_0 = \frac{N}{W} \text{ – нормированная односторонняя спектральная плотность}$$

мощности шума (мощность шума N на 1 Гц полосы W , измеряемая в Вт/Гц).

Отношение мощности сигнала на приемном конце цифровой системы связи P_r к шуму можно выразить как

$$\frac{P_r}{N_0} = \frac{E_b}{N_0} R \text{ (Гц)},$$

где R – скорость передачи битов, Гц.

Отношение принятой мощности цифрового сигнала к шуму можно также выразить в производных единицах дБГц. Нулю дБГц соответствует 1 Гц:

$$\text{дБГц} = 10 \lg \frac{\frac{P_r}{N_0}, \text{Гц}}{1 \text{ Гц}}$$

3.2. Передача сигнала без искажений по каналу связи

Что нужно, для того чтобы среда передачи вела себя как идеальный канал связи?

1. Сигнал на выходе может запаздывать по отношению к сигналу на входе системы
2. Сигналы на выходе и входе системы могут отличаться по амплитуде

3. Выходной сигнал системы передачи не должен быть искажен

Приведенные три условия соответствуют *линейному каналу передачи*. Если канал передачи линеен, то отклик системы постоянен по модулю, а сдвиг фаз, поступающих на выход спектральных составляющих сигнала должен быть линейным по частоте, т.е. все гармоники сигнала должны поступать на выход с одинаковым запаздыванием, чтобы их можно было просуммировать (групповая задержка должна быть постоянной). В реальных системах при передаче по каналу связи сигнал искажается, и для компенсации этих искажений в приемнике применяются эквалайзеры.

Создать идеальный канал связи нереально, так как для рационального использования ширины полосы бесконечный спектр сигнала усекается до конечной величины с помощью фильтра.

Рассмотрим характеристики идеальных фильтров, представленные на рис. 3.1.

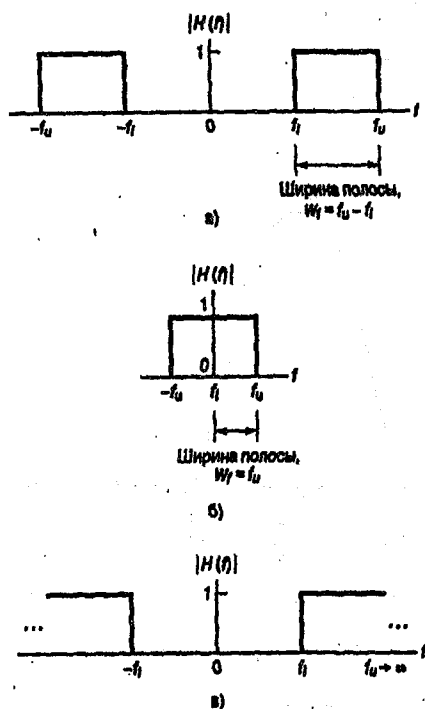


Рис. 3.1. Передаточная функция идеальных фильтров:
а) идеальный пропускающий фильтр; б) идеальный фильтр нижних частот (ФНЧ); в) идеальный фильтр верхних частот (ФВЧ).

Вне диапазона $f_l < f < f_u$, которая называется полосой пропускания (passband), амплитуда отклика идеального фильтра равна нулю. Эффективная ширина полосы пропускания определяется шириной полосы фильтра и составляет $W_f = (f_u - f_l)$ Гц.

Если $f_l \neq 0$; $f_u \neq \infty$, то фильтр называется пропускающим.

Если $f_l = 0$; f_u имеет конечное значение, то фильтр называется фильтром нижних частот.

Если $f_l \neq 0$; $f_u \rightarrow \infty$, то это фильтр верхних частот.

Импульсный отклик идеального фильтра нижних частот, показанный на рис. 3.2, описывается уравнением:

$$h(t) = 2f_u \sin c[2f_u(t - t_0)], \quad (3.1)$$

где функция $\sin c(y) = \frac{\sin(\pi y)}{\pi y}$.

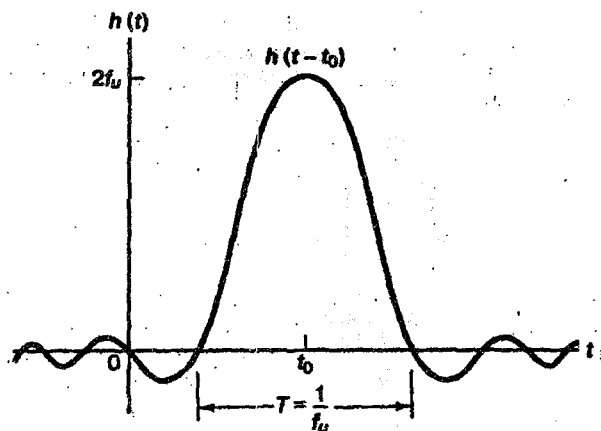


Рис. 3.2. Импульсный отклик идеального фильтра нижних частот.

3.3. Реализуемые фильтры

Простейший реализуемый ФРЧ состоит из сопротивления R и емкости C , как показано на рис. 3.3,а. Этот фильтр называется RC -фильтром и его передаточная функция может быть выражена как

$$H(f) = \frac{1}{1 + 2\pi fRC} = \frac{1}{\sqrt{1 + (2\pi fRC)^2}} e^{-i\theta(f)}, \quad (3.2)$$

где $\theta(f) = \arctg 2\pi fRC$.

Амплитудно-частотная характеристика (АЧХ) $|H(f)|$ и фазо-частотная характеристика $\theta(f)$ изображены на рис. 3.3, б, в. Ширина полосы фильтра нижних частот определяется в точке половинной мощности. Эта точка представляет собой частоту, на которой мощность выходного сигнала равна половине максимального значения, или частоту, на которой амплитуда выходного напряжения равна $\frac{1}{\sqrt{2}}$ максимально-

го значения. В общем случае точка половинной мощности выражается в децибелах как точка -3дБ или точка, находящаяся на 3 дБ ниже максимального значения.

Амплитудный отклик можно выразить в децибелах как

$$|H(f)|_{\text{дБ}} = 20 \lg \frac{V_2}{V_1}, \quad (3.3)$$

где $V_1; V_2$ – напряжения на входе и выходе, а сопротивления на входе и выходе предполагаются равными.

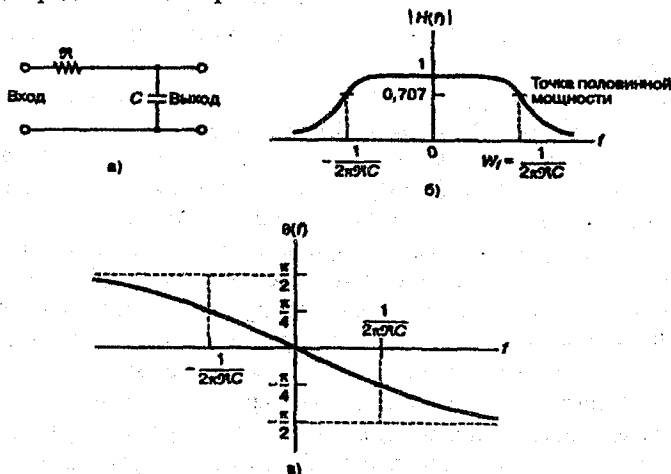


Рис. 3.3. RC-фильтр и его передаточная функция:
а) RC-фильтр б) АЧХ в) ФЧХ.

Точка половинной мощности RC -фильтра нижних частот соответствует $\omega = 1/RC$ рад/с или $f = 1/2\pi RC$ Гц. Таким образом, ширина полосы W_f в герцах равна $1/2\pi RC$.

Форм-фактор фильтра – это мера того, насколько хорошо реальный фильтр аппроксимирует идеальный. Обычно он определяется как отношение ширины полос фильтров по уровню -60 дБ и -6дБ. Достаточно малый форм-фактор (около 2) можно получить в пропускающем фильтре с очень резким срезом. Форм-фактор простого RC -фильтра нижних частот составляет около 600.

Существует несколько полезных аппроксимаций характеристики идеального ФНЧ. Одну из них дает фильтр Баттерворта, аппроксимирующий идеальный ФНЧ функцией:

$$|H_n(f)| = \frac{1}{\sqrt{1 + \left(\frac{f_{\text{сигн}}}{f_u}\right)^{2n}}} \quad n \geq 1, \quad (3.4)$$

где f_u – верхняя частота среза (по уровню -3дБ); n – порядок фильтра.

Чем выше порядок, тем выше сложность и стоимость реализации фильтра. На рис. 3.3 показаны графики $|H(f)|$ для нескольких значений n . По мере роста n АЧХ приближается к характеристикам идеального фильтра.

Фазовый сдвиг и крутизна спада амплитудно-частотной характеристики (АЧХ) за пределами полосы пропускания фильтра определяются его порядком и составляют 90 градусов и 6 дБ/октава на каждый порядок. То есть фильтр первого порядка обеспечивает затухание 6дБ/октава при полном фазовом сдвиге 90 градусов, фильтр второго порядка – 180 градусов и 12 дБ/октава и так далее. На частоте среза затухание фильтра составляет 3 дБ, а фазовый сдвиг – половину от полного (т.е. 45 градусов для фильтра 1-го порядка и 90 – для 2-го). От типа фильтра зависит только плавность изгиба частотной характеристики в области частоты среза и суммарная АЧХ системы, а также фазовые характеристики.

В промышленных конструкциях наибольшее распространение получили построенные на повторителях фильтры Баттерворта, Бесселя и Саллена-Ки. Как правило, используются фильтры 2-го порядка. Каждый из перечисленных типов имеет свои достоинства и недостатки.

Фильтры Бесселя обладают самой гладкой фазовой характеристикой (как у одиночной RC -цепи), но суммарная АЧХ имеет провал величиной 3 дБ на частоте раздела. Фильтры Баттерворта обеспечивают плоскую суммарную АЧХ, но их фазовая характеристика более крутая. Наконец, фильтры Саллена-Ки (равнокомпонентные фильтры) очень удобны в серийном производстве, поскольку (как следует из названия) для них требуются детали одинаковых номиналов и с большим допустимым отклонением, чего нельзя сказать о фильтрах Баттерворта и Бесселя, требующих точных деталей. Однако фазовая и частотная характеристики равнокомпонентных фильтров самые худшие, поэтому их используют только в бюджетных моделях.

Любой фильтр (и активный, и пассивный) использует реактивные элементы — емкости и индуктивности, поэтому вносит в сигнал фазовые и временные искажения. Фильтры НЧ (Low Pass) вносят в сигнал запаздывание и отставание по фазе, которое можно в определенной мере скорректировать фазовращателем. При использовании фильтра Бесселя 2-го порядка в сочетании с таким фазовым корректором можно получить фильтр с абсолютно линейной фазовой характеристикой. Что же касается ФВЧ (High Pass), то они формируют фазовое опережение, которое принципиально невозможно состыковать с имеющимся ФНЧ. Однако в этом случае можно воспользоваться для формирования сигнала полосы высших частот фильтром дополнительной функции (ФДФ). Выходной сигнал такого фильтра получается путем вычитания из входного сигнала его части, прошедшей через ФНЧ. Очевидно, что в этом случае происходит компенсация фазовых искажений и разность фаз сигналов на выходе ФНЧ и ФДФ сохраняется постоянной во всей полосе частот. Однако у фильтров дополнительной функции есть существенный недостаток — крутизна спада АЧХ составляет только 6 дБ/октава, что может иногда оказаться недостаточным.

Для перестройки частоты среза в активном фильтре необходимо синхронно изменять величины частотоподающих звеньев. Для плавной регулировки частоты среза используют потенциометры. Нетрудно подсчитать, что для перестройки фильтра 2-го порядка требуется четырехсекционный потенциометр (для двух каналов). С целью удешевления в последнее время в бюджетных моделях усилителей все чаще используют упрощенные фильтры 2-го порядка, в которых перестраивается по частоте только одно звено. Такие фильтры нельзя отнести к какому-либо конкретному типу, потому что "идеологически выдержанный" фильтр получается только в одном из крайних положений регулятора.

3.4. Дискретизация аналоговой информации

Аналоговый сигнал и его дискретная версия связаны процессом дискретизации (sampling process). Рассмотрим способы дискретизации аналогового сигнала.

1. Естественная дискретизация. Ее можно рассматривать как включение и выключение коммутатора в определенные моменты времени. В этом случае вершина каждого импульса в течение интервала его передачи имеет форму соответствующего аналогового сегмента (рис. 3.4).

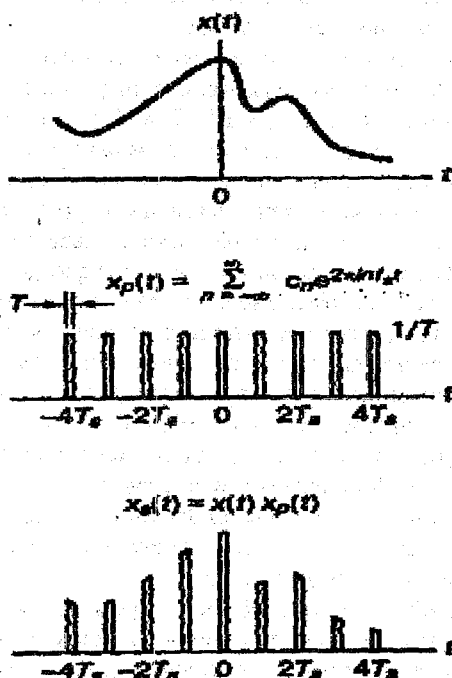


Рис. 3.4. Процесс естественной дискретизации аналогового сигнала.

2. «Выборка-хранение» (sample-and-hold). Умножаем исходный сигнал на последовательность дельта-функций и проделываем операцию свертки с прямоугольным импульсом $p(t)$, имеющим единичную амплитуду и ширину, равную T_s (рис. 3.5).

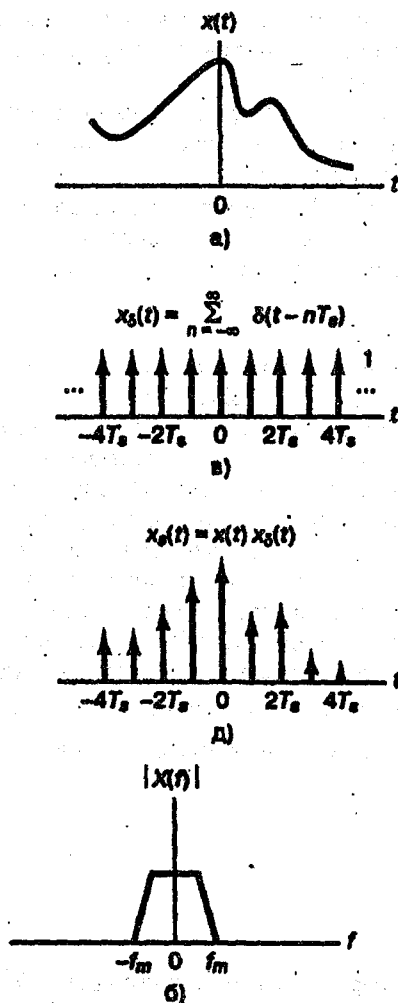


Рис. 3.5. Процесс дискретизации «выборка-хранение»
а) Аналоговый сигнал; б) Фурье-образ аналогового сигнала
в) Единичные импульсы; г) Свертка налогового сигнала с последовательностью единичных импульсов.

Преимущества реализации процесса дискретизации методом «выборка-хранение»:

- Выборка-хранение дает дискретную последовательность импульсов с плоским верхом и затухание высокочастотных спектральных копий, что весьма желательно.

- После операции выборки-хранения требуется дополнительная фильтрация, позволяющая подавить остаточные спектральные компоненты, кратные частоте дискретизации.

- За счет выбора функции $p(t)$ можно получить неоднородное усиление (подавление) спектра нужной полосы частот, а после фильтрации это подавление можно компенсировать путем применения функции, обратной $p(t)$.

Результатом процесса дискретизации является сигнал в амплитудно-импульсной модуляции (pulse-amplitude modulation — PAM). Такое название возникло потому, что выходящий сигнал можно описать как последовательность импульсов с амплитудами, определяемыми выборками входящего сигнала. Аналоговый сигнал можно восстановить (с определенной степенью точности) из модулированного сигнала путем прохождения последнего через фильтр нижних частот. Важно знать, насколько точно отфильтрованный модулированный сигнал совпадает с исходным аналоговым сигналом? Ответ на этот вопрос дает теорема о дискретном представлении (sampling theorem), которая формулируется следующим образом: сигнал с ограниченной полосой, не имеющий спектральных компонентов с частотами, которые превышают f_m Гц, однозначно определяется значениями, выбранными через равные промежутки времени:

$$T_s \leq \frac{1}{2f_m} \text{ с.} \quad (3.5)$$

При другой формулировке верхний предел T_s можно выразить через частоту дискретизации (sampling rate) $f_s = \frac{1}{T_s}$. В этом случае получаем ограничение, именуемое критерием Найквиста или теоремой Котельникова:

$$f_s \geq 2f_m. \quad (3.6)$$

Частота дискретизации $f_s = 2f_m$ также называется частотой Найквиста. Критерий Найквиста — это теоретическое достаточное условие, которое делает возможным полное восстановление аналогового сигнала из последовательности равномерно распределенных дискретных выборок.

После выбора частоты дискретизации каждая спектральная копия отделяется от соседних полосой частот, равной f_s Гц, и аналоговый сигнал полностью восстанавливается из выборок путем фильтрации. В то же время для выполнения этого потребовался бы идеальный фильтр с абсолютно крутыми фронтами. Очевидно, что если $f_s > 2f_m$, копии отдалятся (в частотной области), как показано на рис. 3.5,а, и это облегчит операцию фильтрации. На рисунке также показана типичная характеристика фильтра нижних частот, который может использоваться для выделения спектра немодулированного сигнала. При уменьшении частоты дискретизации до $f_s < 2f_m$ копии начнут перекрываться, как показано на рис. 3.5,б, и информация частично будет потеряна. Явление, являющееся результатом недостаточной частоты выборки (выборки, производимой очень редко), называется наложением (aliasing). Частота Найквиста $f_s = 2f_m$ — это предел, ниже которого происходит наложение; чтобы избежать этого нежелательного явления, следует удовлетворять критерий Найквиста $f_s \geq 2f_m$.

Повышение частоты дискретизации позволяет устранить наложение путем разделения спектральных копий. Существуют специальные фильтры для защиты от наложения (antialiasing filters), фильтрующие как аналоговый сигнал, так и сигнал после дискретизации. Но реализуемые фильтры требуют нулевой полосы для перехода между полосой пропускания и полосой затухания. Но при минимизации полосы перехода резко возрастает сложность фильтров и их стоимость, поэтому оптимальной полосой перехода считается 10-20 % от ширины полосы сигнала. Рассчитав частоту дискретизации Найквиста для 20 % ширины перехода фильтра защиты от наложения спектров, получим инженерную версию критерия Найквиста:

$$f_s \geq 2,2f_m. \quad (3.7)$$

Задача

Сигнал $u(t)$, дискретизованный в соответствии с теоремой Котельникова, имеет два ненулевых отсчета, изображенных на рис. 3.6.

Вычислите мгновенное значение исходного аналогового сигнала в момент времени $t=1$ мкс.

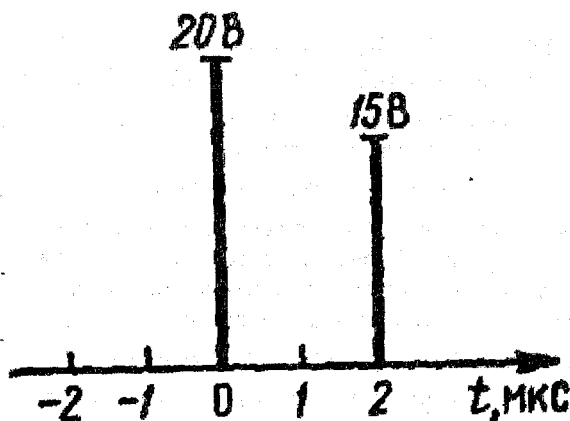


Рис. 3.6. Выборки дискретизованного сигнала.

Решение

В соответствии с рисунком определяем интервал дискретизации $t_0 = 2 \cdot 10^{-6}$ с, откуда находим верхнюю частоту в спектре сигнала

$$\omega_b = \frac{\pi}{t_0} = 1,57 \cdot 10^6 \text{ рад/с.}$$

Ряд Котельникова рассматриваемого сигнала имеет вид:

$$s(t) = 20 \frac{\sin \omega_b t}{\omega_b t} + 15 \frac{\sin(\omega_b t - \pi)}{\omega_b t - \pi},$$

из которого при $t=10^{-6}$ с получаем $s(10^{-6}) = 22,2$ В.

Задача

Дан аналоговый сигнал, который выбирался с частотой Найквиста f_s посредством естественной дискретизации. Докажите, что сигнал, пропорциональный исходному сигналу, может быть восстановлен из

выборок с использованием метода, показанного на рис. 3.7. Параметр mf_s - это частота локального осциллятора, причем m - целое.

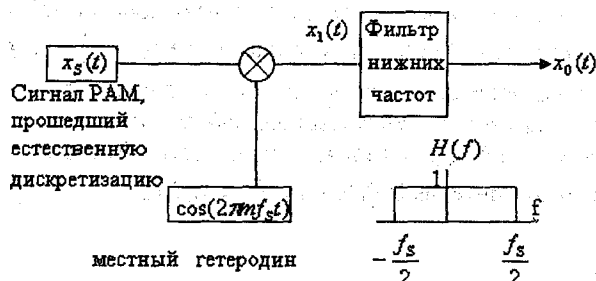


Рис. 3.7.

Решение

По свойству трансляции частоты преобразования Фурье на выходе гетеродина получаем спектр исходного сигнала, периодически повторяющийся по частоте с интервалом f_s Гц (рис. 3.8). ФНЧ затем вырезает нужный спектр, находящийся в частотном интервале $[-\frac{f_s}{2}, \frac{f_s}{2}]$.

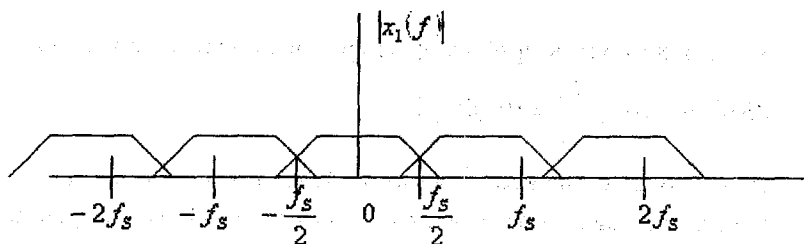


Рис. 3.8.

Так как частота Найквиста равна удвоенной верхней частоте в спектре сигнала, то сигнал, пропорциональный исходному сигналу, может быть восстановлен.

В задаче предполагается идеальная передаточная характеристика фильтра. Однако реализуемые узкополосные фильтры не имеют строго

ограниченной полосы, поэтому фильтруемые спектры всегда включают некоторое наложение (как это показано на рис. 3.8). Однако всегда можно определить полосу частот, вне которой спектральные компоненты сигнала заглушают настолько, что ими можно пренебречь. Наложение спектров присутствует и при частоте дискретизации $f_s < 2f_m$, где f_m — максимальная частота в спектре сигнала. Копии спектра немодулированного сигнала начнут перекрываться, как это показано на рис. 3.8 и часть информации будет потеряно. Для расчета ФНЧ используется инженерная версия критерия Найквиста $f_s < 2,2f_m$.

Задача

Наложение не происходит, если частота дискретизации больше удвоенной ширины полосы сигнала. В то же время сигналов со строго ограниченной полосой не существует. Таким образом, наложение присутствует всегда.

а) Предположим, что фильтрованный сигнал имеет спектр, который описывается фильтром Баттерворта 6-го порядка и верхней частотой среза $f_u = 1000$ Гц. Какая частота дискретизации необходима для снижения наложения до точки -50 дБ в спектре мощностей?

б) Повторите пункт «а» для фильтра Баттерворта 12-го порядка.

Решение

Амплитудная характеристика фильтра в децибелах выражается как

$$|H(f)| \text{ дБ} = 10 \lg \frac{P_2}{P_1} = 10 \lg |H(f)|,$$

где $P_1; P_2$ — мощности на входе и выходе фильтра соответственно.

Переведем численное значение амплитудной характеристики фильтра в точке снижения наложения в разы

$$10 \lg |H(f)| = -50$$

$$\lg |H(f)| = -\frac{50}{10} = -5 \Rightarrow |H(f)| = 10^{-5} = 0,00001.$$

Фильтр Баттерворта аппроксимирует идеальный фильтр нижних частот функцией

$$|H_n(f)| = \frac{1}{\sqrt{1 + \left(\frac{f_{\text{сигн}}}{f_u}\right)^{2n}}}$$

где f_u - верхняя частота среза (по уровню -3дБ); n - порядок фильтра.

1) Для порядка фильтра Баттерворта $n = 6$ имеем

$$|H_n(f)| = \frac{1}{\sqrt{1 + \left(\frac{f_{\text{сигн}}}{f_u}\right)^{12}}} = 0,00001.$$

Решив это уравнение, получим, что

$$\text{соотношение } \frac{f_{\text{сигн}}}{f_u} = 6,8 \Rightarrow f_{\text{сигн}} = 1000 \cdot 6,8 = 6800 \text{ Гц.}$$

Частота дискретизаций по теореме Найквиста должна быть не менее удвоенной верхней частоты в спектре сигнала. Получаем

$$f_{\text{дискр}} = f_s = 2f_{\text{сигн}} = 13600 \text{ Гц.}$$

2) Для порядка фильтра Баттерворта $n = 12$ имеем

$$|H_n(f)| = \frac{1}{\sqrt{1 + \left(\frac{f_{\text{сигн}}}{f_u}\right)^{24}}} = 0,00001$$

Решая это уравнение, получим, что

$$\text{соотношение } \frac{f_{\text{сигн}}}{f_u} = 2,6 \Rightarrow f_{\text{сигн}} = 1000 \cdot 2,6 = 2600 \text{ Гц}$$

Частота дискретизации по теореме Найквиста должна быть не менее удвоенной верхней частоты в спектре сигнала. Получаем

$$f_{\text{дискр}} = f_s = 2f_{\text{сигн}} = 5200 \text{ Гц.}$$

ЛЕКЦИЯ 4

4.1. Цифровые фильтры

Цифровыми называются фильтры, реализованные на основе методов цифровой обработки сигналов. Сначала рассмотрим фильтр с конечной импульсной характеристикой (*фильтр КИХ* – finite impulse response – FIR), схема работы которого представлена на рис. 4.1. Первый отсчет (назовем его g_0) входной числовой последовательности умножается на коэффициент h_0 и появляется на выходе фильтра в виде отсчета k_0 .

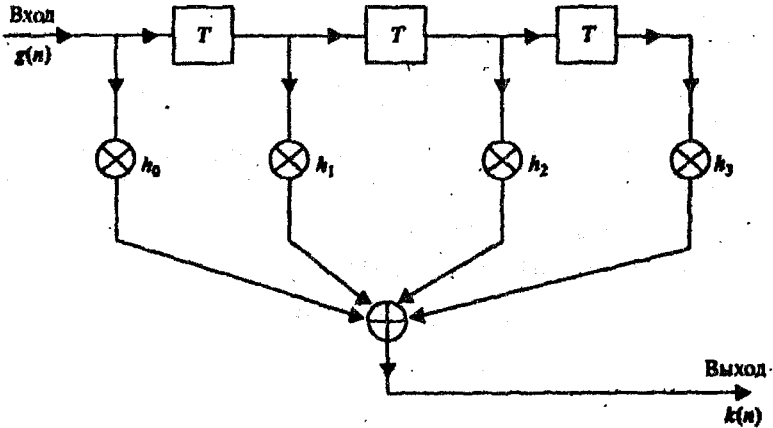


Рис. 4.1. Фильтр КИХ.

Плечи $h_0 - h_3$ называются ветвями (tap). В фильтре присутствует несколько линий задержки, каждая из которых задерживает сигнал на время T , что соответствует периоду цикла выборки сигнала в приемнике. Входной импульс g_0 достигает ветви h_1 после задержки на один период, в это время на входе фильтра появляется следующий импульс g_1 . Следовательно, в момент времени $t=0$ выходной сигнал k_0 равен произведению $g_0 h_0$. В момент времени $t=T$ на выходе фильтра появляется новый сигнал k_1 , который равен произведению $g_0 h_1$ плюс новый входной импульс $g_1 h_0$. Функцию $k(n)$ можно найти с помощью свертки функций $g(n)$ и $h(n)$, она называется сверточной суммой.

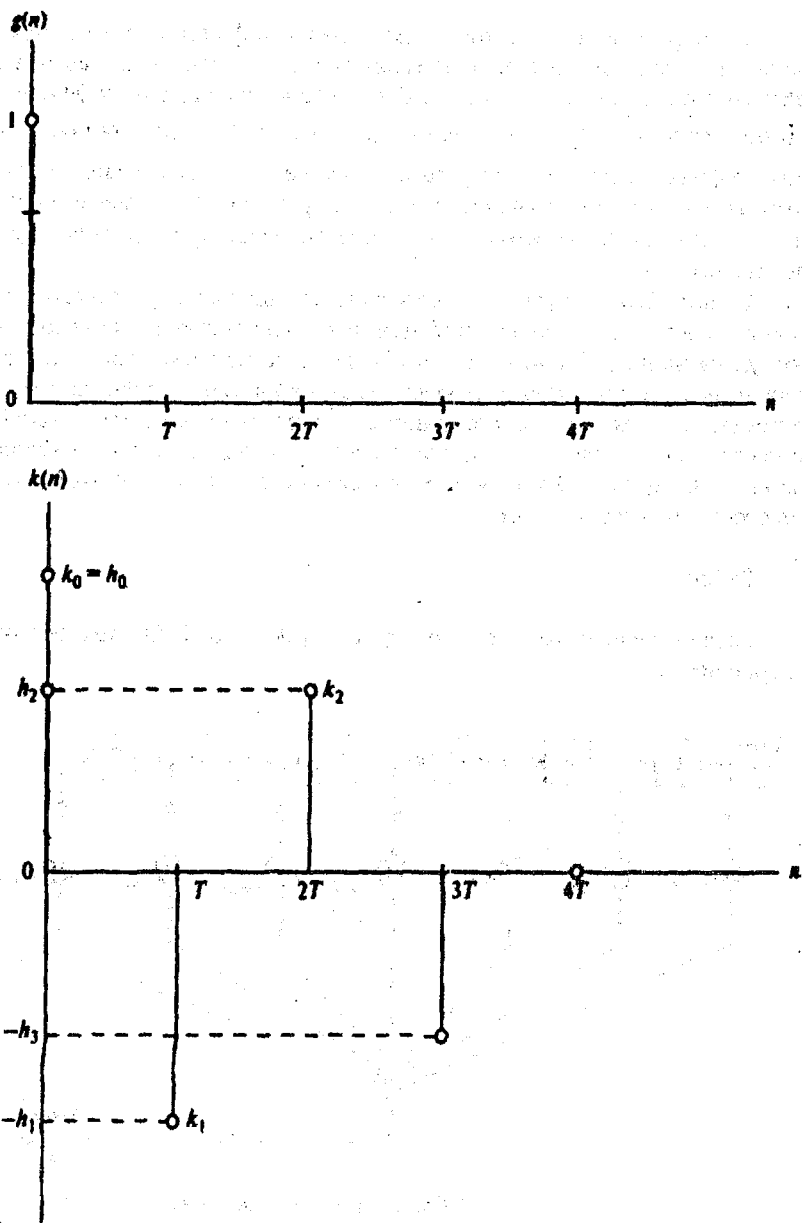


Рис. 4.2. Импульсная характеристика фильтра КИХ.

Для определения частотной характеристики фильтра, нужно подать на его вход сигнал с постоянной амплитудой и бесконечным частотным диапазоном, т.е. дельта-функцию. Если предположить, что коэффициенты передачи h_1 и h_3 отрицательны (см. рис. 4.1), тогда при поступлении на фильтр дельта-функции на его выходе появится сигнал, показанный на рис. 4.2. По истечении интервала времени $3T$ с момента подачи на вход исходного импульса сигнал на выходе фильтра вернется в нулевое положение.

Фильтр КИХ получил свое название благодаря тому, что его входной сигнал исчезает в течение N отсчетов после поступления исходного импульса на вход фильтра. Здесь N – количество элементов задержки, или ветвей фильтра. Для построения фильтрующего устройства необходима система, которая может умножать (с разными весовыми коэффициентами для каждой ветви), складывать и последовательно сохранять входные импульсы. Это можно реализовать с помощью цифрового компьютера и микропроцессора.

Задача

Нарисуйте импульсную характеристику фильтра КИХ, показанного на рисунке 4.3.

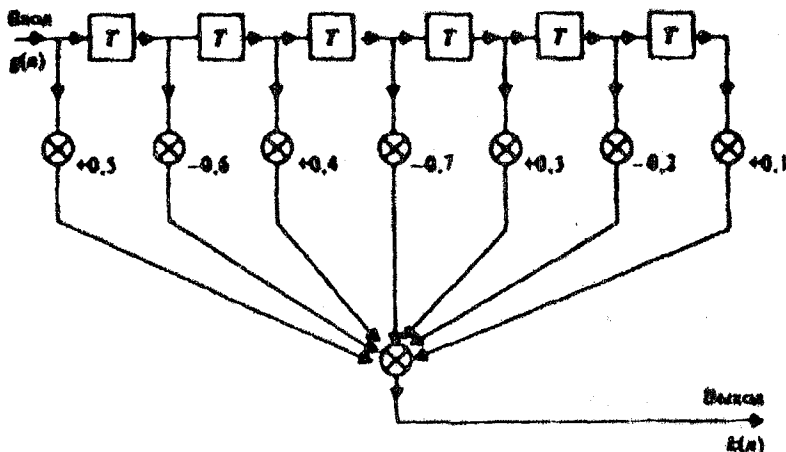


Рис. 4.3. Структура фильтра КИХ.

Решение

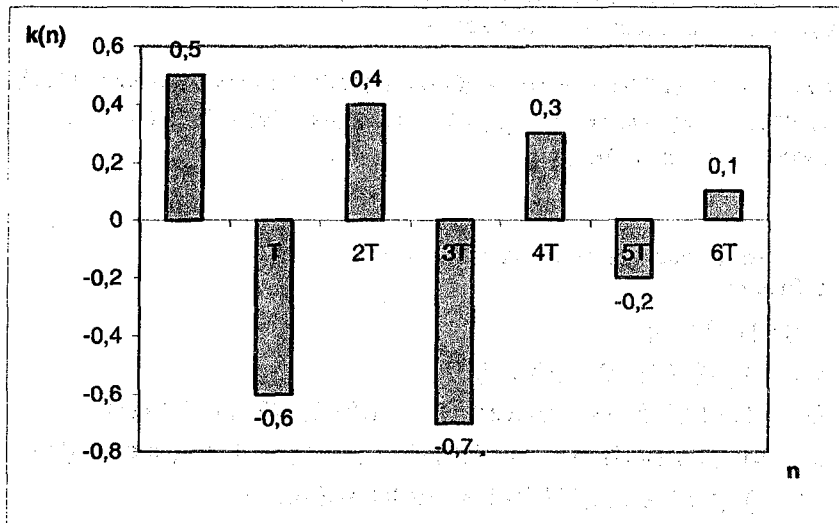


Рис. 4.4. Импульсная характеристика фильтра КИХ.

На рис. 4.5 изображена схема работы второго типа цифрового фильтра – фильтра с бесконечной импульсной характеристикой или фильтра БИХ (finite impulse response – FIR), содержащая три ветви. Ветви в фильтре БИХ являются ветвями обратной связи (feed back) с выхода.

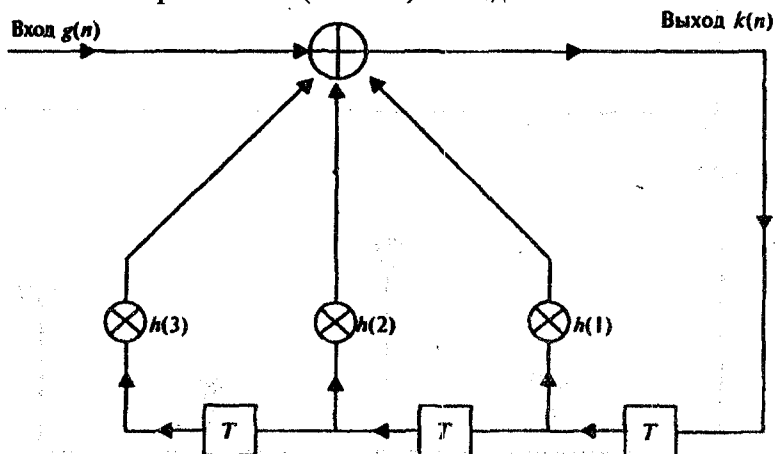


Рис. 4.5. Фильтр с бесконечной импульсной характеристикой.

[illegible]

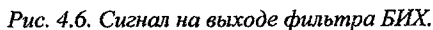
Этот сигнал приведен на рис. 4.6, где

$$k(1) = [g(0)h(1);$$

$$k(2) = [g(0)h(1)]h(1) + [g(0)]h(2);$$

$$k(3) = [g(0)h(1)h(1) + g(0)h(2)]h(1) + [g(0)h(1)]h(2) + [g(0)]h(3);$$

$$k(4) = [g(0)h(1)h(1)h(1) + g(0)h(2)h(1) + g(0)h(2)h(1) + g(0)h(3)]h(1) + \\ + [g(0)h(1)h(1) + g(0)h(2)]h(2) + [g(0)h(1)]h(3) \text{ и т.д.}$$



Для того чтобы создать цифровой фильтр по заданным параметрам (частоте среза и крутизне АЧХ), можно взять аналоговую функцию с известными характеристиками, например полином Чебышева n -порядка. Зная его функцию передачи, можно математически (используя, скажем, z -преобразование) определить функцию передачи эквивалентного цифрового фильтра, которая будет иметь вид числовой последовательности. Теперь несложно нарисовать схему фильтра, которую можно реализовать как на логических элементах, так и программно. Однако такой подход к моделированию электронных систем связан с утомительными математическими расчетами и подвержен ошибкам. Поэтому на практике обычно применяют один из многочисленных пакетов по компьютерному моделированию систем цифровой обработки сигналов, которые сами определяют количество ветвей и их весовые характеристики по заданным техническим требованиям. При более детальном рассмотрении фильтров с БИХ можно сделать вывод, что при заданной частоте среза они имеют меньшее количество ветвей, чем фильтры с КИХ. Зато фильтры с КИХ превосходят фильтры с БИХ по качеству фазовой характеристики.

Задача

Изобразите графически импульсную характеристику фильтра с БИХ, показанную на рис. 4.7.

Решение

$$k(0) = g(0) = 1$$

$$k(1) = 1 * (-0,2) = -0,2$$

$$k(2) = -0,2 * (-0,2) + 1 * 0,3 = 0,34$$

$$k(3) = 0,34 * (-0,2) + 0,3 * (-0,2) - 0,4 = -0,068 - 0,06 - 0,4 = -0,528 \approx -0,53$$

$$k(4) = -0,53 * (-0,2) + (0,04 + 0,3) * 0,3 + (-0,2) * (-0,4) = 0,106 + 0,102 + 0,08 = 0,288$$

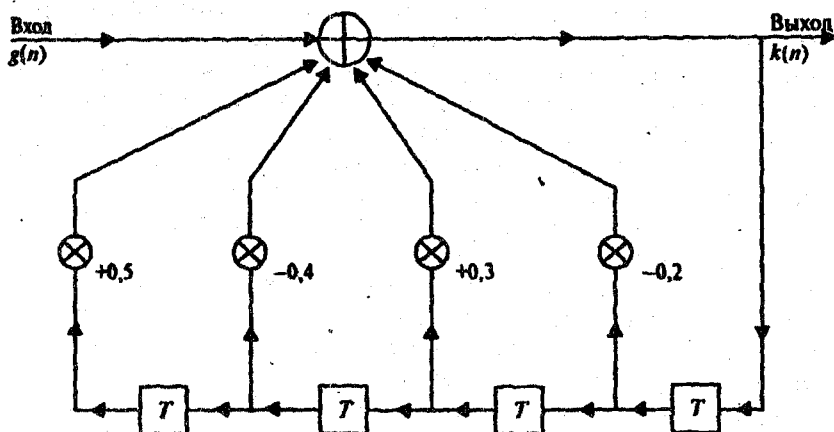


Рис. 4.7. Схема фильтра с БИХ.

По результатам расчета импульсной характеристики имеем импульсную характеристику, показанную графически на рис. 4.8.

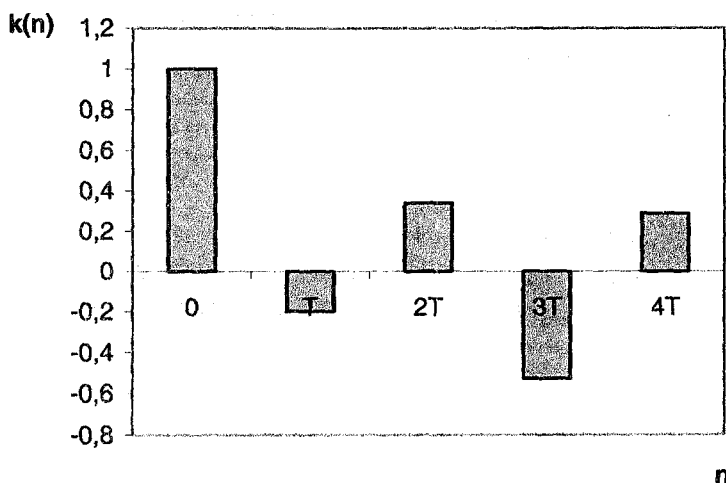


Рис. 4.8. Импульсная Характеристика фильтра БИХ.

4.2. Выборка с запасом

Выборка с запасом (oversampling) — это наиболее экономичное решение задачи преобразования аналогового сигнала в цифровой или цифрового в аналоговый. Это объясняется тем, что обработка сигнала выполняется на высокопроизводительном аналоговом оборудовании, что обычно дороже использования для этой же цели цифрового оборудования обработки сигналов. Рассмотрим преобразование аналоговых сигналов в цифровые. Если это выполняется без выборки с запасом, то процесс дискретизации описывается тремя простыми этапами.

Выборка без запаса

1. Сигнал пропускается через высокопроизводительный аналоговый фильтр нижних частот для ограничения его полосы.

2. Отфильтрованный сигнал дискретизируется с частотой Найквиста с целью создания сигнала с (приблизительно) ограниченной полосой, так как сигнал со строго ограниченной полосой относится к разряду нереализуемых.

3. Выборки квантуются устройством преобразования аналоговых сигналов в цифровые, отображающим выборки, которые могут принимать значения из непрерывного диапазона, в конечный набор дискретных уровней.

Если же выборку производить с запасом, то процесс будет состоять из пяти этапов.

Выборка с запасом

1. Сигнал пропускается через менее производительный (более дешевый) аналоговый фильтр нижних частот (предварительная фильтрация) для ограничения его полосы.

2. Предварительно отфильтрованный сигнал выбирается с частотой в несколько раз выше частоты Найквиста для создания сигнала с ограниченной полосой.

3. Выборки преобразовываются преобразователем аналоговых сигналов в цифровые, отображающим выборки, которые могут принимать значения из непрерывного диапазона, в конечный набор дискретных уровней.

4. Цифровые выборки обрабатываются высокопроизводительным цифровым фильтром для сужения полосы цифровых выборок.

5. Частота дискретизации на выходе цифрового фильтра уменьшается пропорционально сужению полосы, полученному при использовании этого цифрового фильтра.

Полоса пропускания аналогового фильтра, ограничивающая ширину полосы входящего сигнала, равна ширине полосы сигнала плюс область спада. Наличие области перехода приводит к увеличению ширины полосы сигнала на выходе на некоторую величину f_t . Частоту Найквиста f_s для отфильтрованного выхода, обычно равную $2 f_m$ (удвоенной максимальной частоте дискретного сигнала), теперь необходимо увеличить до $2 f_m + f_t$. Ширина полосы спада фильтра является служебными издержками процесса дискретизации. Этот дополнительный спектральный интервал не представляет собой полосу полезного сигнала, а нужен для защиты полосы сигнала путем резервирования спектральной области для накладываются спектра, возникающего в процессе дискретизации. Наложение возникает вследствие того, что реальный сигнал не может быть строго ограниченным. Типичные полосы спада дают 10-20%-ное увеличение частоты дискретизации по сравнению с частотой, определяемой критерием Найквиста. Примером таких служебных издержек может служить цифровая аудиосистема проигрывания компакт-дисков, где двусторонняя полоса равна 40 кГц, а частота дискретизаций – 44,1 кГц, или система проигрывания цифровых аудиокассет, в которой ширина двусторонней полосы также равна 40 кГц, а частота дискретизации — 48,0 кГц.

Естественным желанием является использование аналоговых фильтров с узкой полосой перехода и максимально низкой из возможных частот дискретизации. В то же время аналоговые фильтры имеют две нежелательные особенности. Во-первых, они могут приводить к искажению (нелинейное изменение фазы с частотой), вызванному малыми областями перехода. Во-вторых, цена системы может оказаться высокой, поскольку узкие области перехода подразумевают применение фильтров высоких порядков, требующих большого числа высококачественных составляющих.

Проблема состоит в том, что для уменьшения стоимости хранения данных хотелось бы работать с устройством дискретизации с максимально низкой частотой. Для достижения этой цели можно создать изоэдренный аналоговый фильтр с узкой областью перехода. Однако такой фильтр не только дорог, но и искажает сам сигнал, хотя задачей фильтра как раз является защита сигнала (от нежелательного наложения).

В данном случае выборка с запасом наиболее приемлема – при наличии проблемы, решить которую мы не можем, превращаем ее в проблему, поддающуюся решению. Мы используем дешевый, менее сложный предварительный аналоговый фильтр для ограничения полосы

входящего сигнала. Этот аналоговый фильтр можно упростить за счет выбора более широкой переходной области. При этом увеличивается ширина спектра, из-за чего нам нужно увеличить требуемую частоту дискретизации. Обычно начинают с выбора частоты дискретизации, в 4 раза превышающей исходную, после чего разрабатывают аналоговый фильтр, ширина полосы которого соответствует этой увеличенной частоте дискретизации. Например, вместо дискретизации сигнала компакт-диска на частоте 44,1 кГц при ширине области перехода 4,1 кГц, реализованной с использованием сложнейшего эллиптического фильтра 10-го порядка (подразумевается, что фильтр включает 10 избирательных элементов, таких, как конденсаторы и индуктивности), мы выбираем выборку с запасом. В этом случае устройство дискретизации может работать на частоте 176,4 кГц с областью перехода 136,4 кГц, реализованное простым эллиптическим фильтром 4-го порядка (имеющим всего 4 избирательных элемента).

Итак, у нас есть дискретные данные с большей, чем требуется, частотой дискретизации, и эти данные пропускаются через недорогой высокопроизводительный цифровой фильтр для выполнения фильтрации, необходимой для предотвращения наложения. Цифровой фильтр может реализовать узкую область перехода без искажения, свойственного аналоговым фильтрам, а его эксплуатация недорогая. После того как цифровая фильтрация уменьшила ширину полосы перехода, мы снижаем частоту дискретизации сигнала (повторная выборка). В результате в единую структуру объединяются качественные методы цифровой обработки, фильтрация и повторная выборка.

Рассмотрим теперь вопрос дальнейшего улучшения качества процесса сбора данных. Предварительный аналоговый фильтр приводит к некоторому искажению амплитуды и фазы. Поскольку заранее известно, каково это искажение, цифровой фильтр проектируется не только для защиты (совместно с аналоговым фильтром) от наложения, но и для компенсации усиления и искажения фазы, вносимых аналоговым фильтром. Суммарный результат может, по желанию, улучшаться до любого предела. Таким образом, получаем сигнал более высокого качества (менее искаженный) по более низкой цене. Аппаратура цифровой обработки сигналов, представляющая собой развитие компьютерной индустрии, характеризуется значительным ежегодным снижением цен, чего нельзя сказать об аналоговой аппаратуре.

Подобным образом выборка с запасом используется в процессе преобразования цифрового сигнала в аналоговый. Аналоговый фильтр, через который пропускается преобразованный сигнал, будет искажать сигнал, если последний будет иметь узкую полосу перехода. Но полоса

перехода уже не будет узкой, если данные, полученные после цифро-аналогового преобразования, были оцифрованы с помощью выборки с запасом.

4.3. Устойчивость линейных дискретных систем (ЛДС)

Соотношение вход – выход в ЛДС описывается линейным уравнением. Взаимосвязь между x и y ЛДС, приведенная на рис. 4.9, описывается разностным уравнением (РУ) вида:

$$y(n) = \sum_{i=0}^{N-1} b_i x[(n-i)T] - \sum_{k=1}^{M-1} a_k y[(n-k)T], \quad (4.2)$$

где $b_i; a_k$ – коэффициенты уравнения (вещественные константы); $x[(n-i)T]; y[(n-k)T]$ – воздействие и реакция на это воздействие, задержанные на i и k периодов соответственно.

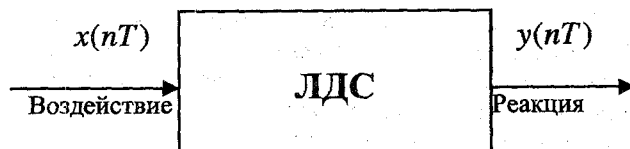


Рис. 4.9. Схема линейной дискретной системы.

ЛДС называется рекурсивной, если хотя бы один из коэффициентов a_k не равен нулю. Порядком рекурсивной ЛДС называют порядок РУ.

Реакция $y(n)$ рекурсивной ЛДС в каждый момент времени определяется:

- Текущим отсчетом воздействия $x(n)$
- Предысторией воздействия $x(n-i)$, $i=1,2,\dots,N-1$
- Предысторией реакции $y(n-k)$, $k=1,2,\dots,M-1$

Импульсная характеристика рекурсивной ЛДС имеет бесконечную длительность (БИХ-система).

Структурная схема рекурсивной ЛДС 2-го порядка состоит из элементов задержки, умножителей на константу и сумматора (см. рис. 4.10).

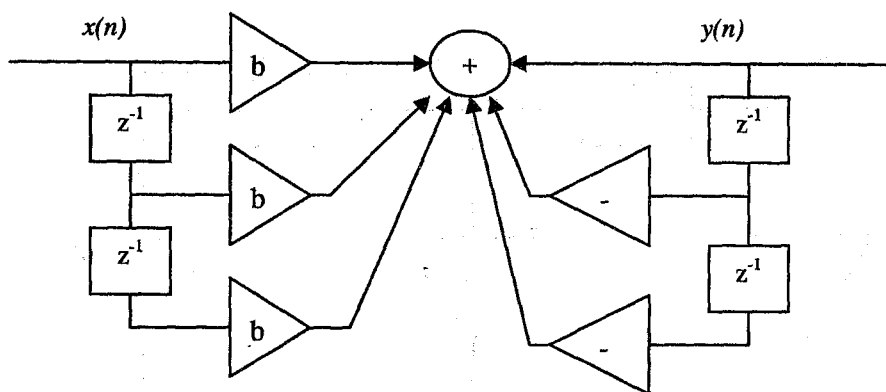


Рис. 4.10. Структурная схема рекурсивной ЛДС 2-го порядка.

ЛДС называется *нерекурсивной*, если все коэффициенты a_k РУ равны нулю.

Реакция *нерекурсивной* ЛДС в каждый момент времени определяется:

- Текущим отсчетом воздействия $x(n)$
- Предысторией воздействия $x(n-i)$

Импульсная характеристика *нерекурсивной* ЛДС имеет конечную длительность (КИХ-система) (см. рис. 4.11).

Оценка устойчивости ЛДС базируется на утверждении, что ограниченное воздействие должно приводить к ограниченной реакции и может проводиться:

- по импульсной характеристике – *нерекурсивные* системы всегда устойчивы. БИХ-системы требуют анализа на устойчивость;
- по Z-изображению импульсной характеристики.

Передаточная функция (ПФ) ЛДС

$$\text{Для звена 2-го порядка } H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}. \quad (4.3)$$

Нулями (0) ПФ называют значения z , при которых $H(z)=0$.

Особыми точками (*), полюсами ПФ называют значения z , в которых знаменатель ПФ равен нулю.

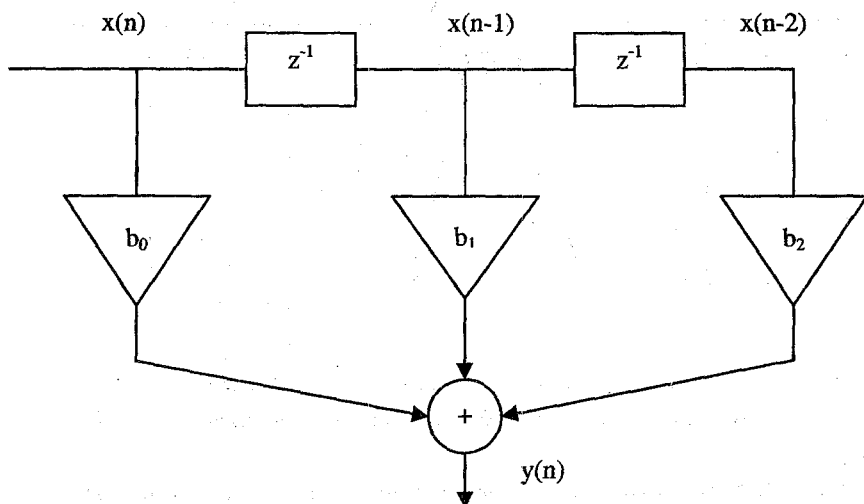


Рис. 4.11. Структурная схема нерекурсивной ЛДС 2-го порядка.

Комплексно-сопряженные полюсы находятся в виде:

$$z_{*1,2} = r_* e^{\pm j\phi_*}, \quad (4.4.)$$

где радиус $r_* = \sqrt{a_2}$;

$$\text{угол } \phi_* = \arccos\left(-\frac{a_1}{2r_*}\right).$$

Вещественные или комплексно-сопряженные нули находятся в виде:

$$z_{01,2} = \frac{-b_1 \pm \sqrt{b_1^2 - 4b_0b_2}}{2b_0} = -\frac{b_1}{2b_0} \pm \frac{\sqrt{b_1^2 - 4b_0b_2}}{2b_0} = \xi_0 \pm j\eta_0 \quad (4.5)$$

или в показательной форме:

$$z_{01,2} = r_0 e^{\pm j\phi_0}, \quad (4.6)$$

где радиус $r_0 = \sqrt{\xi_0^2 + \eta_0^2}$;

$$\text{угол } \phi_* = \arctg\left(\frac{\eta_0}{\xi_0}\right).$$

Частотной характеристикой (ЧХ) ЛДС $H(e^{j\hat{\omega}})$ называется зависимость отношения реакции к дискретному гармоническому воздействию в установившемся режиме.

Модуль ЧХ называется амплитудно-частотной характеристикой (АЧХ) $A(\hat{\omega}) = |H(e^{j\hat{\omega}})|$.

АЧХ – это частотная зависимость отношения амплитуды реакции к амплитуде дискретного гармонического воздействия в установившемся режиме.

$$A(\hat{\omega}) = \sqrt{\frac{[b_0 + b_1 \cos(\hat{\omega}) + b_2 \cos(2\hat{\omega})]^2 + [b_1 \sin(\hat{\omega}) + b_2 \sin(2\hat{\omega})]^2}{[1 + a_1 \cos(\hat{\omega}) + a_2 \cos(2\hat{\omega})]^2 + [a_1 \sin(\hat{\omega}) + a_2 \sin(2\hat{\omega})]^2}} \quad (4.7)$$

Аргумент ЧХ называется фазочастотной характеристикой (ФЧХ)

$$\phi(\hat{\omega}) = \arg\{H(e^{j\hat{\omega}})\}.$$

ФЧХ – это частотная зависимость разности фаз реакции и дискретного гармонического воздействия в установившемся режиме.

$$\phi(\hat{\omega}) = \arctg \frac{a_1 \sin(\hat{\omega}) + a_2 \sin(2\hat{\omega})}{1 + a_1 \cos(\hat{\omega}) + a_2 \cos(2\hat{\omega})} - \arctg \frac{b_1 \sin(\hat{\omega}) + b_2 \sin(2\hat{\omega})}{b_0 + b_1 \cos(\hat{\omega}) + b_2 \cos(2\hat{\omega})} \quad (4.8)$$

Для оценки АЧХ звена 2-го порядка достаточно построить ее график по пяти точкам:

- Двум – на границах основной полосы
- Одной- уточняющей – по середине основной полосы
- Двум – внутри основной полосы, соответствующим максимуму и минимуму (либо нулю) АЧХ.

Определим значения АЧХ и ФЧХ в указанных точках.

1. В точке $\hat{\omega}=0$

$$A(0) = |H(1)| = \left| \frac{b_0 + b_1 + b_2}{1 + a_1 + a_2} \right|; \quad \phi(0) = \arg\{H(1)\} = 0$$

2. В точке $\hat{\omega} = \pi$

$$A(\pi) = |H(-1)| = \left| \frac{b_0 - b_1 + b_2}{1 - a_1 + a_2} \right|; \quad \phi(\pi) = \arg\{H(-1)\} = 0$$

3. В точке $\hat{\omega} = \pi/2$

$$A(\pi/2) = \left| \frac{(b_0 - b_2)^2 + b_1^2}{(1 - a_2)^2 + a_1^2} \right|;$$

$$\phi(\pi/2) = \arg\left(\frac{a_1}{1 - a_2}\right) - \arctg\left(\frac{b_1}{b_0 - b_2}\right)$$

4. В точке $\hat{\omega} = \hat{\omega}_*$ определяется местоположение максимума АЧХ: $\hat{\omega}_* \approx \phi_*$

5. В точке $\hat{\omega} = \hat{\omega}_0$ определяется местоположение минимума АЧХ: при значении $r_0 \neq 1$ $\hat{\omega}_0 \approx \phi_0$. Если в точке ϕ_0 $r_0 = 1$, имеется не минимум, а нуль АЧХ и скачок ФЧХ на π .

Анализ АЧХ по карте нулей и полюсов:

- Если ПФ звена 2-го порядка имеет два комплексно-сопряженных или два вещественных полюса, то максимум АЧХ находится внутри или на границах (два максимума) основного диапазона
- Если ПФ звена 2-го порядка имеет два комплексно-сопряженных или два вещественных нуля, то минимум АЧХ находится внутри или на границах (два минимума) основного диапазона
- Расположение нулей, не лежащих на единичной окружности, приблизительно указывает частоты, на которых АЧХ имеет минимум.
- Расположение нулей на единичной окружности определяет частоты, на которых АЧХ равна нулю, а ФЧХ делает скачок на π . Нуль является наименьшим значением, но не минимумом АЧХ.
- Абсолютная величина полюса влияет на крутизну АЧХ. Если она велика, крутизна АЧХ также велика.

ЛЕКЦИЯ 5

5.1. Сообщения, знаки и символы

Изначально большинство передаваемой информации (кроме информационного обмена между двумя компьютерами) имеет текстовую или аналоговую форму. Если информация является буквенно-цифровым текстом, то используется один или несколько стандартных форматов – методов знакового кодирования: ASCII (American Standard Code for Information Interchange – Американский стандартный код для обмена информацией), EBCDIC (Extended Binary Coded Decimal Interchange Code – расширенный двоичный код для обмена информацией), код Бодо, код Холлерита и т.д. Таким образом текстовый материал преобразуется в цифровой формат.

Знаки вначале кодируются в последовательность битов, которая называется потоком битов или узкополосным сигналом. После этого формируются группы из k бит, называемые символами, причем число всех символов конечно, а их совокупность называется алфавитом. Размер алфавита равен $M = 2^k$. Выбор значений M , k является важным первоначальным этапом проектирования любой цифровой системы. При $k=1$ $M=2$, а система называется бинарной. Модулятор при этом использует один из двух различных сигналов для представления значения «один» и «ноль». В этом частном случае символ и бит – это одно и то же.

На рис. 5.1. приведен пример разделения потока битов для различных значений M и k . Текстовое сообщение на рисунке – это слово «THINK». Использование 6-битовой кодировки ASCII дает поток, состоящий из 30 бит. Если размер алфавита выбрать $M=8$, то биты должны группироваться по три ($k = \log_2 8$), а передатчик должен иметь набор из восьми сигналов $s_i(t)$, которые ставятся в соответствие символам.

Если выбрать $M=32$, то $k = \log_2 32 = 5$. Биты сообщения группируются по пять для создания символов. Разбиение потока битов на символы не обязательно должно совпадать с границами кодированных текстовых знаков, так как передается строка сообщения, и только конечный пользователь приписывает текстовое значение полученной последовательности битов.

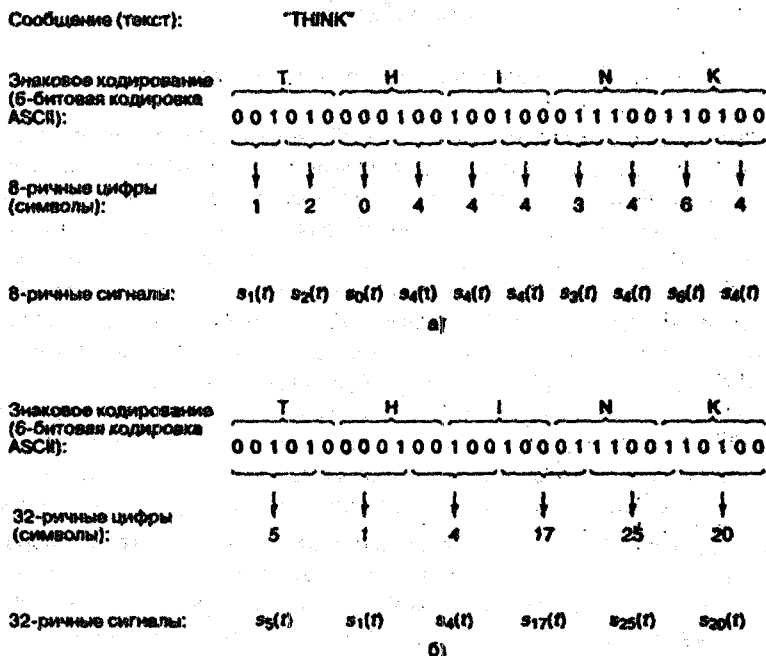


Рис. 5.1. Сообщения, знаки и символы:
а) 8-ричный пример; б) 32-ричный пример.

Если информация является аналоговой, ее знаковое кодирование (как в случае текстовой информации) невозможно. Вначале информацию следует перевести в цифровой формат. Для этого ее нужно дискретизировать по времени и квантовать по амплитуде. Дискретизация аналоговой информации рассмотрена в лекции 3.

5.2. Квантование аналогового сигнала

Рассмотрим процесс квантования сигнала по амплитуде. На рисунке 5.2 изображено L -уровневое устройство квантования аналогового сигнала с полным диапазоном напряжений, равным $V_{pp} = V_p - (-V_p) = 2V_p$ В. Шаг между уровнями квантования, называемый интервалом квантования, составляет q вольт. Если уровни квантования равномерно распределены по всему диапазону, устройство квантования называется равномерным или линейным. Каждое дискретное значение аналогового сигнала аппроксимируется квантованным им-

пульсом: аппроксимация дает ошибку, не превышающую абсолютного значения $q/2$.

Хорошим критерием качества равномерного устройства квантования является его дисперсия (среднеквадратическая ошибка при подразумеваемом нулевом среднем). Если считать, что ошибка квантования, е. равномерно распределена в пределах интервала квантования шириной q (т.е. аналоговый входящий сигнал принимает все возможные значения с равной вероятностью), то дисперсия ошибок для устройства квантования составляет:

$$\sigma^2 = \int_{-q/2}^{q/2} e^2 p(e) de = \int_{-q/2}^{q/2} e^2 \frac{1}{q} de = \frac{q^2}{12}, \quad (5.1)$$

где $p(e)$ – равномерно распределенная плотность вероятности возникновения ошибки квантования.

Дисперсия σ^2 соответствует средней мощности шума квантования. Пиковую мощность аналогового сигнала, нормированную на 1 Ом, можно выразить как:

$$V_p^2 = \left(\frac{V_{pp}}{2}\right)^2 = \left(\frac{Lq}{2}\right)^2 = \frac{L^2 q^2}{4}, \quad (5.2)$$

где L – число уровней квантования.

Объединение выражений (5.1) и (5.2) дает отношение пиковой мощности сигнала к средней мощности квантованного шума:

$$\left(\frac{S}{N}\right)_q = \frac{L^2 q^2 / 4}{q^2 / 12} = 3L^2. \quad (5.3)$$

Отношение $\left(\frac{S}{N}\right)_q$ квадратично растет с числом уровней квантования. При бесконечном росте числа уровней квантования имеем нулевой шум квантования (фактически аналоговый сигнал).

Число бит, выделяемое каждой аналоговой выборке, зависит от допустимого искажения, вызванного квантованием. Пусть величина ошибки квантования $|e|$ определяется как часть p удвоенной амплитуды напряжения аналогового сигнала:

$$|e| \leq pV_{pp}. \quad (5.4)$$

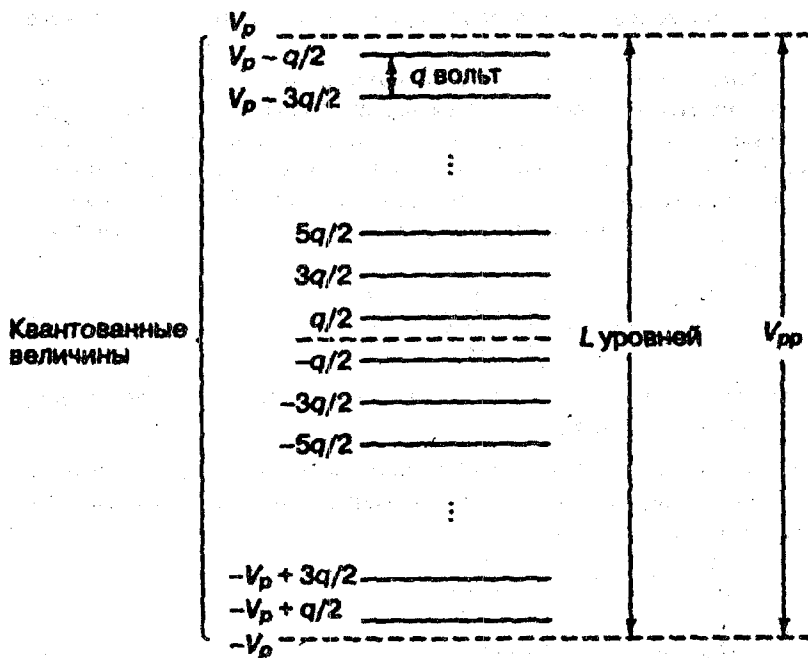


Рис. 5.2. Уровни квантования.

Поскольку ошибка квантования не может быть больше $q/2$, можно записать:

$$|e|_{\max} = \frac{q}{2} = \frac{V_{pp}}{2(L-1)} \approx \frac{V_{pp}}{2L}. \quad (5.5)$$

Если число уровней квантования достаточно велико, то $(L-1)$ можно заменить на L . С учетом (5.4) и (5.5) можно записать:

$$\frac{V_{pp}}{2L} \leq pV_{pp}, \quad (5.6)$$

$$2^l = L \geq \frac{1}{2p} \text{ уровней.} \quad (5.7)$$

$$\text{Количество бит, выделяемое аналоговой выборке, } l \geq \log_2 \frac{1}{2p} \quad (5.8)$$

На практике сигналы не имеют равномерного распределения, поэтому среднеквадратическое напряжение сигнала зависит от типа передаваемого сигнала и распределения напряжения, соответствующего этому сигналу. Независимо от того, известно ли точное распределение напряжений для заданного сигнала, можно хотя бы оценить отношение сигнала к шуму квантования для наибольшей амплитуды. На входе устройства квантования максимальную амплитуду сигнала можно записать как

$$V_{\max} = \frac{L}{2} q \quad (5.9)$$

Отношение максимального напряжения сигнала к среднеквадратическому значению напряжения шума

$$\frac{V_{\max}}{\sigma} = \frac{\frac{L}{2} q}{\frac{q}{\sqrt{12}}} = \frac{L}{2} q \frac{\sqrt{3}\sqrt{4}}{q} = \sqrt{3}L. \quad (5.10)$$

Для определения отношения сигнала к погрешности квантования найдем отношение сигнал/шум в терминах мощности. Для этого возьмем квадраты напряжений сигнала и шума:

$$\left(\frac{S}{N}\right)_q = 3L^2. \quad (5.11)$$

$$\text{Количество уровней квантования } L = 2^l - 1. \quad (5.12)$$

Подставив (5.12) в (5.11) получим:

$$\left(\frac{S}{N}\right)_q = 3L^2 = 3(2^l - 1)^2. \quad (5.13)$$

То же отношение в децибелах равно

$$\left(\frac{S}{N}\right)_q = 10 \lg \{3(2^l - 1)^2\} \approx 10 \lg 3 + 20 \lg 2^l = 4,8 + 6l. \quad (5.14)$$

Задача

Квантование выборок в системе осуществляется с помощью 8-битового АЦП. Оценить отношение сигнал/шум квантования для максимальной амплитуды сигнала.

Решение

$$\left(\frac{S}{N}\right)_q = 4,8 + 6l = 4,8 + 6 \cdot 8 = 52,8 \text{ дБ.}$$

Это отношение сигнал/шум очень велико, но оно справедливо только для максимальной амплитуды сигнала в данной системе, которое хотя и показательно, но не отображает характеристик квантования реальных сигналов.

Задача

В цифровой аудиосистеме проигрывание компакт-дисков аналоговый оцифровывается так, что отношение пиковой мощности сигнала к пиковой мощности шума квантования не менее 96 дБ. Частота дискретизации f_s – 44,1 тысяча выборок в секунду.

- Сколько необходимо уровней квантования аналогового сигнала, чтобы $(S/N_q)_{\text{пик}} = 96$ дБ?
- Какое число бит на выборку необходимо при таком числе уровней?
- Чему равна скорость передачи данных в бит/с?

Решение

Определим число уровней квантования:

$$3L^2 = \left(\frac{S}{N}\right)_q; \left(\frac{S}{N}\right)_q = 96 \text{ дБ},$$

$$10 \lg x = 96 \Rightarrow \lg x = 9,6 \Rightarrow x = 10^{9,6} \approx 4 \cdot 10^9 \text{ раз},$$

$$3L^2 = 4 \cdot 10^9,$$

$$L = \sqrt{\frac{4 \cdot 10^9}{3}} \approx 13,3 \cdot 10^4.$$

Число бит на выборку

$$2^L = L \Rightarrow l = \log_2 L = \log_2(133\,000) \approx 12$$

Скорость передачи данных равна

$$R = l f_s = 12 \cdot 44,1 \cdot 10^3 = 530\,000 \frac{\text{бит}}{\text{с}}$$

Задача

Сигнал речевого диапазона (300–3300 Гц) оцифровывается так, что квантовое искажение $p \leq 0,1\%$ максимального напряжения сигнала. Частота дискретизации равна 8000 выборок/с и используется 32-

уровневая кодировка импульсно-амплитудной модуляции. Определить теоретическую минимальную ширину полосы, при которой еще не возникает межсимвольная интерференция.

Решение

Число бит на уровень квантования

$$l \geq \log_2 \frac{1}{2p} = \log_2 \frac{1}{0,002} = \log_2 500 \approx 9.$$

Скорость передачи битов $R = lf_s = 9 \cdot 8 \cdot 10^3 = 72\,000 \frac{\text{бит}}{\text{с}}$ удовлетворяет требованиям, относящимся к искажениям.

Поскольку используется алфавит сигналов с $M = 2^k = 32$ уровнями, то $k=5$ бит/символ.

Поток битов разбивается на группы по 5 битов с целью формирования новых 32-уровневых символов, и скорость передачи символов равна:

$$R_s = \frac{R}{k} = \frac{72\,000}{5} = 14\,400 \text{ символов/с.}$$

Ширина полосы для передачи символов без межсимвольной интерференции равна:

$$W \geq \frac{R_s}{2} = \frac{14\,400}{2} = 7\,200 \text{ Гц.}$$

Задача

Определить минимальную ширину полосы частот, необходимую для 30-канальной цифровой системы передачи речи. Сравнить эту ширину полосы с полосой частот эквивалентной схемы с частотным разделением.

Решение

30-канальная система, используемая для телефонной связи в Европе, имеет 32 временных интервала (2 из них используется для выравнивания кадров и синхронизации сигналов). Каждый входной аналоговый сигнал речевого диапазона дискретизируется с частотой Найквиста 8 кГц и подается на 8-битовое АЦП.

Скорость передачи $R = 32 \cdot 8 \cdot 8 \cdot 10^3 = 2,048$ Мбит/с, поэтому минимальная полоса частот составляет $2,048/2 = 1,024$ МГц.

Полоса частот эквивалентной 30-канальной системы с частотным разделением при условии, что каждый канал имеет ширину полосы 4 кГц, равна $30 \cdot 4 \text{ кГц} = 120 \text{ кГц}$.

Таким образом, по сравнению с системами с частотным разделением, при цифровой передаче сигналов требуется полоса пропускания на порядок больше, чем при эквивалентной аналоговой передаче.

При увеличении количества уровней квантования улучшается отношение сигнала к шуму дискретизации. Для получения приемлемых значений отношения сигнала к шуму квантования сигналов с малой амплитудой при линейном квантовании нужно большое количество уровней квантования. Это приводит к излишне большим значениям отношения сигнала к шуму квантования для сигналов с большей амплитудой и связано с соответствующим увеличением ширины полосы частот.

Когда существует неопределенность в статистиках входного сигнала, можно создать неравномерное устройство квантования, дающее большее отношение сигнал/шум, чем равномерное, и использующее то же количество бит на уровень. Это реализуется с помощью деления входной динамической области на неравномерные интервалы так, что мощность шума, взвешенная вероятностью появления на каждом интервале, является одинаковой. Можно смоделировать неравномерное квантующее устройство как последовательность операторов, показанных на рис. 5.3.

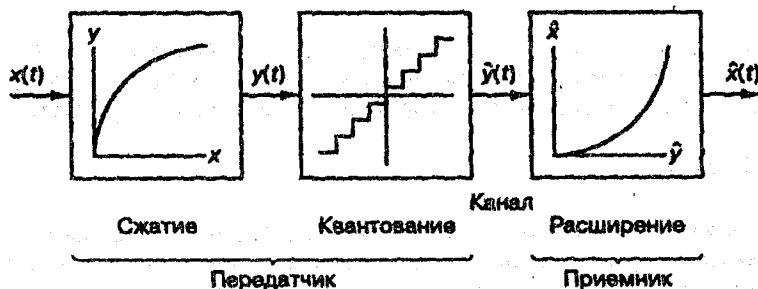


Рис. 5.3. Неравномерное устройство квантования как последовательность операторов: сжатие, равномерное квантование и расширение.

Сначала входной сигнал отображается с помощью нелинейной функции, называемой компрессором; в альтернативную область уровней. Эти уровни равномерно квантуются и квантованные уровни сигнала затем отображаются с помощью дополняющей нелинейной функции, называемой экспандером, в выходную область уровней. Вся операция называется *компандированием*. В ранних системах функции сжатия были гладкими логарифмическими. Большинство современных систем использует кусочно-линейную аппроксимацию функции логарифмического сжатия. В США характеристика устройства сжатия описывается мю-законом, введенным компанией Bell-System. В Европе используется компандер, реализующий А-закон (стандарт CCITT – Consultative Committee for International Telephone and Telegraphy – Международного консультативного комитета по телеграфии и телефонии).

Основные источники искажения аналогового сигнала, восстановленного из дискретных, квантованных и переданных импульсов связаны с устройством АЦП и воздействием канала. С устройством АЦП связаны:

- *Шум квантования* – необходимость аппроксимации исходного сигнала квантованными выборками. Величина шума квантования обратно пропорциональна числу уровней квантования.
- *Насыщение устройства квантования*. Для аппроксимации значений из непрерывного диапазона значениями из конечного множества устройство выделяет L уровней. Диапазон входных значений, для которых разница между входом и выходом незначительна, называется рабочим диапазоном преобразователя. Если входное значение не принадлежит этому диапазону, значения на входе и выходе отличаются сильнее и говорят, что преобразователь работает в режиме насыщения. В общем случае насыщение устраняется путем автоматической регулировки усиления (automatic gain control), которое расширяет рабочий диапазон преобразователя.
- *Синхронизация случайного смещения*. При наличии случайного смещения выборки дискретизация уже не является равномерной и происходит эффект, равносильный частотной модуляции узкополосного сигнала. Управлять синхронизацией случайного смещения можно посредством развязки по питанию и использования кварцевых генераторов.

С воздействием канала связаны искажения:

- *Пороговый эффект* – шумовой компонент на выходе детектора аддитивен, если отношение сигнал/шум достаточно велико и мультипликативен, если отношение сигнал/шум низкое.

- *Межсимвольная интерференция (Intersymbol Interference)*, которая наблюдается, если ширина полосы частот канала меньше полосы частот сигнала.

5.3. Импульсно-кодовая модуляция (ИКМ)

Импульсно-кодовая модуляция (pulse-code modulation — PCM) — это название, данное классу узкополосных сигналов, полученных из сигналов импульсно-амплитудной модуляции (pulse-amplitude modulation — PAM) путем кодирования каждой квантованной выборки цифровым словом. Исходная информация дискретизируется и квантуется в один из L уровней; после этого каждая квантованная выборка проходит цифровое кодирование для превращения в l -битовое ($l = \log_2 L$) кодовое слово. Для узкополосной передачи биты кодового слова преобразовываются в импульсные сигналы. Рассмотрим рис. 5.4, на котором представлена бинарная импульсно-кодовая модуляция. Предположим, что амплитуды аналогового сигнала $x(t)$ ограничены диапазоном от -4 до $+4$ В. Шаг между уровнями квантования составляет 1 В. Следовательно, используется 8 квантовых уровней; они расположены на $-3,5$; $-2,5$; ...; $+3,5$ В. Уровню $-3,5$ В присвоим кодовый номер 0, уровню $-2,5$ В — 1 и так до уровня $3,5$ В, которому присвоим кодовый номер 7. Каждый кодовый номер имеет представление в двоичной арифметике — от 000 для кодового номера 0 до 111 для кодового номера 7. Почему уровни напряжения выбраны именно так, а не с использованием набора последовательных чисел 1, 2, 3, ...? На выбор уровней напряжения влияют два ограничения. Во-первых, интервалы квантования между уровнями должны быть одинаковыми; и, во-вторых, удобно, чтобы уровни были симметричны относительно нуля.

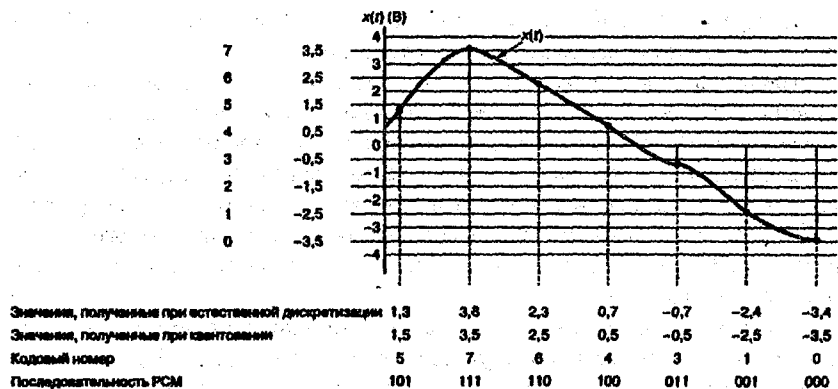


Рис. 5.4. Естественные выборки, квантованные выборки и РСМ.

На оси ординат рис. 5.4 отложены уровни квантования и их кодовые номера. Каждая выборка аналогового сигнала аппроксимируется ближайшим уровнем квантования. Под аналоговым сигналом $x(t)$ изображены четыре его представления: значения выборок в естественной дискретизации, значения квантованных выборок, кодовые номера и последовательность РСМ. Каждая выборка соотнесена с одним из восьми уровней или трехбитовой последовательностью РСМ.

Если передаваемый сигнал является двоичным (т.е. имеет два уровня напряжения), то скорость передачи данных, которая может поддерживаться при полосе W Гц, равна $2W$ бит/с. В качестве примера рассмотрим речевой канал, который посредством модема используется для передачи цифровых данных. Предположим, что ширина полосы равна 3100 Гц. В этом случае пропускная способность C канала равна $2W = 6200$ бит/с. Впрочем, для передачи могут использоваться сигналы с более чем двумя уровнями напряжения; т.е. каждый элемент сигнала может представлять несколько битов. Если, например, в качестве сигналов используются четыре возможных уровня напряжения, то каждый элемент сигнала может представлять два бита. При многоуровневой передаче данных формулировка Найквиста становится следующей:

$$c = 2W \log_2 M, \quad (5.15)$$

где M — количество дискретных уровней сигнала или напряжения.

Следовательно, для $M = 8$ пропускная способность равна 18 600 бит/с — величине, используемой некоторыми модемами.

Итак, при заданной ширине полосы скорость передачи данных может быть увеличена за счет увеличения количества различных элементов сигнала. Впрочем, при таком подходе на приемник возлагается дополнительная нагрузка: вместо распознавания в каждый момент приема двух возможных элементов сигнала приемник должен распознать один из M возможных элементов. Кроме того, на практике величину M ограничивают помехи и прочие искажения сигнала, возникающие в линии передачи.

Рассмотрим задачу, которую должен выполнять приемник. Он должен различать все возможные уровни каждого импульса. Одинаково ли легко приемник различает восемь возможных уровней импульса, приведенного на рис. 5.5, а, и два возможных уровня каждого двоичного импульса приведенного на рис. 5.5, б?

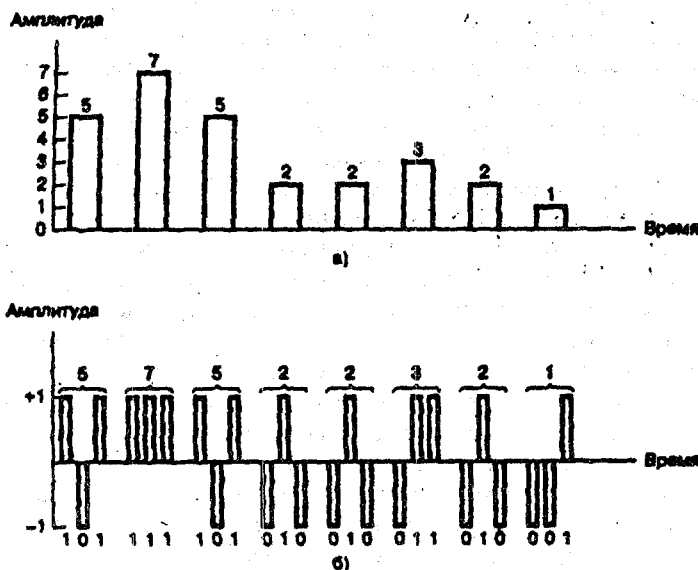


Рис. 5.5. Передача сигналов PCM: а) восьмиуровневая б) двухуровневая.

Передача восьмиуровневого (по сравнению с двухуровневым) импульса требует большей энергии для эквивалентной эффективности обнаружения. При равной средней мощности двоичных и восьмеричных импульсов первые обнаружить проще, поскольку детектор приемника при принятии решения о принадлежности сигнала к одному из двух

уровней располагает большей энергией сигнала на каждый уровень, чем при принятии решения относительно принадлежности сигнала к одному из 8 уровней. Чем расплачивается разработчик системы, если решает использовать более удобную в обнаружении двоичную кодировку РСМ, а не восьмиуровневую кодировку РАМ? Плата состоит в трехкратном увеличении ширины полосы для данной скорости передачи данных по сравнению с восьмиричными импульсами, поскольку каждый восьмиричный импульс должен заменяться тремя двоичными (ширина каждого из которых втрое меньше ширины восьмиричного импульса). Может возникнуть вопрос, почему бы ни использовать двоичные импульсы той же длительности, что и восьмиричные, и разрешить запаздывание информации? В некоторых случаях это приемлемо, но для систем связи реального времени такое увеличение задержки допустить нельзя.

5.4. Дифференциальная импульсно-кодовая модуляция (ДИКМ, Differential PCM – DPCM)

Устройства квантования, рассмотренные выше, называются мгновенными устройствами квантования или *устройствами квантования без памяти*, так как цифровые преобразования основаны на единичной (текущей) входной выборке. Мгновенные квантующие устройства кодируют источник, принимая во внимание плотность вероятности, сопоставленную с каждой выборкой. Методы квантования, которые принимают во внимание корреляцию между выборками, являются *квантующими устройствами с памятью*. Эти квантующие устройства уменьшают избыточность источника сначала посредством превращения коррелированной входной последовательности в связанную последовательность с уменьшенной корреляцией, уменьшенной дисперсией и уменьшенной полосой частот. Затем эта новая последовательность квантуется с использованием меньшего количества бит.

Корреляционные характеристики источника можно представить во временной области с помощью выборки его автокорреляционной функции и в частотной области — его спектром мощности. Если изучается спектр мощности $C_x(f)$ кратковременного речевого сигнала, как изображено на рис. 5.6, то видим, что спектр имеет глобальный максимум в окрестности от 300 до 800 Гц и убывает со скоростью от 6 до 12 дБ/октаву. Изучив этот спектр, можно взглянуть на определенные свойства временной функции, из которой он получен. Видим, что большие изменения сигнала происходят медленно (низкая частота), а быстрые (высокая частота) должны иметь малую амплитуду.

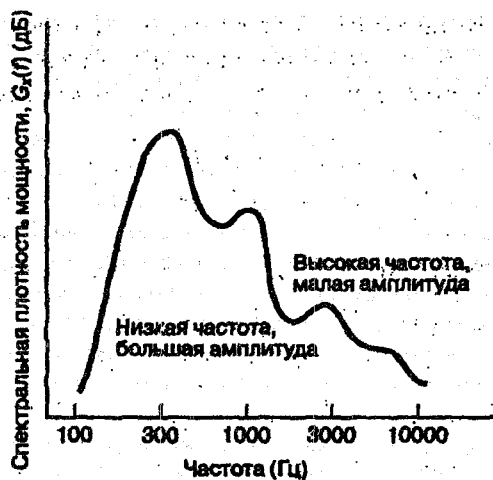


Рис. 5.6. Типичный спектр мощности речевого сигнала.

Эквивалентная интерпретация может быть дана в терминах автокорреляционной функции сигнала $R_x(T)$, как изображено на рис. 5.7.

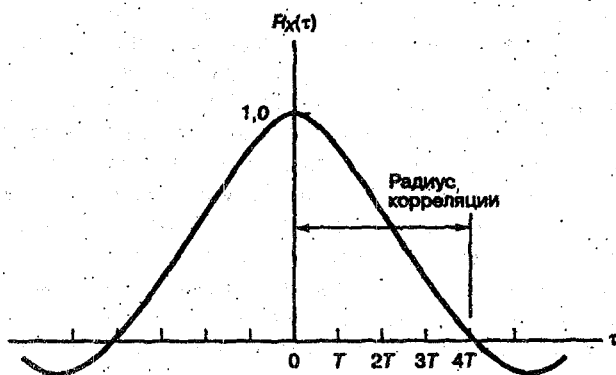


Рис. 5.7. Типичная АКФ речевых сигналов.

Здесь широкая, медленно меняющаяся автокорреляционная функция свидетельствует о том, что при переходе от выборки к выборке будет только слабое изменение и что для полного изменения амплитуды требуется временной интервал, превышающий интервал корреляции. Интервал (или ради-

ус) корреляции, рассмотренный на рис. 5.7, является временной разностью между максимальной и первой нулевой корреляцией. В частности, значение корреляции для типичного единичного выборочного запаздывания лежит в диапазоне примерно от 0,79 до 0,87, а радиус корреляции имеет порядок от 4 до 6 выборочных интервалов, равных T секунд на интервал.

Поскольку разность между соседними временными выборками для речи мала, используемый метод кодирования базируется на передаче от выборки к выборке разностей, а не действительных выборочных значений. В действительности, последовательные разности представляют собой частный случай класса преобразователей с памятью, называемых N -отводными линейными кодерами с предсказанием. Эти кодеры, иногда именуемые кодерами с предсказаниями и поправками, предсказывают следующее входное выборочное значение на основании предыдущих входных выборочных значений. Эта структура показана на рис. 5.8.

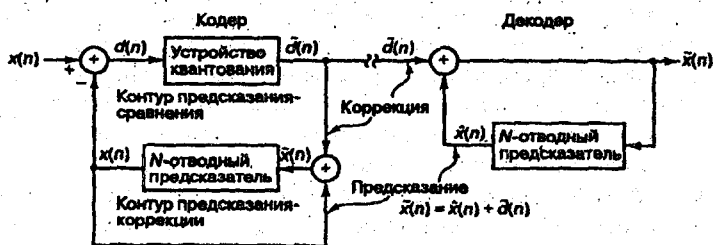


Рис. 5.8. N -отводный дифференциальный импульсно-кодовый модулятор (ДИКМ) с предсказанием.

В этом типе преобразователя передатчик и приемник имеют одинаковую модель предсказания, которая получена из корреляционных характеристик сигнала. Кодер дает ошибку предсказания (или остаток) как разность между следующим измеренным и предсказанным выборочными значениями. Математически контур предсказания описывается следующим образом:

$$d(n) = x(n) - \hat{x}(n), \quad (5.16)$$

где $x(n)$ — n -я входная выборка; $\hat{x}(n)$ — предсказанное значение выборки; $d(n)$ — соответствующая ошибка предсказания. Эта операция производится в контуре предсказания и сравнения. Кодер корректирует свои предсказания, составляя сумму предсказанного значения и ошибки предсказания. Математически контур коррекции описывается следующим образом:

$$\begin{aligned}\tilde{d}(n) &= \text{quant}[d(n)], \\ \tilde{x}(n) &= \hat{x}(n) + \tilde{d}(n).\end{aligned}\tag{5.17}$$

Здесь $\text{quant}(\ast)$ представляет операцию квантования, $\tilde{d}(n)$ — квантованная версия ошибки предсказания, а $\tilde{x}(n)$ — скорректированная и квантованная версия входной выборки.

Это делается в контуре предсказания и поправок. Декодер должен быть также проинформирован об ошибках предсказания, чтобы использовать свой контур коррекции для поправки своего предсказания. Декодер «повторяет» обратный цикл кодера.

Задача связи состоит в передаче разности (ошибки сигнала) между предсказанными и действительными выборочными данными. Преобразователи с предсказанием должны иметь кратковременную память, которая поддерживает проводимые в реальном времени операции, требуемые для алгоритма предсказания. Кроме того, они часто будут иметь долгосрочную память, которая поддерживает медленные, зависящие от данных, операции, такие, как автоматическая регулировка усиления, коррекция коэффициентов фильтра. Предсказатели, которые включают медленные, зависящие от данных, *регулирующие алгоритмы, называются адаптивными*.

Одноотводный линейный кодирующий фильтр с предсказанием (linear prediction coding filter — фильтр LPC) в процессе модуляции DPCM предсказывает следующее входное выборочное значение, основываясь на предшествующем входном выборочном значении.

N-отводный фильтр LPC предсказывает следующее выборочное значение на основе линейной комбинации предшествующих N выборочных значений.

Дельта-модуляция, часто обозначаемая как Δ -модуляция, представляет собой процесс внедрения низкой разрешающей способности аналогоцифрового преобразователя в контур обратной связи дискретных данных, работающий со скоростью, значительно превышающей частоту Найквиста. Причиной возникновения этой технологии стало то, что в процессе преобразования скорость — это менее дорогой ресурс, чем точность, и разумнее будет использовать более быстрые процессы обработки сигналов для получения более высокой точности.

Причина выбора высокой частоты дискретизации заключается в следующем: необходимо убедиться, что выборочные данные имеют вы-

сокую корреляцию, так что простой одноотводный предсказатель будет давать малую ошибку предсказания, которая, в свою очередь, допускает работу устройства квантования с очень малым количеством бит в контуре коррекции ошибок. Простейшей формой устройства квантования является однобитовый преобразователь; по сути, это просто компаратор, который обнаруживает и сообщает знак разности сигнала. Как следствие, ошибкой предсказания сигнала является однобитовое слово, которое имеет интересное преимущество — оно не требует следить за порядком слов при последовательной обработке.

Сигма-дельта-модуляция ($\Sigma-\Delta$ — модуляция). Понять сигма-дельта модулятор можно путем рассмотрения контура обратной связи по шуму. Устройство квантования добавляет ошибку к входному сигналу. Когда выборки образуются с значительным запасом, то высоко коррелируют не только выборки, но и ошибки, которые становятся предсказуемыми и могут быть вычтены из сигнала, отправленного на устройство квантования прежде, чем произойдет процесс квантования. Предшествующая ошибка, образованная как разность между входом и выходом устройства квантования, помещается в регистр запаздывания для использования в качестве следующей ошибки квантования.

Задача

Одноконтурный сигма-дельта-модулятор работает с частотой, в 20 раз превышающей частоту Найквиста для сигнала с полосой частот 10 кГц. Преобразователь представляет собой 1-битовый АЦП.

А) Определите максимальное SNR для входного сигнала в 8 кГц.

Б) Определите максимальное SNR для того же сигнала, если модулятор работает с частотой, в 50 раз превышающей частоту Найквиста.

Решение

Мощность шума на выходе ФНЧ, стоящего после сигма-дельта-модулятора:

$$N(\omega) = \frac{N_0}{3} \left(\frac{f_{BW}}{f_s} \right)^3,$$

где N_0 — мощность шума на входе сигма-дельта-модулятора;

f_{BW} — полоса частот выходного сигнала;

f_s — частота дискретизации.

Выполняя пункт А): $f_{BW}=8000$ Гц; $f_s=400\,000$ Гц (частота Найквиста для сигнала 10 кГц равна 20 кГц, а учитывая превышение в 20 раз, получаем 400 000 Гц).

$$N(\omega) = 10 * \lg\left(\left(\frac{8}{400}\right)^3\right) = -50,9691 \text{ дБ},$$

тогда как при частоте Найквиста $f_s=20000$ Гц

$$N(\omega) = 10 * \lg\left(\left(\frac{8}{20}\right)^3\right) = -11,9382 \text{ дБ},$$

Разность величин составляет $50,9691-11,9382=39,0309$ дБ или 8000 раз.

Выполняя пункт Б): $N(\omega) = 10 * \lg\left(\left(\frac{8}{1000}\right)^3\right) = -62,9073 \text{ дБ}.$

Разность величин составляет $62,9073-11,9382=50,9691$ дБ или 123000 раз.

ЛЕКЦИЯ 6

6.1. Особенности передачи речи по цифровым каналам связи

В случае акустических данных (голоса) информацию можно представить непосредственно с помощью электромагнитного сигнала, имеющего такой же спектр. Впрочем, существует необходимость нахождения компромисса между точностью воспроизведения звука, передаваемого электрическими средствами, и стоимостью самой передачи, повышающейся при увеличении ширины полосы. Спектр речевого сигнала находится в диапазоне приблизительно от 100 Гц до 7000 кГц, хотя и более узкая полоса также даст приемлемое воспроизведение речи. Стандартным для звукового канала является спектр от 300 до 3400 Гц. Этого достаточно для передачи речи, также такой выбор сводит к минимуму требуемую пропускную способность передачи и позволяет использовать достаточно недорогие телефонные аппараты. Телефонный микрофон преобразует входящий акустический сигнал в электромагнитный сигнал, находящийся в диапазоне от 300 до 3400 Гц. Затем этот сигнал передается через телефонную сеть приемнику, который и воспроизводит его как акустический сигнал.

Важнейшее преимущество цифровой связи состоит в том, что цифровая техника позволяет обеспечить любое наперед заданное качество связи. Для цифровой передачи речи необходимо произвести аналого-цифровое преобразование речевого сигнала: подвергнуть аналоговый сигнал дискретизации, квантованию и кодированию. Совокупность этих операций называется импульсно-кодовой модуляцией (ИКМ). Для точного описания формы речевого сигнала, согласно теореме Котельникова, его дискретизацию приходится проводить с частотой 8 кГц (т. е. брать отсчеты через каждые 125 мкс), а для получения нормального качества воспроизведения речи квантовать каждый отсчет по шкале, разбитой на 8192 уровня (при выборе равномерной шкалы квантования). Чтобы закодировать каждое значение отсчета с помощью двоичного числа, потребуется 13 разрядов. В результате для передачи телефонного разговора с помощью последовательности двоичных импульсов необходима скорость $8 \times 13 = 104$ кбит/с (что соответствует при оптимальном кодировании полосе частот 52 кГц). Сравнивая это число с полосой частот в 3100 Гц, которая требуется для аналоговой передачи, нельзя не поразиться колоссальному росту необходимой полосы, которым приходится расплачиваться за преимущества цифровой передачи. Естественно попытаться при реализации цифровой системы передачи понизить скорость передачи.

Первый шаг в этом направлении довольно очевиден. Квантование на 2^{13} уровней оказывается необходимым потому, что уровни аналоговых речевых сигналов могут изменяться в диапазоне 60 дБ. При этом сигналы высокого уровня при равномерной шкале квантования квантуются с таким же шагом, что и сигналы низкого уровня. Однако человеческая речь характеризуется уникальными статистическими свойствами, одно из которых проиллюстрировано на рис. 6.1. На оси абсцисс отложены амплитуды сигнала, нормированные на среднеквадратическое значение величины таких амплитуд в типичном канале связи, а на оси ординат – вероятность. Для большинства каналов речевой связи доминируют очень низкие тона; 50 % времени напряжение, характеризующее энергию обнаруженной речи, составляет менее четверти среднеквадратического значения. Значения с большими амплитудами встречаются довольно редко; только 15 % времени напряжение превышает среднеквадратическое значение. Поэтому можно сделать вывод о том, что так как восприятие сигналов органами слуха человека пропорционально логарифму уровня сигнала, то естественно было бы сигналы высокого уровня квантовать более грубо, а низкого уровня – более точно. Применяв нелинейное квантование с использованием логарифмического закона, можно обойтись восемью разрядами на отсчет, сохранив почти такое

же качество передачи. В результате скорость передачи двоичных разрядов окажется равной 64 кбит/с. Именно эта скорость получила самое широкое распространение, она зафиксирована в рекомендации МККТТ C.711, и на ней работает аппаратура ИКМ во многих странах.



Рис. 6.1. Статистическое распределение амплитуд речи одного лица.

Аналоговый сигнал имеет большую избыточность. Это позволяет предсказывать очередной отсчет и передавать только разницу между фактическим и предсказанным значением каждого отсчета. Если применить хорошую схему предсказания, то изменение амплитуды приращения сигнала окажется меньше изменения амплитуды самого сигнала, что приведет к уменьшению количества передаваемой информации. На этом принципе строится дифференциальная ИКМ (ДИКМ) и адаптивная дифференциальная ИКМ (АДИКМ), которая позволяет понизить скорость передачи речи до 32 кбит/с и ниже за счет дальнейшего усложнения приемопередающей аппаратуры. Продолжив усложнять аппаратуру, можно довести скорость передачи речи до 100-300 бит/с. Можно представить себе, например, на передающей стороне преобразователь речи в текст, а на приемной стороне — читающую машину. Известны пути дальнейшего снижения скорости передачи речи, но не будем на этом останавливаться. Дело в том, что аппаратура цифровой передачи речи со скоростью 64 кбит/с всех удовлетворила потому, что она оказалась работоспособной при использовании самых простых симметричных кабелей с парной скруткой. Аппаратура ИКМ-30 начала свое триумфаль-

ное шествие с уплотнения соединительных линий между городскими телефонными станциями. Там, где раньше по кабельной паре можно было организовать соединительную линию для передачи лишь одного разговора, аппаратура ИКМ-30 позволила организовать по этой же паре передачу 30 разговоров. О лучшем использовании такой пары с помощью аналоговой аппаратуры многоканальной связи не могло быть и речи. Позднее появились аппаратура ИКМ-120 и другие высокопроизводительные системы, работающие по коаксиальным кабелям и волоконным световодам, и острота вопроса об уменьшении скорости передачи разговорных сигналов ниже 64 кбит/с в сетях проводной связи практически была снята. Даже многочисленные разработки аппаратуры цифровой передачи со скоростью 32 кбит/с, реализованные во многих странах на основе принципа АДИКМ, не получили достаточно широкого применения. Баланс между увеличением пропускной способности каналообразующей аппаратуры и сложностью оконечного оборудования в проводной связи пока так и не склонился в пользу первого решения.

Совсем другие перспективы открылись в конце 1980-х – начале 1990-х годов, когда начали развиваться сотовые системы цифровой радиотелефонной связи. В отличие от проводных сетей, где расширение пропускной способности возможно за счет прокладки новых линий, т. е. возобновления ресурсов пропускной способности, в радиосетях действует жесткий закон тесноты в эфире, и приходится иметь дело с невозобновляемым ресурсом радиочастот. Правда, идея сотовой связи как раз и состоит в возобновлении ресурса радиочастот путем повторения частоты передачи на территории, до которой не доходит сигнал той же частоты от мешающей радиостанции. Но возможности такого возобновления ресурса и здесь ограничены, поэтому дальнейшее усложнение аппаратуры ради снижения скорости передачи оказывается оправданным.

Например, в принятой в большинстве стран Европы системе сотовой цифровой связи GSM стандартные скорости передачи речи составляют 13 и 6,5 кбит/с. Для осуществления подобной системы передачи пришлось обратиться к более глубокому проникновению в механизм речеобразования.

Как известно, один из важнейших результатов современной теории передачи информации состоит в рекомендации разделения задач кодирования источника и кодирования канала. В задачу кодирования источника информации входит описание передаваемого сообщения в максимально экономной форме, т. е. удаление избыточности в сообщении. Полученное таким образом сжатое сообщение становится более уязвимым к воздействию помех и может оказаться искаженным при передаче. Поэтому после кодирования источника применяется кодирование кана-

ла, в задачу которого входит защита передаваемого сообщения от помех. Кодирование канала требует внести в передаваемое сообщение некоторую избыточность, но не случайную, которая присутствовала в первоначальном сообщении, а строго обоснованную теоретически, которая гарантирует оговоренное качество передачи.

Итак, имеется цифровая версия аналогового речевого сигнала, т. е. функция, описывающая, например, закон изменения тока во времени. Из такого сигнала нужно попытаться удалить избыточность. Эту задачу можно решать несколькими методами. Один из них – попытаться найти избыточность путем чисто математического анализа рассматриваемой функции. Другой путь решения задачи – анализ акустических характеристик этой функции (с точки зрения ее восприятия органами слуха). Наконец, можно искать избыточность моделированием самого процесса речеобразования. Именно последний из перечисленных методов нашел применение в современных системах цифровой радиосвязи.

6.2. Процесс речеобразования

В формировании речи участвуют легкие, трахеи, гортань, рот и нос. При изменении положения губ, языка, челюстей меняется конфигурация ротовой полости, которая представляет собой нелинейную акустическую систему, а следовательно, и резонансные частоты акустической системы. При формировании звонких звуков (гласных и согласных) голосовые связки колеблются, пропуская поток воздуха в виде квазипериодической последовательности треугольной формы. Эти последовательности содержат много гармоник частоты основного тона, затухающих приблизительно по закону 12 дБ/октава. Резонансная система голосового тракта усиливает одни из этих гармоник и подавляет другие. Говорящий управляет частотой основного тона только в определенных пределах. Так, у взрослого мужчины частота основного тона меняется в пределах от 50 до 250 Гц, у взрослой женщины от 120 до 300 Гц и до 350 Гц у детей.

В простейшем случае, при произнесении нейтральных (безударных) гласных голосовой тракт можно представить в виде трубы длиной L , с одной стороны которой в узле колебаний расположен источник звука (голосовые связки), а другая, где находится пучность звуковых волн, открыта (раскрытые губы). Такая труба резонирует на нечетных гармониках собственной частоты: f_0 , $3f_0$, $5f_0$. Тогда частоту f_0 можно найти как $f_0 = v_{зв} / 4L$ где $v_{зв} = 340$ м/с – скорость звука. При длине

голосового тракта 17 см. получаем, $f_0 = 500$ Гц и другие резонансные частоты: 1500 Гц, 2500 Гц и т.д. Эти резонансные частоты в акустике называются формантами. В процессе речи частоты формант постоянно изменяются.

Спектр звонких звуков в основном расположен в нижней полосе частот речевого сигнала и сгруппирован вокруг формант, однако скорость изменений формант такова, что спектр речи остается практически постоянным на промежутках менее 16 мс.

Глухие звуки представляют собой слабые шумоподобные колебания. При формировании глухих звуков голосовые связки не колеблются, а в голосовом тракте возникает сжатие, проходя которое поток воздуха генерирует широкий спектр частот (шум). В зависимости от того, где возникает сжатие (в передней части рта или в районе голосовых связок), воздействие голосового тракта на формируемый шум различно. Однако во всех случаях в спектре глухих звуков подчеркнуты верхние частоты.

При произнесении взрывных звуков голосовой тракт перекрыт в каком-то месте и затем, после мгновенного сжатия препятствия, в нем возникает переходный акустический процесс. Если к этому добавляется вибрация голосовых связок, то формируются звонкие взрывные звуки (например, твердые согласные).

Слуховая система человека (ухо) устроена так, что воспринимает весь спектр звуковых частот набором из большого количества биологических полосовых фильтров, настроенных по всему диапазону звуковых частот. При восприятии речи проявляются эффекты маскирования отдельных тонов сигналов. Так, сильные звуки подавляют слабые, что позволяет отбрасывать их при передаче без существенного ущерба для качества звучания восстановленного сигнала. Кроме того, отдельные частоты могут маскироваться сигналами близкостоящих от них частот, динамический диапазон речевых сигналов для различных людей достигает 60 дБ при изменении уровня речи отдельного человека в пределах 20 – 30 дБ.

Рассмотренные кратко механизмы образования и восприятия речи позволяют сформулировать основные идеи, использованные в технике анализа-синтеза речи при ее передаче в сжатом виде по каналам связи. Речь можно синтезировать с помощью генератора периодических импульсов, следующих с частотой основного тона для воспроизведения звонких звуков и генератора шума для воспроизведения глухих звуков. Оба исходных сигнала следует пропустить через фильтры с меняющимися во времени характеристиками, соответствующими характеристикам речевого тракта.

В вокодерах (от английского voice coder) измеряются параметры речевого аппарата человека, по которым на приемной стороне создается аналог такого же аппарата и синтезируется звук. При этом постулируется, что в течение 20–30 мс эти параметры остаются постоянными. Поэтому перед сжатием поток бит ИКМ нарезается на сегменты, содержащие по 160–240 отсчетов для оценки постоянных параметров.

Разработка вокодеров основана на использовании различных моделей голосового тракта. На рис. 6.2 показана однотрубная модель речеобразующего аппарата человека. Модель представляет собой M последовательно соединенных секций одинаковой длины, имеющих постоянную (в пределах секции) площадь сечения. Если вдувать воздух в такой свисток (подавать сигнал возбуждения), то в зависимости от конфигурации соединения и сигнала возбуждения на выходе формируется тот или иной звук. Таким образом, звук, получаемый с помощью данной модели, можно задать параметрами сигнала возбуждения и M значениями площадей секций свистка. От точности описания основного тона зависит и степень сжатия, и качество синтезированной речи.

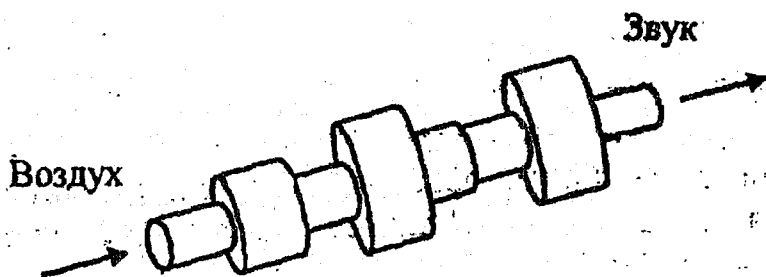


Рис. 6.2. Модель голосового тракта.

В системе GSM первых выпусков исходный цифровой поток речевого сигнала со скоростью передачи 104 кбит/с разбивался на отдельные блоки по 160 отсчетов, которые записывались. Каждый из таких блоков занимал промежуток времени 20 мс (иначе говоря, запоминались последовательности по $160 \times 13 = 2080$ разрядов). Зарегистрированные последовательности подвергались анализу, в результате которого для каждой из них находились восемь коэффициентов фильтрации, определяющие соответствующие резонансы, и возбуждающий сигнал. Именно эта информация передавалась приемнику, который воспроизводил по ней исходный речевой сигнал наподобие того, как это происходит в органах

речи человека (этот орган как бы настраивается с помощью восьми параметров, а затем при его возбуждении получается звук).

Однако упомянутый анализ распространяется на сравнительно короткие отрезки времени и не может выявить долгие гласные звуки, захватывающие соседние блоки. Поэтому для устранения избыточности при произнесении долгих гласных применяется долгосрочное предсказание. С этой целью в передатчике запоминаются переданные последовательности длительностью по 15 мс, с которыми сравниваются текущие последовательности. Из уже переданных выбирается последовательность, имеющая наибольшую корреляцию с текущей (т. е. больше других похожая на текущую), и передается только разность между текущей и выбранной последовательностями. Поскольку записанные в передатчике последовательности приемнику известны, нужно передать лишь указатель о том, с какой из записанных последовательностей проведено сравнение. Таким образом достигается дальнейшее сокращение объема передаваемой информации. В результате описанной обработки получается блок цифрового речевого сигнала продолжительностью 20 мс, содержащий 260 разрядов и имеющий скорость передачи всего 13 кбит/с (т. е. в восемь раз ниже исходной). Описанная процедура получила название регулярного импульсного возбуждения с долгосрочным предсказанием (английское сокращение RPE-LTP, которое расшифровывается в виде Regular Pulse Excitation – Long Term Prediction).

Наиболее эффективным методом сжатия речевого сигнала до скоростей 4—16 кбит/с считается алгоритм CELP (code excited linear prediction) — линейное предсказание и многоимпульсное возбуждение от кода. Сигнал возбуждения берется из заранее созданной кодовой книги, которая содержит от 2^{10} до 2^{16} векторов. Содержимое кодовой книги определяется в результате анализа достаточно длинных реальных речевых сигналов (детерминированная книга). В стохастическую (алгебраическую) книгу записываются коды отсчетов нормального белого шума с нулевым средним и единичной дисперсией.

Поиск оптимального вектора возбуждения требует проведения достаточно объемных вычислений. За 5—8 мс для каждого вектора надо определить ошибки предсказания, вычислить взвешенную среднеквадратическую ошибку и в качестве оптимального выбрать вектор, обеспечивающий минимальную ошибку. Номер (индекс) такого вектора передается по каналу. Для уменьшения числа операций вводятся дополнительные упрощенные книги. В приемнике имеется точная копия кодовой книги, из которой по принятому индексу извлекается оптимальный вектор возбуждения и посылается в синтезатор, идентичный используемому в кодере.

На рис. 6.3 представлена блок-схема преобразователя звука типа вокодер.

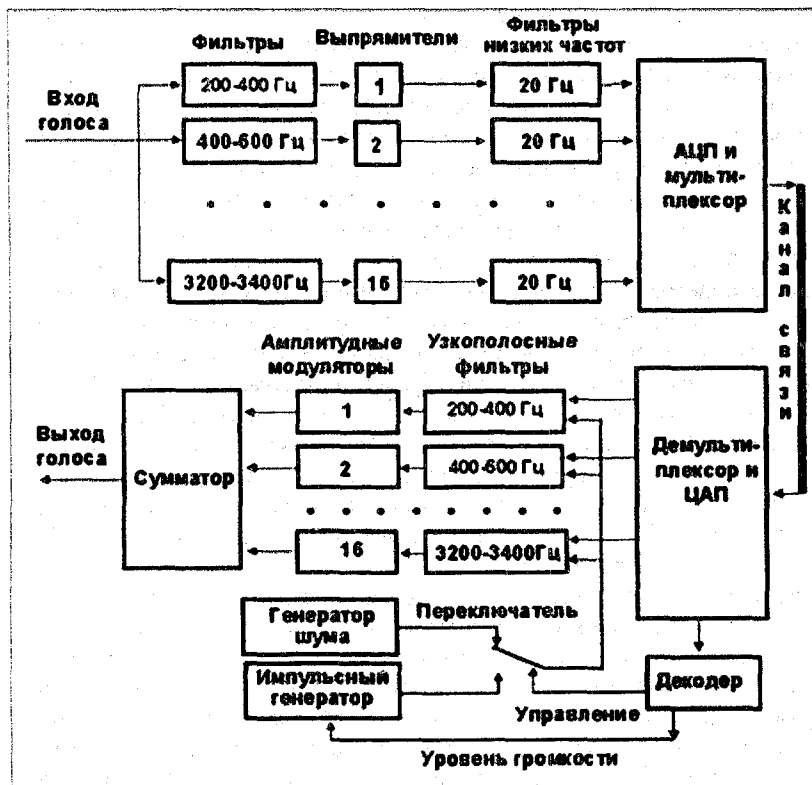


Рис. 6.3. Блок-схема кодирования/декодирования человеческого голоса (Vocoder).

Исходный спектр человеческого голоса здесь делится на ряд субдиапазонов (на рис. 6.3 их число равно 16) по 200 Гц каждый. Эти субдиапазоны выделяются узкополосными фильтрами, за которыми следуют выпрямители и фильтры низких частот (20 Гц). Выходные сигналы этих фильтров мультиплексируются и преобразуются в цифровую форму. Частота стробирования этих сигналов составляет примерно 50 Гц. Разрядность АЦП в этом случае может составлять 3 бита. На принимающей стороне осуществляется цифро-аналоговое

преобразование (ЦАП) и демультимплексирование. Сбалансированные амплитудные модуляторы, управляемые ЦАП и переключателем, выдают сигналы на узкополосные фильтры. Все эти сигналы смешиваются в сумматоре, а результат воспроизводится. В данном случае необходимое быстродействие передающей линии составляет $3 \text{ бита} \cdot 50 \text{ Гц} \cdot 16 \text{ каналов} = 2,4 \text{ кбит/с}$. Дальнейший выигрыш может быть получен за счет цифрового сжатия. Число каналов (фильтров) и ширина пропускаемой полосы частот может варьироваться, соответственно будет меняться и качество воспроизведения звука. Минимально возможная полоса пропускания передающей линии, при которой значение передаваемого текста еще воспринимается правильно, лежит ниже 1 кбит/с . В случае необходимости передачи индивидуальных особенностей голоса сначала должен проводиться анализ этих персональных отличий. Особенности голоса в закодированном виде передаются принимающей стороне, где эти данные используются в дальнейшем при воспроизведении закодированного текста.

ЛЕКЦИЯ 7

7.1. Типы сигналов импульсно-кодовой модуляции

Чтобы передать двоичные цифры по узкополосному каналу, их представляют электрическими импульсами. На рис. 7.1, а показаны разделенные во времени интервалы передачи кодовых слов, причем каждое кодовое слово является 4-битовым представлением квантованной выборки. На рис. 7.1, б каждая двоичная единица представляется импульсом, а каждый двоичный ноль – отсутствием импульса.

Задача приемника – определить в каждый момент приема бита имеется ли импульс в канале передачи. Вероятность точного определения наличия импульса является функцией энергии принятого импульса (или площади под графиком импульса). Следовательно, ширину импульса T' (рис. 7.1, б) выгодно делать как можно больше. Если увеличить ширину импульса до максимально возможного значения, равного времени передачи бита T , то получится сигнал, показанный на рис. 7.1, в. Вместо того, чтобы описывать сигнал как последовательность импульсов и их отсутствий (униполярное представление), можно описать его как последова-

тельность переходов между двумя ненулевыми уровнями (биполярное представление). Если сигнал находится на верхнем уровне напряжения, он представляет двоичную единицу, а если на нижнем — двоичный ноль.

Существует три способа модулирования информации в последовательность импульсов: можно варьировать амплитуду, положение или длительность импульсов, что дает соответственно следующие схемы: амплитудно-импульсная модуляция (pulse-amplitude modulation — PAM), фазово-импульсная модуляция (pulse-position modulation — PPM) и широтно-импульсная модуляция (pulse-width modulation — PWM). Фазово-импульсная модуляция осуществляется через задержку появления импульса на время, соответствующее значению информационных символов. Широтно-импульсная модуляция осуществляется посредством изменения ширины импульса на величину, соответствующую значению импульса. Для кодировок PPM и PWM амплитуда импульса фиксируется.

Существует несколько типов сигналов PCM. Они изображены на рис. 7.2. В приложениях телефонной связи эти сигналы часто называются кодами канала.

При цифровом кодировании дискретной информации применяют потенциальные и импульсные коды. В потенциальных кодах для представления логических единиц и нулей используется только значение потенциала сигнала, а его перепады, формирующие закон, импульсы, во внимание не принимаются. Импульсные коды позволяют представить двоичные данные либо импульсами определенной полярности, либо импульса — перепадом потенциала определенного направления.

При использовании прямоугольных импульсов для передачи дискретной информации необходимо выбрать такой способ кодирования, который одновременно достигал бы нескольких целей:

- имел при одной и той же битовой скорости наименьшую ширину спектра результирующего сигнала,
- обеспечивал синхронизацию между передатчиком и приемником,
- обладал способностью распознавать ошибки,
- обладал низкой стоимостью реализации.

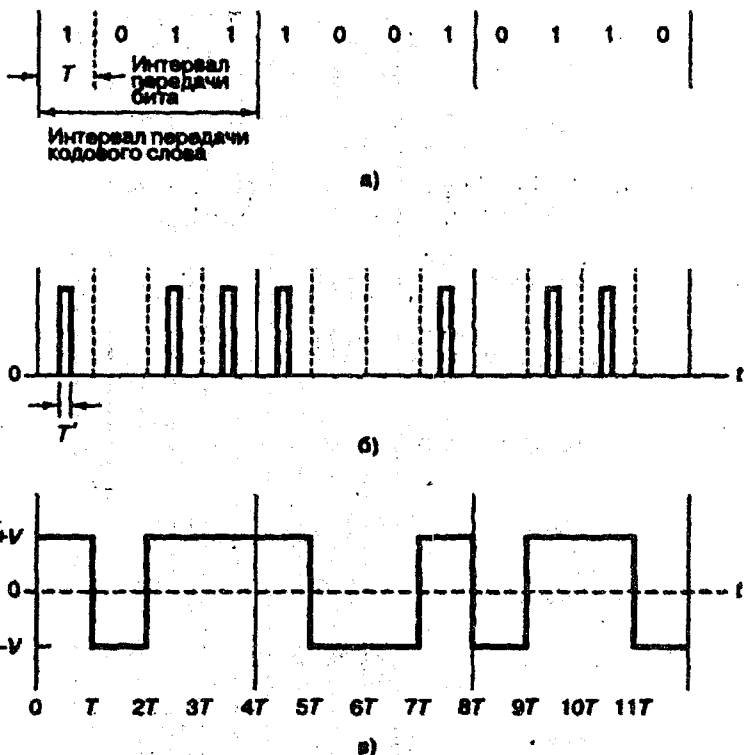


Рис. 7.1. Пример представления двоичных цифр в форме сигналов.

а) Последовательность РСМ.

б) Импульсное представление последовательности РСМ.

в) Импульсный сигнал (переход между двумя уровнями).

Более узкий спектр сигналов позволяет на одной и той же линии (с одной и той же полосой пропускания) добиваться более высокой передачи данных. Кроме того, часто к спектру сигнала предъявляется требование отсутствия постоянной составляющей, т. е. наличия постоянного тока между передатчиком и приемником. В частности, применение различных трансформаторных схем гальванической развязки препятствует прохождению постоянного тока.

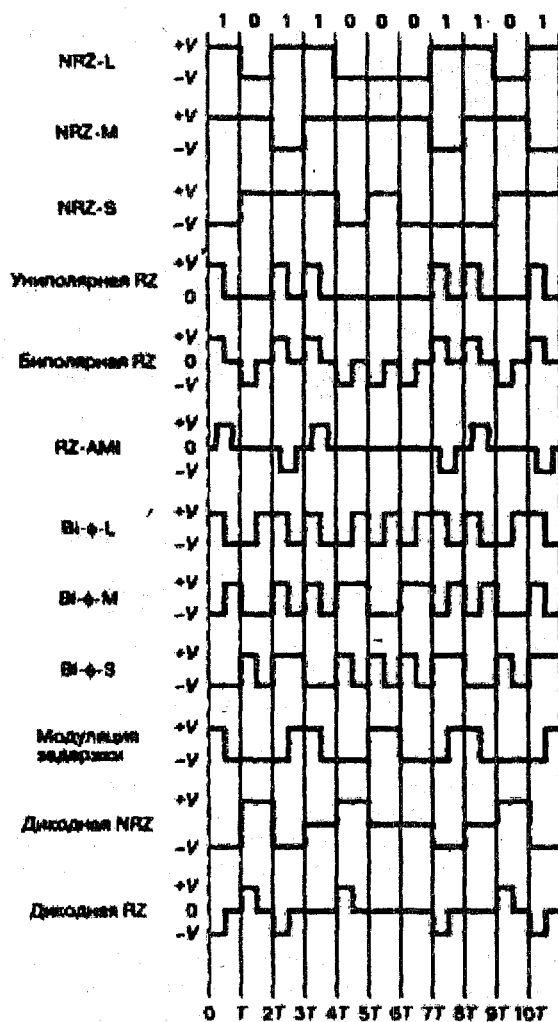


Рис. 7.2. Различные сигналы PCM.

Синхронизация передатчика и приемника нужна для того, чтобы приемник точно знал, в какой момент времени необходимо считывать новую информацию с линии связи. На небольших расстояниях хорошо работает схема, основанная на отдельной тактирующей линии связи, так что

информация снимается с линии только в момент прихода тактового импульса. Однако на больших расстояниях неоднородности характеристик канала связи приводят к неравномерности скорости распространения сигнала и, следовательно, к нарушению синхронизирующих свойств тактирующих импульсов. Кроме того, экономия проводников в дорогостоящих кабелях заставляет отказываться от использования дополнительной линии тактирующих импульсов. В этой связи важное значение имеет свойство самосинхронизации кодов канала, когда сигналы PCM несут для приемника указания о том, в какой момент времени нужно осуществлять распознавание очередного бита. Любой резкий перепад сигнала – фронт – может служить хорошим указанием для синхронизации.

Сигналы PCM делятся на четыре группы:

Без возврата к нулю (nonreturn – to – zero – NRZ)

С возвратом к нулю (return – to – zero – RZ)

Фазовое кодирование

Многоуровневое бинарное кодирование

7.2. Характеристики кодов канала

Группа кодировок NRZ включает следующие подгруппы: NRZ-L (L – level – уровень), NRZ-M (M-mark-метка), NRZ-S (S-space – пауза). Кодировка NRZ-L широко используется в цифровых логических схемах. Двоичная единица в этом случае представляется одним уровнем напряжения, а двоичный ноль – другим. При использовании кодировки NRZ-M двоичная единица (или метка – mark) представляется изменением уровня, а ноль или пауза (space) – отсутствием изменения уровня. Такая кодировка называется *дифференциальной*. Применяется кодировка NRZ-M преимущественно при записи на магнитную ленту. Кодировка NRZ-S является обратной по отношению к кодировке NRZ-M: двоичная единица представляется отсутствием изменения уровня, а двоичный ноль – изменением уровня.

Метод NRZ прост в реализации, обладает хорошей распознаваемостью ошибок (из-за двух резко отличающихся потенциалов), но не обладает свойством самосинхронизации. При передаче длинной последовательности единиц или нулей сигнал на линии не изменяется, поэтому приемник лишен возможности определять по входному сигналу моменты времени, когда нужно в очередной раз считывать данные. Другим серьезным недостатком метода NRZ является наличие низкочастотной составляющей, которая приближается к нулю при передаче длинной последовательности единиц или нулей. Большин-

ство линий связи сопрягаются с аппаратурой через реактивные элементы, например через трансформаторы. Поскольку униполярные сигналы всегда содержат в своем спектре постоянную составляющую и значительную долю низкочастотных компонентов при передаче длинной последовательности нулей или единиц, такое сопряжение затруднено или вовсе невозможно – реактивные элементы на этих частотах представляют собой почти “обрыв” или почти “короткое замыкание”. Из-за этого многие каналы связи, не обеспечивающие прямого гальванического соединения между приемником и передатчиком, этот вид кодирования не поддерживают. Привлекательность кода NRZ, из-за которой имеет смысл заняться его улучшением, состоит в достаточно низкой частоте основной гармоники $f_0 = \frac{R}{2}$ Гц,

где R – битовая скорость передачи данных.

Группа кодировок RZ включает униполярную кодировку RZ, биполярную кодировку RZ и кодировку RZ-AMI. В униполярной кодировке RZ единица представляется наличием импульса, длительность которого равна половине ширины бита, а ноль – его отсутствием. В биполярной кодировке RZ единицы и нули представляются импульсами противоположных уровней, длительность каждого из которых также составляет половину ширины бита. В каждом интервале передачи бита присутствует импульс.

В методе биполярного кодирования с альтернативной инверсией (**Bipolar Alternate Mark Inversion, AMI**) используется три уровня потенциала – отрицательный, нулевой и положительный. Для кодирования логического нуля используется нулевой потенциал, а логическая единица кодируется либо положительным потенциалом, либо отрицательным, при этом потенциал каждой новой единицы противоположен потенциалу предыдущей. Код AMI частично ликвидирует проблемы постоянной составляющей и отсутствия самосинхронизации, присущие коду NRZ. Это происходит при передаче длинных последовательностей единиц. В этих случаях сигнал на линии представляет собой последовательность разнополярных импульсов с тем же спектром, что и у кода NRZ и с основной гармоникой $f_0 = \frac{R}{2}$ Гц. Длинные же последовательности нулей так же

опасны для кода AMI, как и для кода NRZ. В целом для различных комбинаций бит на линии использование кода AMI приводит к более узкому спектру сигнала, чем для кода NRZ, а значит, к более высокой пропускной способности линии. Например, при передаче чередующихся нулей и

единиц основная гармоника $f_0 = \frac{R}{4}$ Гц. Код АМІ предоставляет некото-

рые возможности по распознаванию ошибочных сигналов. Так, нарушение строгого чередования полярности сигналов говорит о ложном импульсе. Три уровня сигнала в коде АМІ требуют дополнительного увеличения мощности передатчика на 3 дБ для обеспечения той же достоверности приема сигналов, что и для кодов с двумя состояниями.

Существует код, похожий на АМІ, но с двумя уровнями сигнала. При передаче нуля он передает потенциал, который был установлен в предыдущем такте (т. е. не меняет его), а при передаче единицы потенциал инвертируется на противоположный. Этот код называется потенциальным кодом с инверсией при единице (Non Return to Zero with ones Inverted, NRZI).

Сигналы в дикодной кодировке для кодирования двоичных данных также используют три уровня сигнала. В этом случае переходы в передаваемой информации от единицы к нулю и от нуля к единице меняют полярность импульсов. При отсутствии переходов передается сигнал нулевого уровня.

Группа фазового кодирования включает в себя следующие кодировки: bi- ϕ -L (двухфазный уровень) – более известная как *манчестерское кодирование*, bi- ϕ -M, bi- ϕ -S и модуляция задержки (delay modulation – DM или кодировка Миллера).

Рассмотрим подробнее **манчестерское кодирование**.

При использовании манчестерского кодирования единица кодируется отрицательным перепадом сигнала в середине битового интервала, нуль – положительным перепадом. На границах битовых интервалов сигнал, если это необходимо, меняет значение, готовясь к отображению следующего битового интервала. Поскольку сигнал не “замирает” на одном логическом уровне, постоянная составляющая равна нулю. Подстройка системы синхронизации осуществляется при передаче каждого бита. Критерием ошибки в канале связи может являться “замирание” сигнала на одном из логических уровней на время, превышающее время передачи одного информационного бита. Платой за эти полезные качества является удвоение требуемой пропускной способности канала связи.

На рис. 7.3 приведены схемы кодирования и декодирования сигнала кодом «Манчестер II», а на рис. 7.4 изображены временные диаграммы сигналов в различных точках схем.

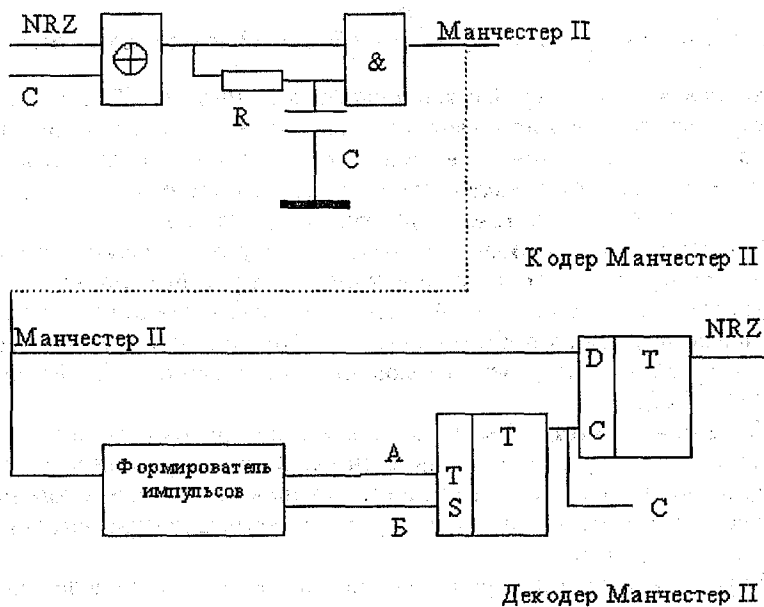


Рис. 7.3. Схема кодирования и декодирования кода "Манчестер II".

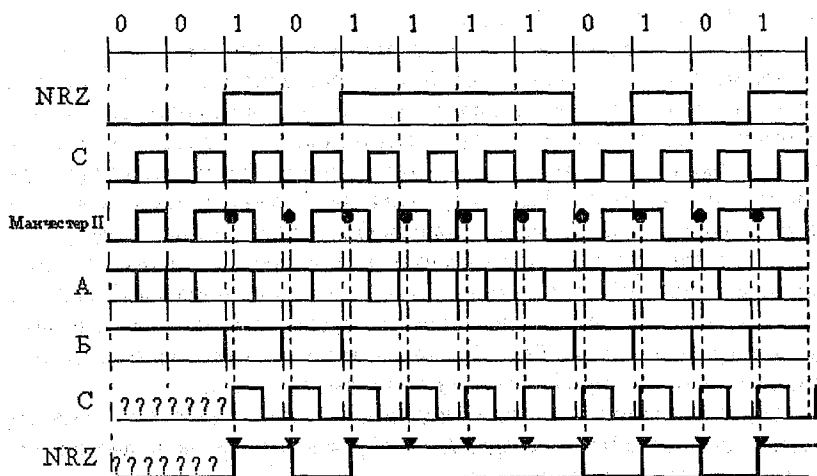


Рис. 7.4 Временные диаграммы кодирования и декодирования для кода "Манчестер II".

Сигнал в коде "Манчестер – II" может быть получен суммированием по модулю двух сигналов: NRZ и С (сигнал синхронизации). Другими словами, сигнал в коде "Манчестер II" принимает единичные значения в тех интервалах времени, в которых сигналы NRZ и С имеют противоположные логические значения. Вследствие этого схема шифратора очень проста и легко может быть дополнена схемой подавления помех, предназначенной для подавления помех, возникающих из-за несовпадения фронтов суммируемых сигналов.

Схема дешифрации кода "Манчестер II" несколько сложнее и содержит формирователь импульсов из входного потока информации и два триггера, счетный и D-триггер для декодирования информации. В начальный момент выходной код не определен, но по приходу первого тактового импульса, соответствующего переходу из логического нуля в логическую единицу во входном коде NRZ, система начинает генерировать правильный выходной код NRZ, декодированный из входной последовательности "Манчестер II" и совпадающий с исходным кодом. Дополнительно декодер восстанавливает и сигнал синхронизации, что упрощает подстройку тактового генератора приемника.

Таким образом, можно сформулировать основные преимущества кода "Манчестер II" перед кодом NRZ в дополнение к тому, что он устраняет недостатки кода NRZ, перечисленные выше.

- Синхросигнал и информация передаются по одному каналу связи, в то время как кодирование NRZ требует двух линий передачи.

- Диапазон частот при использовании кодирования "Манчестер II"

состоит из двух составляющих – $\frac{R}{2}$ и R Гц, в отличие от кодирования NRZ, в выходном спектре которого находятся частоты от нуля (постоянная составляющая) до $\frac{R}{2}$ Гц. Это позволяет, кроме всего прочего,

сделать приемник кода "Манчестер II" более узкополосным и, следовательно, более помехоустойчивым. Легко реализуется и трансформаторная связь с магистралью. Поскольку передатчик импульсный, можно повысить его мощность, используя его с двукратной перегрузкой (примерно 50 % времени он находится в выключенном состоянии).

Из недостатков можно отметить наличие схем шифрации и дешифрации, удвоенную по сравнению с кодированием NRZ, верхнюю частоту в канале связи и необходимость предустановки приемника посылкой перехода из нуля в единицу.

7.3. Логическое кодирование для улучшения характеристик потенциальных кодов

Для улучшения характеристик потенциальных кодов используется два метода.

Первый метод — избыточные коды — основан на добавлении в исходный код избыточных бит, содержащих логические единицы. В этом случае длинные последовательности нулей прерываются и код становится самосинхронизирующимся для любых передаваемых данных. Исчезает также постоянная составляющая, а значит, еще более сужается спектр сигнала. Однако этот метод снижает полезную пропускную способность линии связи, так как избыточные единицы полезной информации не несут.

При применении избыточных кодов исходная последовательность бит разбивается на порции — символы. Затем каждый исходный символ заменяется на новый, который имеет большее количество бит, чем исходный. Например, логический код 4В/5В, используемый в технологиях FDDI и Fast Ethernet, заменяет исходные символы длиной в 4 бит на символы длиной в 5 бит. Так как результирующие символы содержат избыточные биты, то общее количество битовых комбинаций в них больше, чем в исходных. Так, в коде 4В/5В результирующие символы могут содержать 32 битовых комбинации, в то время как исходные символы — только 16. Поэтому в результирующем коде можно отобрать 16 таких комбинаций, которые не содержат большого количества нулей, а остальные считать запрещенными кодами. Кроме устранения постоянной составляющей и придания коду свойства самосинхронизации, избыточные коды позволяют приемнику распознавать искаженные биты. Если приемник принимает запрещенный код, значит, на линии произошло искажение сигнала. Символы кода 4В/5В длиной 5 бит гарантируют, что при любом их сочетаний на линия не могут встретиться более трех нулей подряд. Буква В в названии кода означает, что элементарный сигнал имеет 2 состояния — от английского binary — двоичный. Имеются также коды и с тремя состояниями сигнала, например, в коде 8В/6Т для кодирования 8 бит исходной информации используется код из 6 сигналов, каждый из которых имеет три состояния. Избыточность кода 8В/6Т выше, чем кода 4В/5В, так как на 256 исходных кодов приходится $3^6=729$ результирующих символов.

Для обеспечения заданной пропускной способности линии передатчик, использующий избыточный код, должен работать с повышенной тактовой частотой. Так, для передачи кодов 4В/5В со скоростью 100 Мбит/с передатчик должен работать с тактовой частотой 125 МГц. При этом спектр сигнала на линии расширяется по сравнению со случаем, когда на линии передается чистый, не избыточный код. Тем не менее спектр избыточного потенциального кода оказывается уже спектра манчестерского

кода, что оправдывает дополнительный этап логического кодирования, а также работу приемника и передатчика на повышенной тактовой частоте.

Второй метод улучшения свойств потенциальных кодов — скремблирование.

Перемешивание данных скремблером перед передачей их в линию с помощью потенциального кода является другим способом логического кодирования.

Методы скремблирования заключаются в побитовом вычислении результирующего кода на основании бит исходного кода и полученных в предыдущих тактах бит результирующего кода. Например, скремблер может реализовывать следующее соотношение:

$$B_i = A_i \oplus B_{i-3} \oplus B_{i-5}, \quad (7.1)$$

где B_i — двоичная цифра результирующего кода, полученная на i -м такте работы скремблера; A — двоичная цифра исходного кода, поступающая на i -м такте на вход скремблера; B_{i-3} и B_{i-5} — двоичные цифры результирующего кода, полученного на предыдущих тактах работы скремблера, соответственно на 3 и 5 тактов ранее текущего такта; $\oplus$ — операция исключающего ИЛИ (сложение по модулю 2).

Например, для исходной последовательности 110110000001 скремблер даст следующий результирующий код:

$B_1 = A_1 = 1$ (первые три цифры результирующего кода будут совпадать с исходным, так как еще нет нужных предыдущих цифр)

$$B_2 = A_2 = 1$$

$$B_3 = A_3 = 0$$

$$B_4 = A_4 \oplus B_1 = 1 \oplus 1 = 0$$

$$B_5 = A_5 \oplus B_2 = 1 \oplus 1 = 0$$

$$B_6 = A_6 \oplus B_3 \oplus B_1 = 0 \oplus 0 \oplus 1 = 1$$

$$B_7 = A_7 \oplus B_4 \oplus B_2 = 0 \oplus 0 \oplus 1 = 1$$

$$B_8 = A_8 \oplus B_5 \oplus B_3 = 0 \oplus 0 \oplus 0 = 0$$

$$B_9 = A_9 \oplus B_6 \oplus B_4 = 0 \oplus 1 \oplus 0 = 1$$

$$B_{10} = A_{10} \oplus B_7 \oplus B_5 = 0 \oplus 1 \oplus 0 = 1$$

$$B_{11} = A_{11} \oplus B_8 \oplus B_6 = 0 \oplus 0 \oplus 1 = 1$$

$$B_{12} = A_{12} \oplus B_9 \oplus B_7 = 1 \oplus 1 \oplus 1 = 1$$

Таким образом, на выходе скремблера появится последовательность 110001101111, в которой нет последовательности из шести нулей, присутствовавшей в исходном коде.

После получения результирующей последовательности приемник передает ее дескремблеру, который восстанавливает исходную последовательность на основании обратного соотношения:

$$C_i = B_i \oplus B_{i-3} \oplus B_{i-5} = (A_i \oplus B_{i-3} \oplus B_{i-5}) \oplus B_{i-3} \oplus B_{i-5} = A_i \quad (7.2)$$

Различные алгоритмы скремблирования отличаются количеством слагаемых, дающих цифру результирующего кода, и сдвигом между слагаемыми. Так, в сетях ISDN при передаче данных от сети к абоненту используется преобразование со сдвигами в 5 и 23 позиции, а при передаче данных от абонента в сеть — со сдвигами 18 и 23 позиции.

Существуют и более простые методы борьбы с последовательностями единиц, также относимые к классу скремблирования.

Для улучшения кода Bipolar AMI используются два метода, основанные на искусственном искажении последовательности нулей запрещенными символами.

На рис. 7.5 показано использование метода B8ZS (Bipolar with 8 Zeros Substitution) и метода HDB3 (High Density Bipolar 3 Zeros) для корректировки кода AMI.

Исходный код состоит из двух длинных последовательностей нулей. В первом случае — из 8, а во втором — из 5.

Код B8ZS исправляет только последовательности, состоящие из 8 нулей. Для этого он после первых трех нулей вместо оставшихся пяти нулей вставляет пять цифр: $V-1^*-0-V-1^*$. Здесь V обозначает сигнал единицы, запрещенной для данного такта полярности, т. е. сигнал, не изменяющий полярность предыдущей единицы, 1^* — сигнал единицы корректной полярности, а знак звездочки отмечает тот факт, что в исходном коде в этом такте была не единица, а ноль. В результате на 8 тактах приемник наблюдает 2 искажения — очень маловероятно, что это случилось из-за шума на линии или других сбоев передачи. Поэтому приемник считает такие нарушения кодировкой 8 последовательных нулей и после приема заменяет их на исходные 8 нулей. Код B8ZS построен так, что его постоянная составляющая равна нулю при любых последовательностях двоичных цифр.

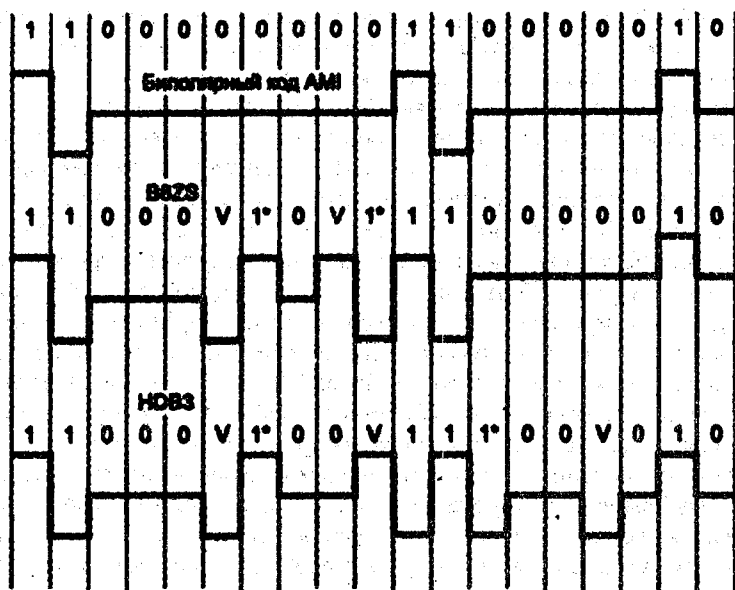


Рис. 7.5. Коды B8ZS и HDB3. V – сигнал запрещенной полярности;
1* – сигнал единицы корректной полярности, но заменившей ноль в исходном коде.

Код HDB3 исправляет любые четыре подряд идущих нуля в исходной последовательности. Правила формирования кода HDB3 более сложные, чем кода B8ZS. Каждые четыре нуля заменяются четырьмя сигналами, в которых имеется один сигнал V. Для подавления постоянной составляющей полярность сигнала V чередуется при последовательных заменах. Кроме того, для замены используются два образца четырехтактных кодов. Если исходный код содержал не четное число единиц, то используется последовательность 000V, а если число единиц было четным — последовательность 1*00V.

7.4. Спектральные параметры сигналов PCM

Наиболее распространенными критериями, используемыми при сравнении кодировок PCM и выборе подходящего типа сигнала из многих доступных, являются спектральные характеристики, возможности битовой синхронизации и выявления ошибок, устойчивость к интерференции и помехам, а также цена и сложность реализации. Спектральные характеристики некоторых распространенных кодировок PCM показаны

на рис. 7.6. Здесь изображена зависимость спектральной плотности мощности (измеряется в Вт/Гц) от нормированной ширины полосы, WT где W — ширина полосы, а T — длительность импульса.

Произведение WT часто называют базой сигнала. Поскольку скорость передачи импульсов или сигналов R_s , обратна T , нормированную ширину полосы можно также выразить как W/R_s . Из последнего выражения видно, что нормированная ширина полосы измеряется в герц/(импульс/с) или в герц/(символ/с). Это относительная мера ширины полосы; она описывает, насколько эффективно используется полоса пропускания при интересующей нас кодировке. Считается, что любой тип кодировки, требующий менее 1,0 Гц для передачи одного символа в секунду, эффективно использует полосу. Например, модулирование задержки и двубинарная кодировка. Любая кодировка, требующая более 1,0 Гц полосы для передачи одного символа в секунду, неэффективно использует полосу. Пример: двухфазная (манчестерская) кодировка. На рис. 7.6 можно также видеть распределение энергии сигналов в различных кодировках по спектру. Например, двубинарная кодировка и схема NRZ имеют значительное число спектральных компонентов около постоянной составляющей и на низких частотах, тогда как двухфазная кодировка вообще не содержит энергии на частоте постоянной составляющей.

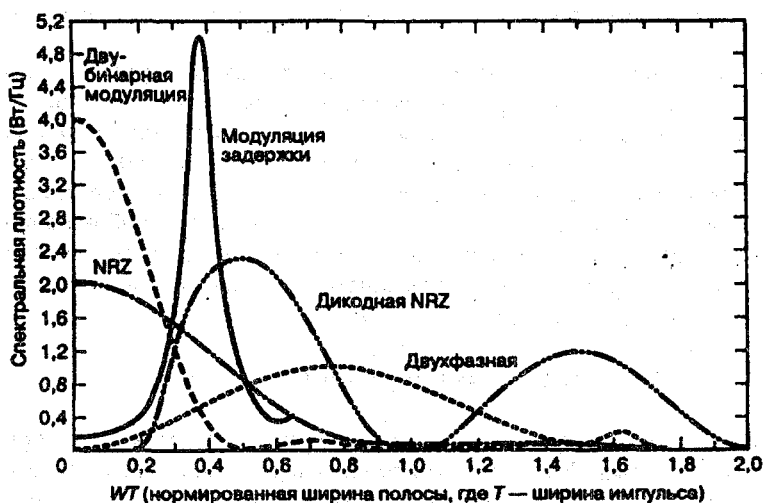


Рис. 7.6. Спектральные плотности различных кодировок PCM.

Важным параметром измерения эффективности использования полосы является отношение R/W (измеряется в бит/с/герц). Эта мера характеризует скорость передачи данных, а не скорость передачи сигналов. Для данной схемы передачи сигналов отношение R/W описывает, какой объем данных может быть передан из расчета на каждый герц доступной полосы.

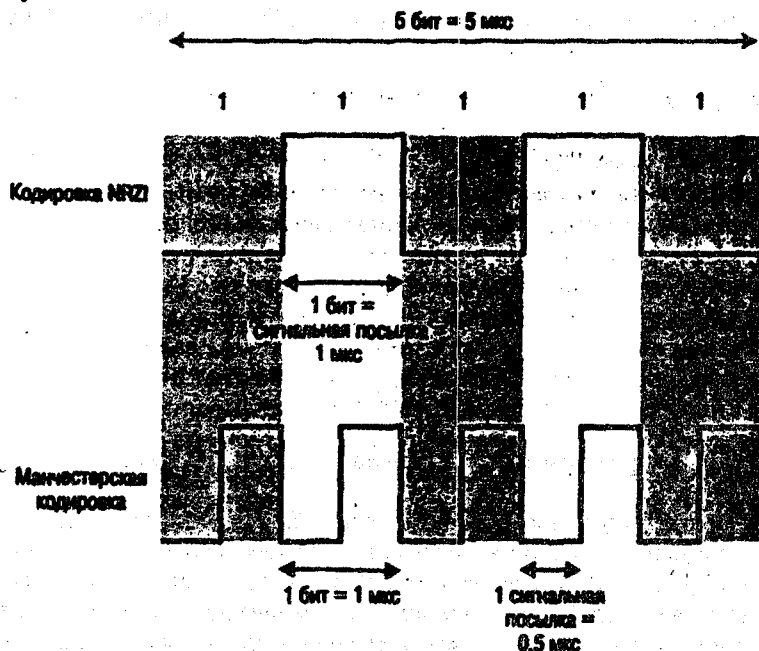


Рис. 7.7. Поток двоичных единиц при скорости 1 Мбит/с.

При использовании кодированных сигналов необходимо различать скорость передачи данных (выражается в битах в секунду) и скорость модуляции (выражается в бодах). Скорость передачи данных, или скорость передачи битов, равна $R = \frac{1}{t_b}$ где t_b — длительность бита. Скорость

модуляции — это скорость генерации сигнальных посылок. Рассмотрим, например, манчестерскую кодировку. Сигнальная посылка минимального размера представляет собой импульс, продолжительность которого равна половине длительности бита. Для строки, состоящей только из двоичных нулей или только из двоичных единиц, генери-

руется постоянный поток, состоящий из таких же импульсов. Следовательно, максимальная скорость модуляции при манчестерской кодировке равна $\frac{2}{t_b}$. Данная ситуация иллюстрируется на рис. 7.7, на котором

показана передача потока двоичных единиц при скорости передачи данных 1 Мбит/с с использованием кодировки NRZI и манчестерской кодировки.

$$\text{Вообще } D = \frac{R}{b}, \quad (7.3)$$

где D – скорость модуляции (бод),

R – скорость передачи данных, бит/с,

b – число битов в сигнальной посылке.

ЛЕКЦИЯ 8

8.1. Причины применения полосовой модуляции

Цифровая модуляция – это процесс преобразования цифровых символов в сигналы, совместимые с характеристиками канала. При узкополосной модуляции (baseband modulation) эти сигналы обычно имеют вид импульсов заданной формы. При полосовой модуляции (bandpass modulation) импульсы заданной формы модулируют синусоиду, называемую несущей волной (carrier frequency), затем следует передача на нужное расстояние с использованием радиочастот. Для этого несущая преобразовывается в электромагнитное поле. Несущая для передачи узкополосных сигналов нужна по следующим причинам:

Передача электромагнитного поля через пространство выполняется с помощью антенн. Размер антенны зависит от длины волны λ и текущей задачи. Для переносных телефонов размер антенны обычно равен $\lambda/4$, а длина волны c/f , где $c = 3 \cdot 10^8$ м/с – скорость света. Рассмотрим передачу узкополосного сигнала (скажем, имеющего частоту $f = 3000$ Гц) путем сопряжения его непосредственно с антенной без использования несущей. Какая антенна нам понадобится? Возьмем стандарт телефонной промышленности, $\lambda/4$. Получаем, что для узкопо-

лосного сигнала $\lambda = \frac{c}{f} = \frac{3 \cdot 10^8}{3 \cdot 10^3} = 10^5$ м. Следовательно, длина антенны

$\frac{\lambda}{4} = \frac{10^5}{4} = 25$ км. Если же узкополосная информация модулируется несущей более высокой частоты, например 900 МГц, размер антенны будет составлять $\lambda = \frac{c}{f} = \frac{3 \cdot 10^8}{9 \cdot 10^8} = 0,33 \Rightarrow \frac{\lambda}{4} = 0,083$ м ≈ 8

см. Таким образом, приведенные расчеты показывают, что *модулирование несущей частоты, или полосовая модуляция, — это этап, необходимый для всех систем, использующих радиопередачу.*

Полосовая модуляция имеет и другие важные преимущества при передаче сигналов. При использовании канала более чем одним сигналом, модуляция может применяться для выделения различных сигналов. Этот метод называется уплотнением с частотным разделением. Модуляция может использоваться для преодоления последствий межсимвольной интерференции (intersymbol interference — ISI). Вследствие системной фильтрации принятые импульсы могут перекрываться (как это показано на рис. 8.1), а вследствие искажения при проходе по каналу связи импульсы могут «размываться» на соседние интервалы передачи символов, таким образом мешая процессу обнаружения.

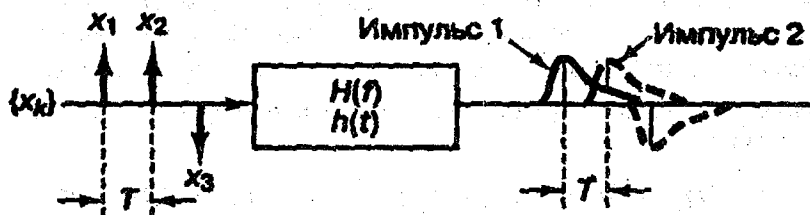


Рис. 8.1 Межсимвольная интерференция.

Класс схем модулирования, известный как модулирование расширенным спектром, требует полосы, значительно превышающей минимальную полосу, необходимую для передачи сообщения.

Кроме того, модуляция может использоваться для перемещения сигнала в диапазон частот, в котором легко удовлетворяются специфици-

ческие конструктивные требования, например относящиеся к фильтрации и усилению.

Полосовая модуляция – это процесс преобразования информационного сигнала в синусоидальную волну. При цифровой модуляции синусоида на интервале T называется цифровым символом. Синусоиды могут отличаться по амплитуде, частоте и фазе. Таким образом, полосовую модуляцию можно определить как процесс варьирования амплитуды, частоты или фазы (или их комбинаций) радиочастотной несущей согласно передаваемой информации.

8.2. Амплитудная модуляция

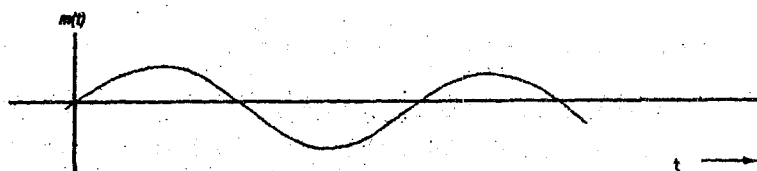
Амплитудная модуляция (АМ) представляет собой простейшую форму модуляции и изображена на рис. 8.2.

Математически данный процесс можно описать следующим образом:

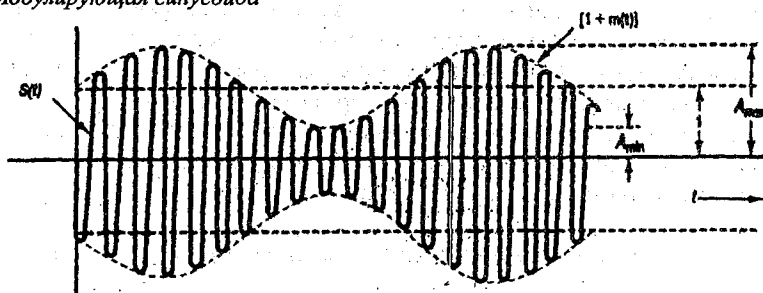
$$s(t) = [1 + nx(t)] \cos 2\pi f_c t, \quad (8.1)$$

где $\cos 2\pi f_c t$ – несущая; $x(t)$ – информационный сигнал, оба приведенные к единичной амплитуде. Параметр n называется коэффициентом модуляции и представляет собой отношение амплитуд информационного и несущего сигналов. Входной сигнал равен $m(t) = nx(t)$. Единица в формуле (8.1) представляет собой постоянную составляющую, предупреждающую потерю информации. Огибающая результирующего сигнала на рис. 8.2, б равна $[1 + nx(t)]$, поскольку $n < 1$, то огибающая точно воспроизводит исходный сигнал. Если $n > 1$, то огибающая пересечет временную ось, и информация будет потеряна. Данная схема также известна как двухполосная связь с передачей несущей (DSBTC – double sideband transmitted carrier). Если принять $x(t) = \cos 2\pi f_u t$, то из выражения (8.1), используя тригонометрические преобразования, можно получить:

$$s(t) = \cos 2\pi f_c t + \frac{n}{2} \cos 2\pi(f_c - f_u)t + \frac{n}{2} \cos 2\pi(f_c + f_u)t. \quad (8.2)$$



а) Модулирующая синусоида



б) Результирующий амплитудно-модулированный сигнал

Рис. 8.2. Амплитудная модуляция.

То есть сигнал в амплитудной модуляции состоит из гармоник исходной несущей частоты и двух гармоник, смещенных относительно несущей на f_u Гц (рис. 8.3). Часть спектра $|f| > |f_c|$ называется верхней боковой полосой, а часть спектра $|f| < |f_c|$ – нижней боковой полосой. И нижняя, и верхняя боковые полосы точно копируют спектр исходного сигнала, но в нижней боковой полосе частоты обращены.

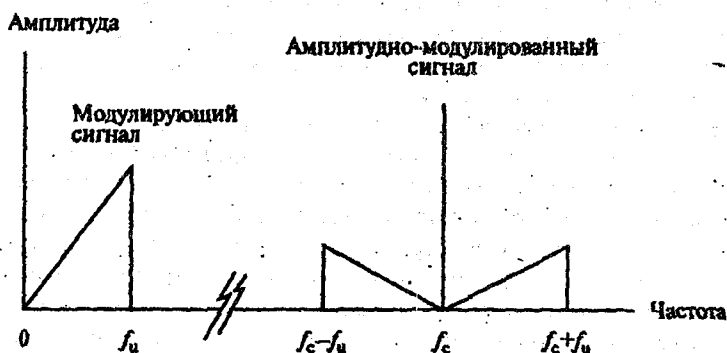


Рис. 8.3. Спектр амплитудно-модулированного сигнала.

$$\text{Отметим важное соотношение } P_i = P_c \left(1 + \frac{n^2}{2}\right), \quad (8.3)$$

где P_i – общая переданная мощность сигнала $s(t)$, а P_c – мощность, переданная на несущей частоте. Хотелось бы, чтобы n было максимально большим, чтобы большая часть мощности сигнала использовалась для передачи информации, но все же должно выполняться неравенство $n < 1$.

Модуляция амплитуды осуществляется как линейными, так и нелинейными методами. В основе нелинейных методов лежит использование активных устройств с нелинейными передаточными характеристиками, что, однако, приводит к гармоническим искажениям модулированного сигнала. В линейных же методах, наоборот, получается модулированный сигнал большой мощности с меньшими искажениями.

8.3. Демодуляция амплитудно-модулированных сигналов

Если умножить амплитудно-модулированный сигнал на его составляющую несущей частоты, то часть полученных спектральных составляющих будет содержать исходный информационный сигнал. Одним из способов подобного перемножения является применение устройства с квадратичной характеристикой. Наличие квадратичной составляющей передаточной функции приводит к перемножению компоненты несущей частоты с соответствующей компонентой модулированного сигнала. В результате появляется немодулированная компонента. Частоте не требуется никаких компонент модулирующего сигнала, кроме компоненты несущей.

Наиболее часто для демодуляции слабых амплитудно-модулированных сигналов в качестве устройства с квадратичной передаточной характеристикой используется полупроводниковый диод. В этом случае диод работает в режиме малых токов, т.е. все время остается в открытом состоянии.

На рис. 8.4 показана схема, которая используется для реализации другого метода демодуляции амплитудно-модулированных сигналов. В этом случае диод работает в режиме больших токов. Входное напряжение растягивается в соответствии с передаточной функцией таким образом, что диод начинает непрерывно переключаться из проводящего состояния в непроводящее, и наоборот. Схема реагирует на огибающую (образованную положительными или отрицательными полупериодами, в зависимости от включения диода) принятого сигнала и поэтому носит название детектора огибающей.

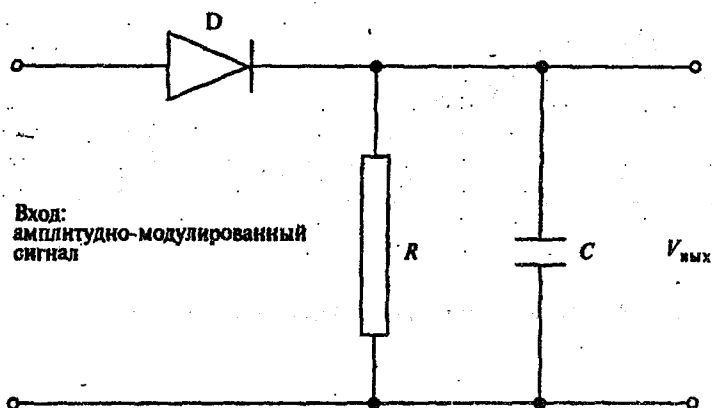


Рис. 8.4. Диодный детектор.

Опишем работу такой схемы, считая для упрощения, что диод идеальный и его сопротивление бесконечно в режиме отсечки и равно нулю в открытом состоянии. В течение положительного полупериода напряжения модулированного сигнала диод открыт и напряжение V_k на конденсаторе равняется $V_{вх}$. Постоянная времени RC специально подбирается достаточно большой по сравнению с периодом модулированного сигнала. Когда напряжение на входе схемы начинает падать, конденсатор не может разрядиться через закрытый диод, а благодаря большой величине постоянной времени напряжение на нем не может повторять форму входного напряжения. В результате конденсатор сравнительно медленно разряжается через сопротивление R . График напряжения изображен на рис. 8.5.

Результатом последующей низкочастотной фильтрации, которая убирает высокочастотную рябь, является демодулированный (восстановленный) сигнал. Его форма с большой степенью точности повторяет форму огибающей, которая фактически эквивалентна модулирующему сигналу. Однако, если постоянная времени выбрана слишком большой, мы столкнемся с явлением, которое называется срезанием отрицательных пиков и приводит к искажениям сигнала (рис. 8.6).

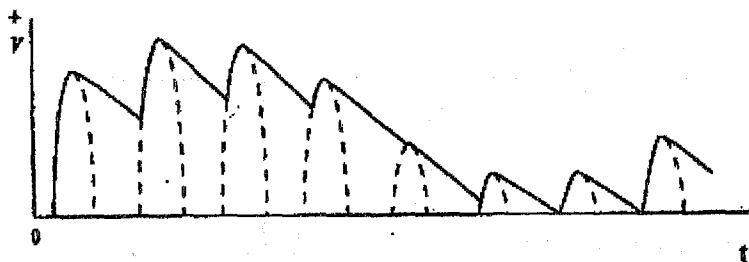


Рис. 8.5. Напряжение на выходе диодного детектора.

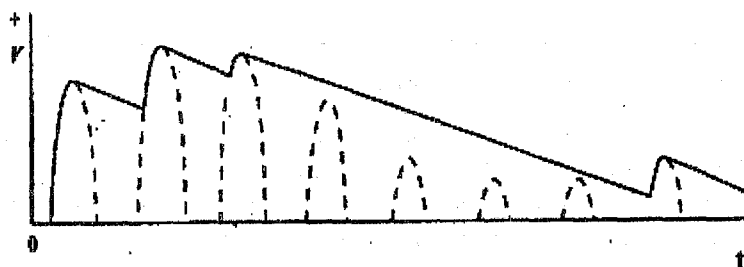


Рис. 8.6. Срезание отрицательных пиков.

Значение постоянной времени, необходимое для нормальной работы демодулятора, зависит от скорости изменения формы огибающей принятого сигнала. Эта скорость, в свою очередь, определяется глубиной модуляции n и частотой модулирующего сигнала. Величина этой постоянной в идеальном случае должна равняться

$$RC \leq \frac{\sqrt{(1-n)^2}}{n\omega_u}, \quad (8.4)$$

где $\omega_u = 2\pi f_u$ — круговая частота.

Это соотношение в целом справедливо для небольших значений n .

8.4. Балансная модуляция

Амплитудная модуляция привлекательна прежде всего своей простотой и, следовательно, дешевизной, которую обеспечивает использование в приемнике диодного детектора. Однако большая доля передаваемой

мощности приходится на несущую частоту, т.е. расходуется зря, так как несущая сама по себе не несет полезной информации. По этой причине были разработаны другие виды амплитудной модуляции. В одном из них используется балансный модулятор. Такая модуляция характеризуется полным подавлением несущей и наличием в спектре только боковых частот. Передача с помощью балансной модуляции получила название передачи с двумя боковыми и подавлением несущей (double sideband suppressed carrier — DSBSC).

Один из методов осуществления этого вида АМ заключается в использовании модулятора Коуэна, схема которого приведена на рис. 8.7. В течение одного полупериода несущей в точке А схемы создается положительное напряжение, а в точке В — отрицательное, что приводит к смещению всех четырех диодов. При этом обмотки двух трансформаторов, подключенных к диодному мосту, замыкаются накоротко, и на выходе мы получаем нулевое напряжение. В течение 2-го полупериода диоды смещаются в обратном направлении, т.е. закрываются; диодный мост размыкается, и модулирующий сигнал проходит на выход схемы. Выходной сигнал эквивалентен произведению модулирующего сигнала на положительный период высокочастотной несущей, как показано на рис. 8.8.

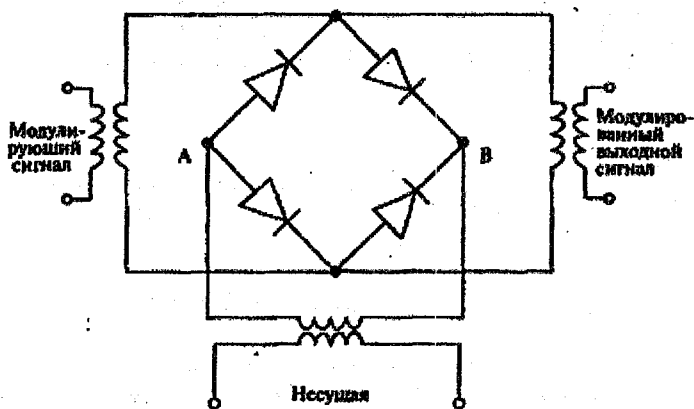


Рис. 8.7. Электрическая схема модулятора Коуэна.

Для того чтобы проанализировать работу модулятора Коуэна, рассмотрим умножение модулирующего сигнала на несущую, которая представляется в виде разложения в ряд Фурье:

$$g(t) = kV_u \cos \omega_u t \left[\frac{1}{2} + \frac{2}{\pi} (\sin \omega_c t + \frac{1}{3} \sin 3\omega_c t + \dots) \right] =$$

$$= \frac{kV_u \cos \omega_u t}{2} + \frac{kV_u}{\pi} [\sin(\omega_c - \omega_u)t + \sin(\omega_c + \omega_u)t] + \dots \quad (8.5)$$

где k — постоянная схемы.

Уравнение (8.5) содержит исходный сигнал, составляющие верхней и нижней боковых частот, которые и представляют сигнал с подавлением несущей, а также частотные компоненты высших порядков с центрами на частотах нечетных гармоник, кратных частоте ω_c . Как видно, в полученной функции отсутствует компонента с частотой несущей. После полосовой фильтрации в спектре сигнала остаются только требуемые частотные компоненты, все остальные подавлены.

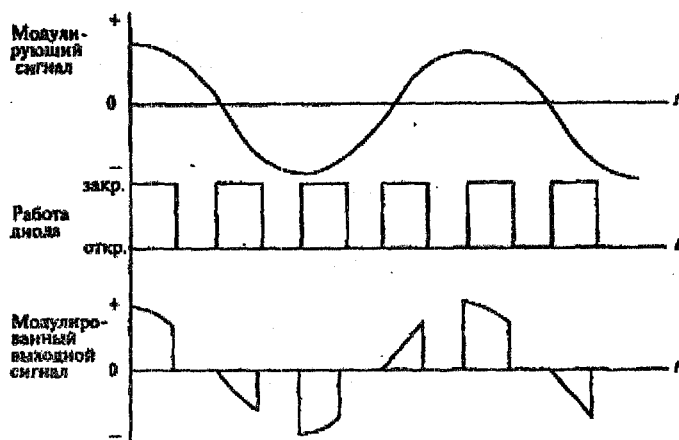


Рис. 8.8. Выходная функция модулятора Коузена.

На рис. 8.9 показан еще один тип модулятора — так называемый кольцевой модулятор, состоящий из тех же четырех диодов, но соединенных в виде кольца. В зависимости от полуцикла несущей в каждый момент времени открыта только одна пара диодов: например, D1 и D2 открыты, а D3 и D4 в этот момент закрыты, и наоборот. В результате сигнал на выходе равен либо положительному мгновенному значению модулирующего сигнала, если открыты диоды D1 и D2, либо отрицательному, если открыты D3 и D4 (рис. 8.10)). Схема с кольцевым расположением

диодов работает как переключатель полярности с частотой переключения, равной частоте несущей. Анализ работы такого модулятора аналогичен анализу схемы Коуэна, но несущий сигнал здесь является биполярным, а не униполярным, как в схеме Коуэна. Мы опять получаем сигнал с двумя спектральными составляющими и подавлением несущей, но теперь, вследствие биполярности несущей, энергия его в два раза больше.

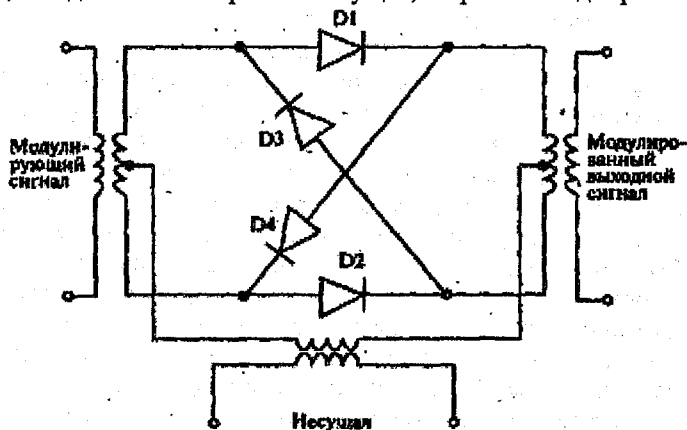


Рис. 8.9. Кольцевой модулятор – электрическая схема.

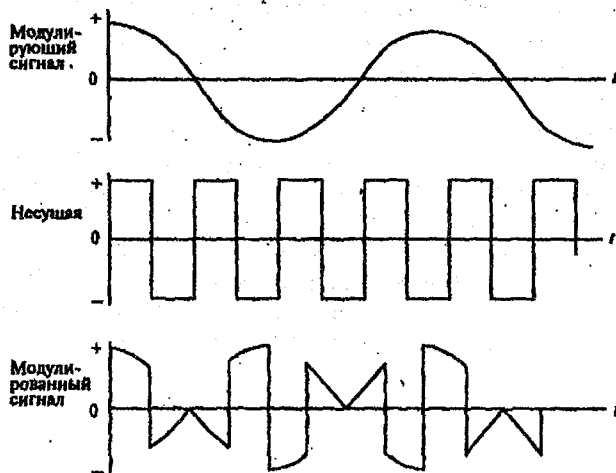


Рис. 8.10. Выходная функция кольцевого модулятора.

Детектор огибающей не может восстановить исходный сигнал из DSBSC-сигнала (с двумя боковыми и подавлением несущей), полученного при исполь-

зовании диодного детектора, потому что, как это видно из рисунков 8.8 и 8.10, огибающая имеет неподходящую форму. Поэтому прием осуществляется с помощью балансного модулятора другого типа. Принятый модулированный сигнал умножается на сигнал-версию исходного несущего сигнала, который должен быть точно сфазирован в соответствии с принятым DSBSC-сигналом. Такой метод демодуляции называется когерентным детектированием. Для получения несущего сигнала в приемнике имеются два способа.

1. Передача несущей по отдельному каналу.

2. Формирование передатчиком пилот-сигнала на частоте несущей уменьшенной амплитуды (примерно 10 % от амплитуды несущего колебания). Это сохраняет значительную часть энергии по сравнению с обычной и нет нужды в дополнительном канале, как в первом способе.

В цифровой связи если для обнаружения сигналов приемник использует информацию о фазе несущей, то процесс называется когерентным обнаружением. При идеальном когерентном обнаружении приемник содержит прототипы каждого возможного сигнала. Эти сигналы-прототипы дублируют алфавит переданных сигналов по всем параметрам, даже по радиочастотной фазе. В этом случае говорят, что приемник автоматически подстраивается под фазу входного сигнала. В процессе демодуляции приемник перемножает и интегрирует входной сигнал с каждым прототипом (определяет корреляцию). Некогерентная демодуляция не требует знания абсолютной величины фазы несущей входящего сигнала, что в некоторой степени упрощает схему обнаружения, но одновременно повышает вероятность ошибки.

Следует отметить, что каждая боковая амплитудно-модулированного сигнала несет полную информацию об исходном модулированном сигнале. Значит, можно достичь удовлетворительного приема, работая с одной боковой (single sideband – SSB). При этом уменьшается полоса частот и потенциально удваивается информационная емкость канала в заданной полосе, а также уменьшается расход передаваемой энергии. Один из методов, используемых для организации связи на одной боковой частоте, заключается в том, что с помощью балансного модулятора сначала получают сигнал с двумя боковыми и подавлением несущей (DSBSC), а затем отфильтровывают одну из них. Связь только на одной боковой частоте тоже требует использования в приемнике когерентного детектора.

8.5. Угловая модуляция

Частотная и фазовая модуляции являются частными случаями угловой модуляции. Модулированный сигнал выражается следующим образом:

$$s(t) = A \cos[2\pi f_c t + \phi(t)]. \quad (8.6)$$

При фазовой модуляции фаза пропорциональна модулирующему сигналу:

$$\phi(t) = n_p m(t), \quad (8.7)$$

где n_p – коэффициент фазовой модуляции; $m(t)$ – модулирующий сигнал.

Поскольку частоту можно определить как скорость изменения фазы сигнала, то при частотной модуляции модулирующему сигналу пропорциональна производная фазы:

$$\phi'(t) = n_f m(t), \quad (8.8)$$

где n_f – коэффициент частотной модуляции.

На рис. 8.11 показаны амплитудная, частотная и фазовая модуляции синусоидальной волны. Формы сигналов с частотной и фазовой модуляцией весьма схожи. В действительности их невозможно различить без знания модулирующей функции.

Пример

Рассмотрим сигнал с угловой модуляцией:

$$s(t) = 10 \cos[(10^6)\pi + 5 \sin 2\pi(10^3)t].$$

Определим максимальное отклонение фазы и максимальное отклонение частоты.

Максимальное отклонение фазы равно $n_p = 5$.

Для вычисления коэффициента частотной модуляции возьмем производную от фазы: $n_f = 10^4 \pi$,

$$\begin{aligned} \phi'(t) &= (5 \sin 2\pi(10^3)t)' = 5 \cos 2\pi(10^3)t \cdot 2\pi(10^3) = \\ &= 10^4 \pi \cos 2\pi(10^3)t = n_f \cos 2\pi(10^3)t \end{aligned}$$

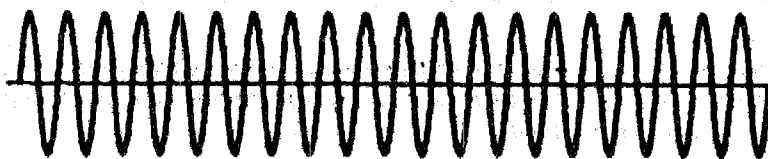
$$n_f = 10^4 \pi$$

Максимальное отклонение частоты равно $n_f = 10^4 \pi$ рад/с.

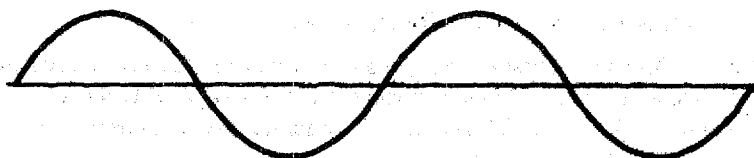
Так же как и в случае амплитудной модуляции, частотная и фазовая модуляции порождают сигнал с полосой, центрированной на частоте f_c . Но размеры этих полос различаются. Амплитудная модуляция представляет собой линейный процесс, поэтому она порождает частоты, равные сумме и разности несущего сигнала и гармоник модулирующего сигнала, поэтому полоса пропускания системы с АМ равна

$$W_T = 2W, \quad (8.9)$$

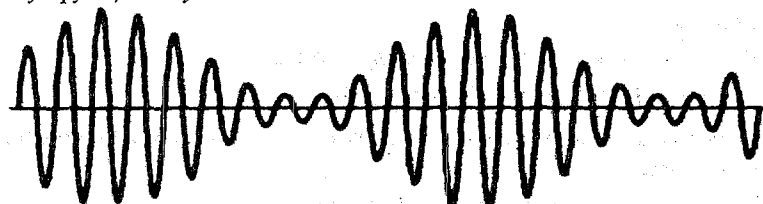
где W_T – полоса пропускания системы; W – полоса модулирующего сигнала.



а) Несущий сигнал



б) Модулирующий синусоидальный сигнал



в) Амплитудно-модулированный сигнал (двухполосный с подавлением несущей)



г) Фазомодулированный сигнал



д) Частотномодулированный сигнал

Рис. 8.11. Амплитудная, фазовая и частотная модуляции синусоидального сигнала синусоидальной несущей.

Угловая модуляция включает в себя член вида $\cos(\phi(t))$, который нелинеен и порождает большой диапазон частот. По сути, для модулируемой синусоиды с частотой f_m сигнал $s(t)$ будет содержать гармоники с частотами $f_c + f_m; f_c + 2f_m$ и т.д. В более общем случае для передачи частотно- или фазоманипулированного сигнала требуется бесконечная полоса. На практике хорошим эмпирическим правилом, известным как правило Карсона, является следующее:

$$W_T = 2(\beta + 1)W, \quad (8.10)$$

$$\text{где } \beta = \begin{cases} n_p A_m; \text{ фазовая модуляция} \\ \frac{\Delta F}{B} = \frac{n_f A_m}{2\pi B}; \text{ частотная модуляция} \end{cases}$$

ΔF – смещение модулируемой частоты относительно несущей;

A_m – максимальное значение модулирующего сигнала.

Формулу для частотной модуляции можно переписать в виде:

$$W_T = 2\Delta F + 2W. \quad (8.11)$$

Следовательно, и частотная, и фазовая модуляции требуют большей полосы, чем амплитудная.

Амплитудная модуляция требует большей передаваемой мощности при заданном отношении сигнал/шум на выходе детектора, чем другие виды модуляции. В некоторой мере это вызвано потерями мощности при передаче несущей. Зато частотная модуляция обладает улучшенными качественными характеристиками (Сигнал/Шум), которые можно и далее улучшать, изменяя соответствующим образом коэффициент модуляции. Однако “расплатой” за это служит увеличение полосы частот. Именно по этой причине системы с частотной модуляцией известны под названием систем с расширением полосы.

8.6. Квадратурная амплитудная модуляция

Квадратурная амплитудная модуляция (QAM- quadrature amplitude modulation) является популярным методом аналоговой передачи сигналов, используемым в асимметричных цифровых абонентских линиях (ADSL- asymmetric digital subscriber line). Данный метод модуляции со-

вмещает в себе амплитудную и фазовую модуляции. В методе QAM использованы преимущества одновременной передачи двух различных сигналов на одной несущей частоте, но при этом используется две копии несущей частоты, сдвинутые друг относительно друга на 90° . При квадратурной амплитудной модуляции обе несущие являются амплитудно-модулированными. В общих чертах модуляционная схема QAM показана на рис. 8.12. Со скоростью R бит/с на вход поступает поток двоичных чисел. Этот поток разбивается на два (биты попеременно распределяются по двум отдельным потокам), передаваемых со скоростью $\frac{R}{2}$ бит/с. Верхний поток на рисунке 8.12 модулируется на несущей частоте f_c так, что двоичный ноль представляется отсутствием несущей волны, а двоичная единица – наличием несущей волны постоянной амплитуды. Для модулирования нижнего потока на рис. 8.12 та же несущая волна смещается на 90° , после чего также производится модуляция. После этого два модулированных сигнала складываются и передаются вместе.

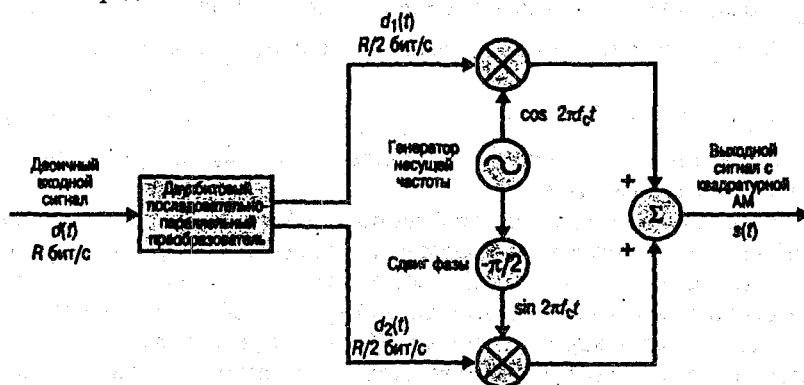


Рис. 8.12. Квадратурная амплитудная модуляция.

При использовании двухуровневой амплитудной модуляции каждый из двух потоков может находиться в одном из двух состояний, а объединенный поток – в одном из $2 \cdot 2 = 4$ состояний. При использовании четырехуровневой манипуляции (т.е. четырех различных уровней амплитуды) объединенный поток будет находиться в одном из $4 \cdot 4 = 16$ состояний. Разумеется, чем больше число состояний, тем выше потенциальная частота возникновения ошибок.

ЛЕКЦИЯ 9

9.1. Важнейший параметр цифровой связи – отношение сигнал/шум

В аналоговой связи критерием качества является отношение средней мощности сигнала к средней мощности шума (S/N или SNR). В цифровой связи в качестве критерия качества чаще используется нор-

мированная версия SNR , $\frac{E_b}{N_0}$. E_b — это энергия бита, и ее можно опи-

сать как мощность сигнала S , умноженную на время передачи бита T_b . N_0 — это спектральная плотность мощности шума, и ее можно выразить как мощность шума N , деленную на ширину полосы W . Поскольку время передачи бита и скорость передачи битов R_b , взаимно обратны,

T_b , можно заменить на $\frac{1}{R_b}$.

$$\frac{E_b}{N_0} = \frac{S T_b}{N / W} = \frac{S / R_b}{N / W}. \quad (9.1)$$

Ещё одним параметром, часто используемым в цифровой связи, является скорость передачи данных в битах в секунду. В целях упрощения выражений для представления скорости передачи битов вместо записи R_b будем писать просто R . С учетом сказанного перепишем выражение

(9.1) так, чтобы было явно видно, что отношение $\frac{E_b}{N_0}$ представляет со-

бой отношение S/N , нормированное на ширину полосы и скорость передачи битов.

$$\frac{E_b}{N_0} = \frac{S}{N} \frac{W}{R}. \quad (9.2)$$

Одной из важнейших метрик производительности в системах цифровой связи является график зависимости вероятности появления оши-

бочного бита P_b от $\frac{E_b}{N_0}$. На рис. 9.1 показан “водопадоподобный” вид

большинства подобных кривых. При $\frac{E_b}{N_0} \geq x_0$ $P_b \leq P_0$. Безразмерное

отношение $\frac{E_b}{N_0}$ — это стандартная качественная мера производительности систем цифровой связи. Следовательно, необходимое отношение

$\frac{E_b}{N_0}$ можно рассматривать как метрику, позволяющую сравнивать про-

изводительность различных систем; чем меньше требуемое отношение

$\frac{E_b}{N_0}$, тем эффективнее процесс обнаружения при данной вероятности

ошибки.

Рассмотрим полезность метрики $\frac{E_b}{N_0}$. Аналоговый сигнал отно-

сится к сигналам, представляемым через мощность. Почему это имеет смысл? Об аналоговом сигнале можно думать как о сигнале, имеющем бесконечную длительность, который не требуется разграничивать во времени. Неограниченно длительный аналоговый сигнал содержит бесконечную энергию; следовательно, использование энергии — это не самый удобный способ описания характеристик такого сигнала. Значительно более удобным параметром для аналоговых волн является мощность (или скорость доставки энергии).

В то же время в системах цифровой связи мы передаем (и принимаем) символы путем передачи некоторого сигнала в течение конечного промежутка времени, времени передачи символа T_s . Концентрируя внимание на одном символе, видим, что мощность (усредненная по времени) стремится к нулю. Значит, для описания характеристик цифрового сигнала мощность не подходит. Для подобного сигнала нужна метрика, “достаточно хорошая” в пределах конечного промежутка времени. Другими словами, энергия символа (мощность, проинтегрированная по T_s — это гораздо более удобный параметр описания цифровых сигналов.

Цифровой сигнал — это транспортное средство, представляющее цифровое сообщение. Сообщение может содержать один бит (двоичное сообщение), два (четверичное), ..., 10 бит (1024-ричное). В аналоговых системах нет ничего подобного такой дискретной структуре сообщения. Аналоговый информационный источник — это бесконечно квантованная непрерывная волна. Для цифровых систем критерий качества должен позволять сравнивать одну систему с другой на битовом уровне. Следовательно, описывать цифровые сигналы в терминах SNR практически бесполезно, поскольку сигнал может иметь однобитовое, 2-битовое или 10-битовое значение. Предположим, что для данной вероятности возникновения ошибки в цифровом двоичном сигнале требуемое отношение SNR равно 20. Будем считать, что понятия сигнала и его значения взаимозаменяемы. Поскольку двоичный сигнал имеет однобитовое значение, требуемое отношение SNR на бит равно 20 единицам. Предположим, что наш сигнал является 1024-ричным, с теми же 20 единицами требуемого отношения SNR. Теперь, поскольку сигнал имеет 10-битовое значение, требуемое отношение SNR на один бит равно всего 2.

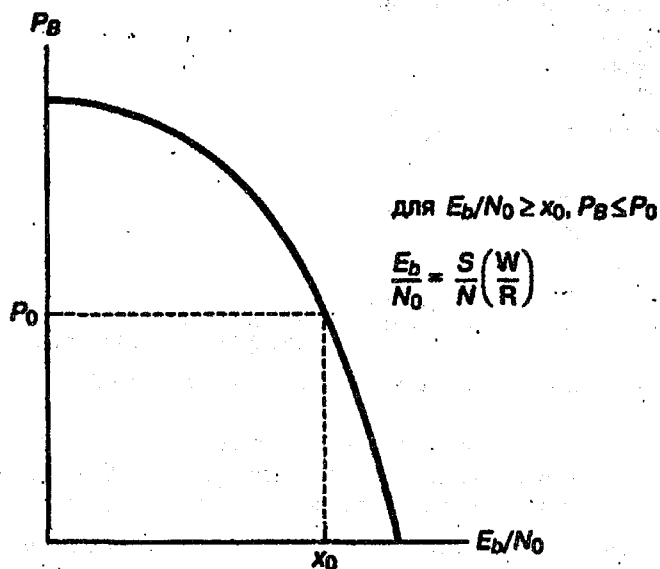


Рис. 9.1. Общий вид зависимости вероятности битовой ошибки от битового отношения сигнал/шум.

9.2. Векторное представление синусоиды

Использував теорему Эйлера, введем комплексную запись синусоидальной несущей:

$$e^{i\omega_0 t} = \cos \omega_0 t + i \sin \omega_0 t. \quad (9.1)$$

При комплексной записи указаны два важнейших компонента любой синусоидальной несущей волны, называемые взаимно ортогональными: синфазной (действительной) и квадратурной (мнимой). Немодулированная несущая представляется в полярной системе координат в виде единичного вектора с постоянной скоростью ω_0 рад/с, вращающегося против часовой стрелки. При переходе от времени t_0 к t_1 можно изобразить переменные во времени проекции вращающегося вектора на синфазной (I) и квадратурной (Q) осях. Эти декартовы оси обычно называют синфазным (I channel) и квадратурным (Q channel) каналами, а их проекции представляют взаимно ортогональные составляющие сигнала, связанные с этими каналами (рис. 9.2).

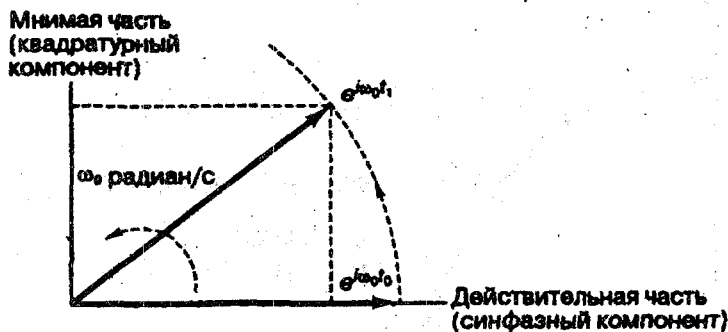


Рис. 9.2. Векторное представление синусоидальной несущей.

Процесс модуляции несущей можно рассматривать как систематическое возмущение вращающегося вектора и его проекций. При амплитудной модуляции синусоидой с единичной амплитудой и частотой $\omega_m \ll \omega_0$ вращающийся вектор $e^{i\omega_0 t}$ возмущается двумя боковыми векторами: $e^{i\omega_m t}$, вращающимся против часовой стрелки, и $e^{-i\omega_m t}$, вращающимся по часовой стрелке (рис. 9.3).

Квадратурный
компонент

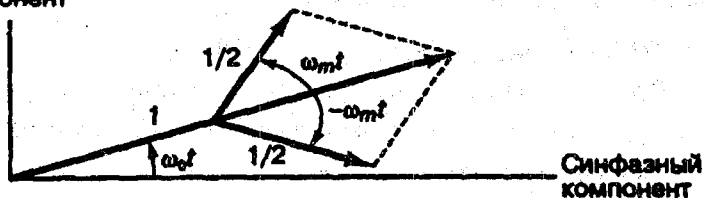


Рис. 9.3. Векторное представление амплитудной модуляции.

Боковые векторы вращаются намного медленнее, чем вектор несущей волны. В результате модулированный вектор несущей растет и уменьшается согласно указаниям боковых полос, но частота его вращения остается постоянной.

В случае частотной модуляции вектор несущей ускоряет или замедляет свое вращение согласно указаниям боковых полос, при этом амплитуда вектора несущей остается неизменной (рис. 9.4).

Квадратурный
компонент

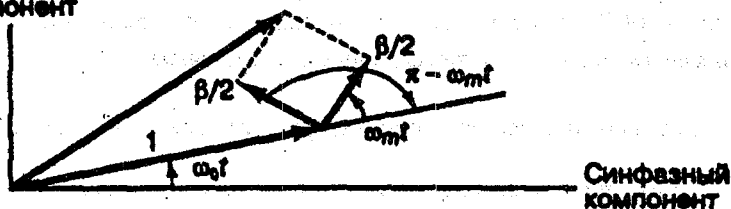


Рис. 9.4. Узкополосная частотная модуляция.

Амплитуда цифровых символов часто представляется в виде:

$$A = \sqrt{\frac{2E_i}{T}}, \quad (9.1)$$

где E_i – энергия i -того символа из алфавита M : $i = 1, 2, \dots, M$; T – длительность символа.

Выведем это. Сигнал описывается формулой $s(t) = A \cos \alpha t$, где A – максимальная амплитуда сигнала. Поскольку максимальное значе-

ние в $\sqrt{2}$ раз превышает его среднеквадратическое (root-mean-square – rms) значение, то можно записать:

$$s(t) = \sqrt{2}A_{rms} \cos \omega t = \sqrt{2A_{rms}^2} \cos \omega t. \quad (9.2)$$

Пусть сигнал выражается через колебания тока или напряжения, так что A_{rms}^2 представляет среднюю мощность P , нормированную на 1 Ом, поэтому можно записать следующее:

$$s(t) = \sqrt{2P} \cos \omega t. \quad (9.3)$$

Так как $P(\text{Ватт}) = \frac{E(\text{джоули})}{T(\text{секунды})}$, то выражение (9.3) можно переписать в виде:

$$s(t) = \sqrt{\frac{2E}{T}} \cos \omega t. \quad (9.4)$$

Поскольку ключевой параметр при определении вероятности ошибки в процессе обнаружения – это энергия принятого сигнала, то форма записи (9.4) является удобной для выражения вероятности ошибки в цифровом символе P_E как функцию энергии сигнала.

9.3. Расстояние между тонами цифровой частотной манипуляции

Для того чтобы цифровые сигналы некоторого алфавита хорошо различались между собой, они должны быть ортогональными. Что подразумевается под ортогональностью цифровых сигналов в частотной манипуляции (frequency shift keying – FSK)? Частотные тоны f_1 и f_2 являются ортогональными, если при переданном тоне f_1 дискретная отгибающая на выходе принимающего фильтра, согласованного с f_2 , дает нуль (т.е. отсутствуют перекрестные помехи). Подобная ортогональность между тонами сигнального множества FSK обеспечивается, если любая пара тонов множества разделена по частоте расстоянием, кратным $\frac{1}{T}$ Гц. Спектры подобных соседствующих тонов показаны на рис. 9.5.

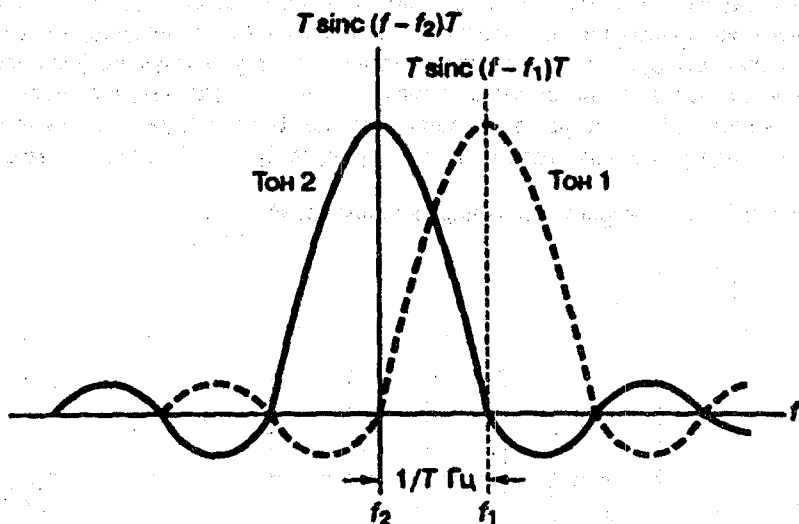


Рис. 9.5. Минимальное расстояние между тонами для ортогональной передачи сигналов FSK с некогерентным обнаружением.

Максимум спектра тона 1 должен совпадать с одним из переходов через нуль спектра тона 2, а максимум спектра тона 2 должен приходиться на один из переходов через нуль спектра тона 1. Расстояние по частоте между центром главного лепестка спектра и первым переходом через нуль является минимально необходимым расстоянием между ними. При некогерентном обнаружении это соответствует минимальному расстоянию между тонами, которое равно $\frac{1}{T}$ Гц.

Несмотря на то, что использование схемы FSK подразумевает передачу в течение каждого интервала передачи символа всего одного однополосного тона, когда говорят о ширине полосы сигнала, подразумевают спектр, достаточный для всех тонов M -арного множества. Следовательно, для модуляции FSK требования к полосе при некогерентном обнаружении равны $\frac{M}{T}$.

В случае когерентного обнаружения использование информации о фазе несущей означает, что частотные тоны могут быть расположены ближе друг к другу по частоте при соблюдении ортогональности в наборе тонов FSK. Рассмотрим пример, изображенный на рис. 9.6. В дан-

$$\frac{M}{2T},$$

$$-1-1+1-1+1+1+1+1-1+1-1-1=0$$

при когерентном обнаружении тонов FSK.

Пример

Являются ли они ортогональными?

на время передачи бита на частоте $f_1 = 10000$ Гц. В данном примере рассматривается некогерентное обнаружение сигналов.

9.4. Вероятность ошибки при использовании для связи наборов ортогональных и неортогональных сигналов

Если вместо двоичного алфавита используется алфавит из M ортогональных сигналов, то при данном SNR увеличивается число возможностей для появления символьной ошибки, всего существует $(M-1)$ возможность допустить ошибку. График, демонстрирующий ухудшение вероятности символьной ошибки P_E с увеличением M приведен на рис. 9.7. Фиксированный SNR означает фиксированный объем энергии на символ, следовательно, при увеличении M этот объем энергии нужно распределять уже между большим числом битов, т.е. на каждый бит приходится меньше энергии.

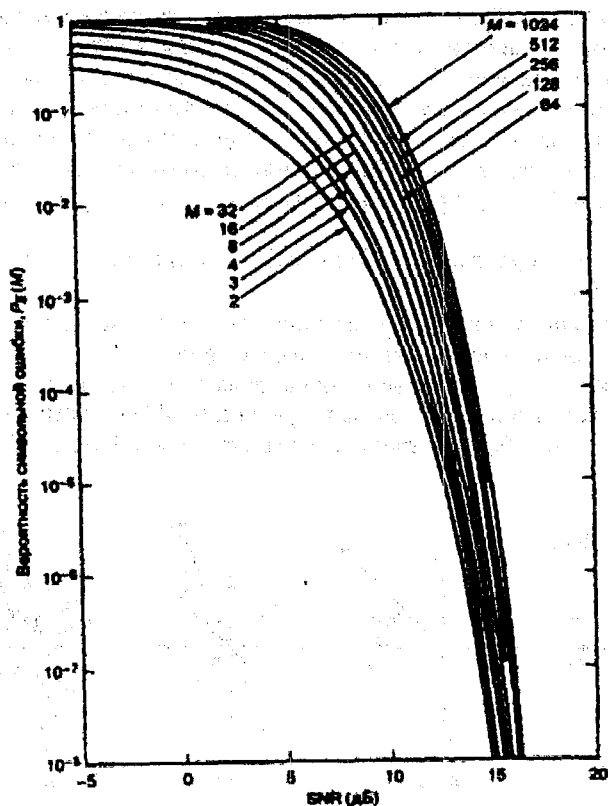


Рис. 9.7. Зависимость вероятности символьной ошибки от SNR для когерентной передачи сигналов FSK.

Однако ортогональная передача приводит к повышению достоверности обнаружения при увеличении объема алфавита M . Рассмотрим аналогию – приобретение товара. Выбор качества соответствует выбору точки на оси P_E рис. 9.7, скажем, 10^{-3} . Проведем из этой точки горизонтальную линию через все кривые (от $M=2$ до $M=1024$). Пусть мы покупаем самую маленькую упаковку творога, которая содержит две пачки и стоит \$1. Обращаясь к рис. 9.7, можно сказать, что такая покупка соответствует пересечению проведенной горизонтальной линии с графиком для $M=2$. Смотрим вниз на соответствующее значение параметра SNR и называем пересечение с этой осью ценой \$1. При следующем походе за покупками мы решаем, что в прошлый раз стоимость творога была высокой – по 50 центов за пачку. Поэтому решаем купить большую упаковку (8 пачек) за \$2. Данная покупка соответствует пересечению горизонтальной линии с кривой $M=8$ на рис. 9.7. Смотрим вниз и называем соответствующее значение SNR ценой \$2. Хотя мы и купили большую емкость, заплатив за нее большую цену, все же стоимость одной пачки упала (и составляет теперь 25 центов). Эту аналогию можно продолжать. Мы можем приобретать все большие и большие упаковки, при этом их цена (SNR) будет расти, а стоимость за пачку (требуемое для обнаружения $\frac{E_b}{N_0}$) будет падать. Это – эффект масштаба: приобретение

за один раз большого количества товара соответствует закупкам по оптовым ценам, при этом цена единицы товара падает.

Рассмотрим теперь векторное представление неортогональных сигналов – сигналов многофазной манипуляции (multiple phase shift keying – MPSK). На рис. 9.8 показаны наборы сигналов MPSK для $M=2, 4, 8$ и 16 .

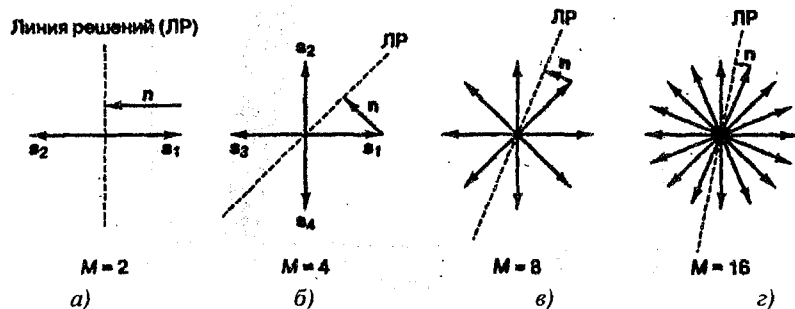


Рис. 9.8. Наборы сигналов MPSK для $M=2, 4, 8$ и 16 .

На рис. 9.8, а показаны бинарные ($k=1$; $M=2$) антиподные векторы s_1 и s_2 , угол между которыми равен 180° . Граница областей решений разделяет сигнальное пространство на две области. На рисунке также показан вектор шума n , равный по амплитуде сигналу s_1 . При указанных направлениях и амплитуде энергия вектора шума является минимальной для того, чтобы допустить ошибку в символе при приеме. На рис. 9.8, б показаны 4-арные ($k=2$; $M=4$) векторы, расположенные друг к другу под углом 90° . Границы областей решений (на рисунке изображена только одна) делят сигнальное пространство на четыре области. Здесь также изображен вектор шума n (начало – в вершине вектора сигнала, направление перпендикулярно ближайшей границе областей решений), являющийся вектором минимальной энергии, достаточной для того, чтобы допустить символическую ошибку при приеме. Вектор шума минимальной энергии на рис. 9.8, б меньше вектора шума на рис. 9.8, а, что свидетельствует о большей уязвимости 4-арной системы по сравнению с бинарной (энергии сигналов в обоих случаях взяты равными). Изучая рис. 9.8 в, г можно заметить следующую закономерность. При многофазной передаче сигналов по мере роста величины M на сигнальную плоскость помещается все больше сигнальных векторов. По мере того как векторы располагаются плотнее, для появления ошибки вследствие шума требуется все меньше энергии.

Размещение большего числа векторов сигналов в сигнальном пространстве эквивалентно повышению скорости передачи данных без увеличения системной ширины полосы (все векторы ограничиваются одной и той же плоскостью), т.е. повышению использования полосы за счет вероятности ошибки. Чем можно поступить, чтобы «выкупить» возросшую вероятность ошибки? Иными словами, чем мы можем поступить, чтобы расстояние между соседними векторами сигналов на рис. 9.8, г стало таким же, как на рис. 9.8 а? Можно увеличивать интенсивность сигнала (сделать векторы сигналов длиннее), пока минимальное расстояние от вершины вектора сигнала до линии решений не станет равным размеру вектора шума на рис. 9.8, а. Таким образом, для неортогонального набора сигналов по мере роста M можно увеличивать производительность полосы либо за счет повышения вероятности ошибки, либо за счет увеличения от-

ношения $\frac{E_b}{N_0}$.

ЛЕКЦИЯ 10

10.1. Демодуляция и обнаружение

Типичные функции демодуляции и обнаружения цифрового приемника показаны на рис. 10.1.

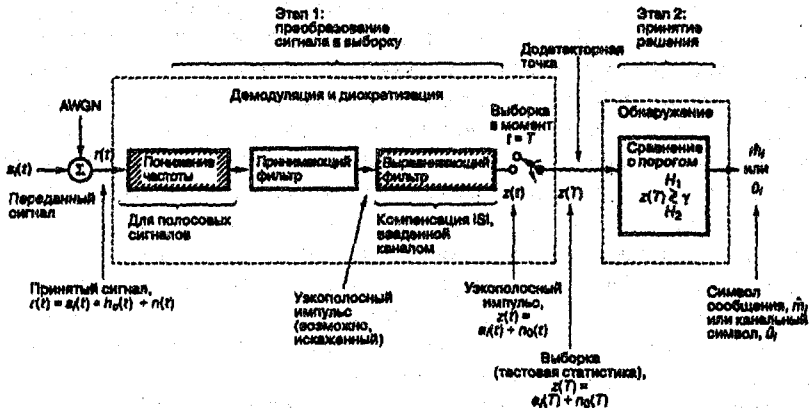


Рис. 10.1. Этапы процесса демодуляции/обнаружения цифровых сигналов.

Принятый сигнал $r(t)$ искажается вследствие воздействия шума $n(t)$, который предполагается аддитивным гауссовским (additive white gaussian noise - AWGN) и неидеальной импульсной характеристикой канала $h_c(t)$. На рис. 10.1 процесс демодуляции показан как восстановление сигнала в неискаженный узкополосный импульс, а обнаружение – как процесс принятия решения относительно цифрового значения этого сигнала, хотя в литературе часто под демодуляцией уже подразумевается обнаружение сигнала. В блоке демодуляции и дискретизации показан блок преобразования частоты, который не является обязательным, так как отвечает за трансляцию полосовых сигналов, работающих на определенных радиочастотах. Принимающий фильтр является обязательным устройством для демодуляции, так как его задачей является восстановление узкополосного импульса с максимально возможным уровнем сигнал/шум (SNR) и без межсимвольной интерференции. Оптимальный принимающий фильтр, выполняющий такую задачу, называется согласованным (matched) или коррелятором (correlator). За принимающим фильтром может находиться выравнивающий фильтр (эквалайзер – equalizer). Он не является обязательным блоком в демодуляторе, так как

необходим только в тех системах, в которых сигнал может искажаться вследствие межсимвольной интерференции, введенной каналом. Принимающий и выравнивающий фильтр показаны как два отдельных блока для подчеркивания различия их функций. Конструктивно же они могут быть выполнены как один фильтр.

В конце каждого интервала передачи символа T на выход устройства дискретизации, додетекторную точку, поступает выборка

$$z(T) = a_i(T) + n_0(T) \quad i=1,2, \quad (10.1)$$

где $a_i(T)$ – желаемый компонент сигнала; $n_0(T)$ – шум.

Плотности условных вероятностей $p(z|s_1)$ и $p(z|s_2)$ определим как:

$$p(z|s_1) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z - a_1}{\sigma_0} \right)^2 \right], \quad (10.2)$$

$$p(z|s_2) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z - a_2}{\sigma_0} \right)^2 \right], \quad (10.3)$$

Эти плотности условных вероятностей показаны на рис. 10.2. Плотность $p(z|s_1)$, изображенная справа, называется правдоподобием s_1 и показывает плотность вероятности случайной переменной $z(T)$ при условии передачи символа s_1 . Подобным образом функция $p(z|s_2)$ (справа) называется правдоподобием s_2 и показывает плотность вероятности случайной переменной $z(T)$ при условии передачи символа s_2 . Ось абсцисс представляет полный диапазон возможных значений выборки, взятой в течении этапа 1.

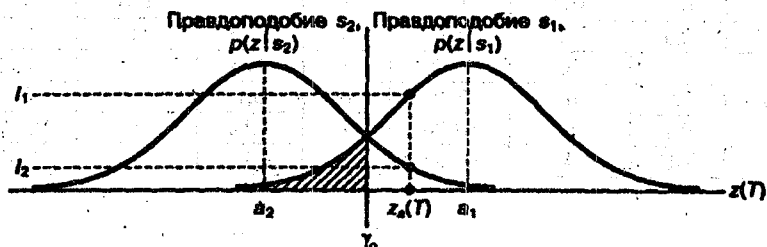


Рис. 10.2. Плотности условных вероятностей $p(z|s_1)$ и $p(z|s_2)$.

Пример

Построим плотности условных вероятностей сигналов 0 и 1 на выходе принимающего фильтра. Количество переданных символов пусть равняется 1000. В табл. 10.1 приведены результаты измерений.

На рис. 10.3 показаны построенные по данным табл. 10.1 кривые условных плотностей вероятностей передачи сигнала «0» ($p(z|s_1)$) и передачи сигнала «1» ($p(z|s_2)$) при условии приема искаженного сигнала z .

Из приведенного примера видно, что оптимальный принимающий фильтр – согласованный фильтр – отображает все сигналы с равными энергиями в одну и ту же точку $z(T)$, следовательно, для процесса обнаружения важна энергия сигнала; $z(T)$ является сигналом напряжения, пропорциональным энергии принятого символа.

Таблица 10.1

Результаты измерения энергии на выходе согласованного фильтра

При поступлении на вход сигнала «1»				При поступлении на вход сигнала «0»			
№ значения	Значение энергии на выходе фильтра	Зарегистрированное количество раз из 1000 (N)	Вероятность $p = \frac{N}{1000}$	№ значения	Значение энергии на выходе фильтра	Зарегистрированное количество раз из 1000 (N)	Вероятность $p = \frac{N}{1000}$
1	1	600	0,6	1	0	600	0,6
2	0,8	200	0,2	2	0,01	200	0,2
3	0,7	100	0,1	3	0,02	100	0,1
4	0,5	50	0,05	4	0,03	50	0,05
5	0,3	30	0,03	5	0,05	30	0,03
6	0,2	10	0,01	6	0,1	10	0,01
7	0,1	6	0,006	7	0,12	6	0,006
8	0,05	4	0,004	8	0,15	4	0,004

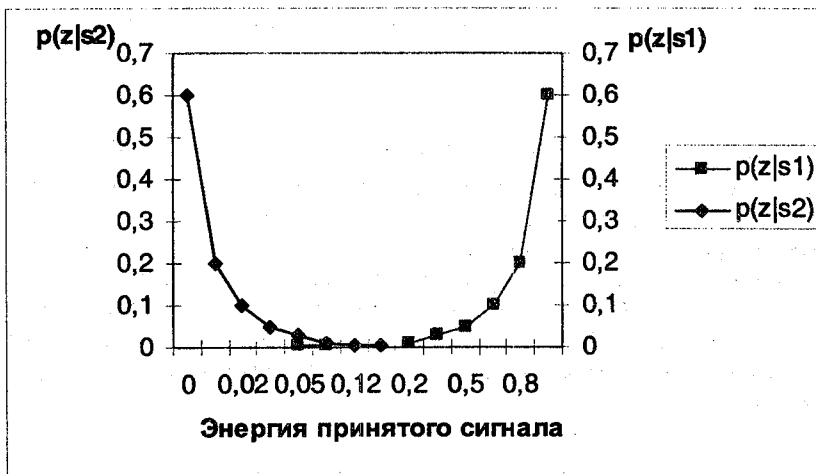


Рис. 10.3. Условные плотности вероятности, рассчитанные для примера.

На этапе 2 обнаружение выполняется посредством выбора гипотезы, являющейся следствием порогового измерения,

$$\begin{array}{c}
 H_1 \\
 z(T) > \gamma \\
 H_2 \\
 z(T) < \gamma
 \end{array} \quad (10.4)$$

где H_1 и H_2 — две возможные бинарные гипотезы о том, был ли передан сигнал $s_1(t)$ или $s_2(t)$; γ — порог.

При симметричных правдоподобиях (как это показано на рисунке 10.2) оптимальный порог выбирается из критерия минимальной ошибки как

$$\gamma_0 = \frac{a_1 + a_2}{2}, \quad (10.5)$$

где a_1 — сигнальный компонент при передаче $s_1(t)$; a_2 — сигнальный компонент при передаче $s_2(t)$.

В приведенном выше примере можно выбрать порог $\gamma_0 = \frac{1+0}{2} = 0,5$.

Для равновероятных сигналов оптимальный порог γ_0 проходит через пересечение функций правдоподобия, а этап принятия решения заключается в эффективном выборе гипотезы, соответствующей сигналу с максимальным правдоподобием.

В данном случае рассмотрен обнаружитель в виде жесткой схемы принятия решений, когда выход демодулятора квантуется на два уровня – нулевой и единичный. Если выход демодулятора имеет более двух уровней квантования, то это мягкая схема принятия решения (рис. 10.4).

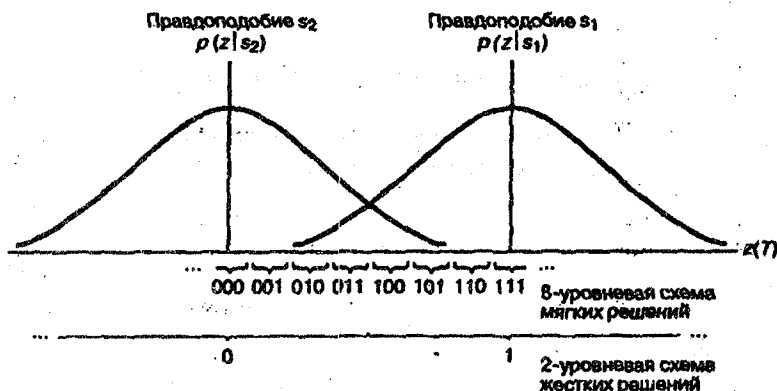


Рис. 10.4. Жесткая и мягкая схема принятия решений.

Поступление трехбитового символа вместо одного двоичного символа эквивалентно передаче обнаружителю меры достоверности вместе с решением относительно значения сигнала. Это приводит к улучшению на 2 дБ требуемого для обнаружения отношения сигнал/шум. В течение процесса обнаружения при использовании мягкой схемы принятия решений в данном случае нужно успеть обработать в три раза больше данных, поэтому за использование мягкой схемы принятия решений приходится платить увеличением объемов памяти и скоростью обработки.

10.2. Вероятность ошибки

Детектор, минимизирующий вероятность ошибки для классов равновероятных сигналов, называется детектором максимального правдоподобия. В процессе принятия бинарного решения, показанного на рис. 10.2, существует две возможности возникновения ошибки. При передаче сигнала $s_1(t)$ ошибка происходит при выборе гипотезы H_2 (заштрихованная область слева от γ_0 на рис. 10.2). При передаче сигнала $s_2(t)$ ошибка происходит при выборе гипотезы H_1 . Для равных априорных вероятностей (т.е. $P(s_1) = P(s_2) = \frac{1}{2}$) имеем следующую вероятность битовой ошибки:

$$P_B = \frac{1}{2}P(H_2 | s_1) + \frac{1}{2}P(H_1 | s_2). \quad (10.6)$$

Используя симметричность плотностей вероятности, можно записать:

$$P_B = P(H_2 | s_1) = P(H_1 | s_2). \quad (10.7)$$

Вероятность появления ошибочного бита P_B численно равна площади «под хвостом» любой функции правдоподобия $-p(z | s_1)$ или $p(z | s_2)$, «заползающим» на «неправильную» сторону порога. Таким образом, для вычисления P_B можно проинтегрировать $p(z | s_1)$ от $-\infty$ до γ_0 или $p(z | s_2)$ от γ_0 до ∞ . Если выбрать второй вариант и заменить правдоподобие $p(z | s_2)$ его гауссовским эквивалентом, то вероятность битовой ошибки можно записать как:

$$P_B = \int_{\gamma_0=(a_1+a_2)/2}^{\infty} \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z-a_2}{\sigma_0} \right)^2 \right] dz, \quad (10.8)$$

где σ_0^2 — дисперсия шума вне коррелятора.

Сделав замену $u = (z - a_2) / \sigma_0$; $\sigma_0 du = dz$, получим:

$$P_B = \int_{u=(a_1-a_2)/2\sigma_0}^{u=\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] du = Q\left(\frac{a_1-a_2}{2\sigma_0}\right), \quad (10.9)$$

$$\text{где } Q(x) \approx \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du \quad (10.10)$$

называется гауссовским интегралом ошибок и часто используется при описании вероятности с гауссовской плотностью распределения. Функция $Q(x)$ приведена в таблице приложения 1.

Аппроксимация функции $Q(x)$:

$$Q(x) = \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (10.11)$$

справедлива для значений $x > 3$. Программа для пакета MATLAB вычисления значений аппроксимированной функции $Q(x)$ приведена в приложении 2.

При униполярной передаче сигналов (ортогональный случай), когда $s_1(t) = A$; $s_2(t) = 0$ вероятность появления ошибочного бита равна:

$$P_B = Q\left(\sqrt{\frac{E_b}{N_0}}\right). \quad (10.12)$$

Для биполярной передачи сигналов (антиподный случай), когда $s_1(t) = A$; $s_2(t) = -A$ вероятность битовой ошибки равна:

$$P_B = Q\left(\sqrt{\frac{2E_b}{N_0}}\right). \quad (10.13)$$

Задача 10.1

Равновероятные двоичные импульсы в кодировке RZ когерентно обнаруживаются в гауссовом канале с $N_0 = 10^{-8}$ Вт/Гц. Предполагается, что синхронизация идеальна, амплитуда принятых импульсов равна 100 мВ и вероятность ошибки $P_B = 10^{-3}$; найдите наибольшую скорость передачи данных, возможную в описанной системе.

Решение

Сигналы в кодировке RZ можно рассматривать как униполярные (ортогональные), поэтому можно записать выражение для вероятности битовой ошибки как:

$$P_{\text{в}} = Q\left(\sqrt{\frac{E_b}{N_o}}\right) = 10^{-3}.$$

Из таблицы приложения 1 найдем значение $Q=3,08$. Получим, что

$$\sqrt{\frac{E_b}{N_o}} = 3,08; \text{ откуда } \frac{E_b}{N_o} \approx 9,5 \text{ и энергия бита } E_b = 9,5 \cdot 10^{-8} \frac{\text{Вт}}{\text{Гц}};$$

В то же время энергия бита равна:

$$E_b = \frac{S}{R} = 9,5 \cdot 10^{-8}$$

Мощность сигнала при нормировке к сопротивлению 1 Ом найдем как: $S = U^2 = 10\,000 \cdot 10^{-3} B^2 = 10 \text{ Вт}$

Теперь найдем битовую скорость передачи данных как

$$R = \frac{10 \text{ Вт}}{9,5 \cdot 10^{-8} \frac{\text{Вт}}{\text{Гц}}} = 1,0526 \cdot 10^8 \text{ Гц} = 105,26 \text{ МГц} = \frac{\text{Мбит}}{c}$$

Задача 10.2

Двоичная система связи передает сигналы $s_i(t)$ ($i = 1, 2$). Тестовая статистика приемника $z(T) = a_i + n_0$, где компонент сигнала a_i равен $a_1 = +1$ или $a_2 = -1$, а компонент шума n_0 имеет равномерное распределение. Плотности условного распределения $p(z|s_i)$ даются выражениями:

$$p(z|s_1) = \begin{cases} \frac{1}{2}; & -0,2 \leq z \leq 1,8 \\ 0; & \text{для других } z \end{cases} \quad p(z|s_2) = \begin{cases} \frac{1}{2}; & -1,8 \leq z \leq 0,2 \\ 0; & \text{для других } z \end{cases}$$

Определите вероятность появления ошибки P_B для равновероятной передачи сигналов и использования оптимального порога принятия решения.

Решение

На рис. 10.5 показаны функции правдоподобия, заданные в условии. Оптимальный порог рассчитаем как $\gamma = \frac{a_1 + a_2}{2} = \frac{1-1}{2} = 0$.

Вероятность появления ошибочного бита P_B численно равна площади под «хвостом» любой функции правдоподобия, «заползающей» на «неправильную» сторону порога, т.е. площади, заштрихованной на рис. 10.4 области:

$$P_B = 0,2 \cdot \frac{1}{2} = 0,1.$$

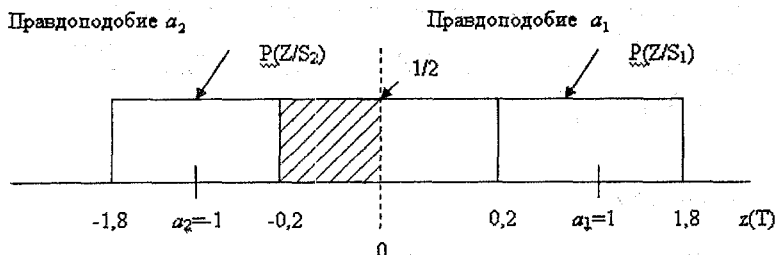


Рис. 10.5. Плотности условных вероятностей.

Задача 10.3

Биполярный двоичный сигнал $s_1(t)$ – это импульс $+1$ или -1 В на интервале $(0, T)$. К сигналу добавляется белый аддитивный гауссов шум с двусторонней спектральной плотностью мощности 10^{-3} Вт/Гц. Если обнаружение принятого сигнала производится с помощью согласованного фильтра, определите максимальную скорость передачи битов, которую можно поддерживать при вероятности появления ошибочного бита $P_B \leq 10^{-3}$.

Решение

Аналогично решению задачи 10.1.

Для антиподных сигналов $P_B = Q\left(\sqrt{\frac{2E_6}{N_0}}\right) = 10^{-3}$,

$$\sqrt{\frac{2E_6}{N_0}} = 3,08 \Rightarrow \frac{2E_6}{N_0} \approx 9,5 \Rightarrow E_6 = \frac{9,5 \cdot 10^{-3}}{2} = 0,005 \text{ Вт/Гц}$$

$S = U^2 = 1$ Вт мощность сигнала

$$E_6 = \frac{S}{R_6} \Rightarrow R_6 = \frac{1}{0,005} = 200 \text{ бит/с.}$$

Максимальная скорость передачи битов при вероятности появления ошибочного бита $P_B \leq 10^{-3}$ должна быть 200 бит/с.

Задача 10.4

Двоичные импульсы в кодировке NRZ передаются по кабелю, ослабляющему сигнал на 3 дБ (на пути от передатчика к приемнику). Эти импульсы когерентно обнаруживаются приемником, а скорость передачи данных равна 56 Кбит/с . Шум считать гауссовым с $N_0 = 10^{-6} \text{ Вт/Гц}$. Чему равна минимальная мощность, необходимая для передачи с вероятностью ошибки $P_B = 10^{-3}$?

Решение

Для антиподных сигналов $P_B = Q\left(\sqrt{\frac{2E_B}{N_0}}\right)$. Из таблицы приложения 1 находим, что для $P_B = 10^{-3}$ функция $Q=3,08$. Отсюда

$$\frac{2E_B}{N_0} = 9,5. \text{ Так как } E_6 = \frac{S}{R} = 9,5$$

$$2(S/R_B) = N_0 \cdot 9,5 = 10^{-6} \cdot 9,5 \Rightarrow \frac{S}{R_B} = 4,75 \cdot 10^{-6}.$$

Средняя мощность сигнала $S = 4,75 \cdot 10^{-6} \cdot 56 \cdot 10^3 = 0,266$ Вт, что при переводе в децибелы равно: $S_{\text{дБ}} = 10 \lg 0,266 = -0,575$ дБ.

С учетом ослабления сигнала минимальная мощность, необходимая для передачи, равна $S_{\text{дБ}} = 3 \text{ дБ} - 0,575 \text{ дБ} = 2,425 \text{ дБ}$ или при переводе в Ватты:

$$\lg S = 2,425/10$$

$$S = 1,75 \text{ Вт.}$$

10.3. Согласованный фильтр

Согласованный фильтр (matched filter) – это линейное устройство, спроектированное для того, чтобы давать на выходе максимально возможное для данного передаваемого сигнала отношение сигнал/шум. Импульсная характеристика такого фильтра $h(t)$ является зеркальным отображением сигнала сообщения $s(t)$, запаздывающим на время передачи символа T . Запаздывание на T секунд делает $h(t)$ функцией положительного времени (рис. 10.6).

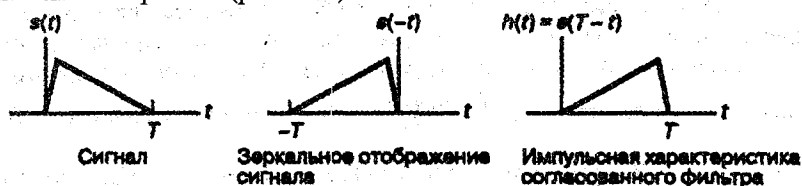


Рис. 10.6. Импульсная характеристика согласованного фильтра.

Выход $z(T)$ причинного фильтра во временной области можно описать как свертку принятого входного сигнала $r(t)$ с импульсной характеристикой фильтра.

Почему импульсная характеристика фильтра представляет собой зеркальное отображение сигнала? При представлении упорядоченных во времени событий существует проблема путаницы между самым старшим битом и самым младшим. На рис. 10.7, а показано, как обычно изображается функция времени: самое раннее событие представлено слева, а наиболее позднее – справа. На рис. 10.7, б показано, как импульсы поступают в сеть (или канал) и покидают ее. Здесь самое раннее событие изображено справа, а наиболее позднее – слева. Обращение во времени импульсной характеристики согласованного фильтра можно рассматривать как «предварительную коррекцию», чтобы поступающий

сигнал и импульсный отклик фильтра или «нога в ногу» (ранний с ранним, поздний с поздним).

Довольно часто термин «согласованный фильтр» используется как синоним термина «коррелятор». Дело в том, что свертка в согласованном фильтре с обращенной во времени функцией дает еще одно обращение во времени, подавая на выход 9 в конце интервала передачи символа то, что является корреляцией сигнала с собственной копией. Выходы коррелятора и согласованного фильтра одинаковы только в момент времени $t=T$.

Для синусоидального входа выход коррелятора на промежутке $0 \leq t \leq T$ приблизительно описывается линейной функцией, а выход согласованного фильтра приблизительно описывается синусоидой, амплитуда которой в том же промежутке времени модулирована линейной функцией (рис. 10.8).

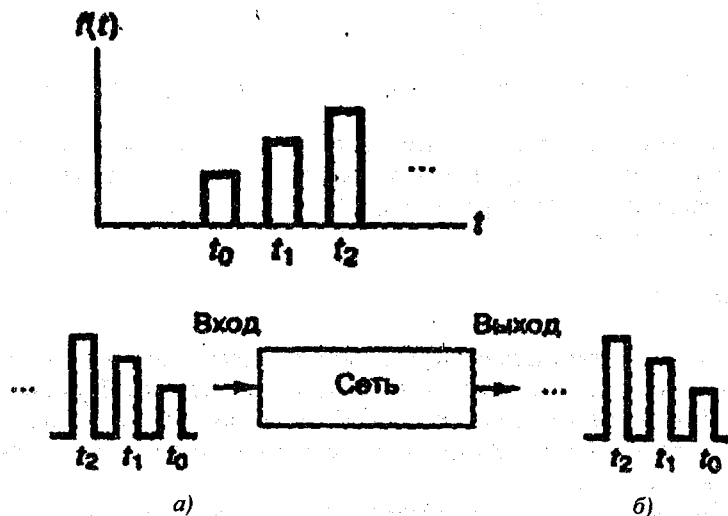


Рис. 10.7. Дилемма в представлении упорядоченных во времени событий.

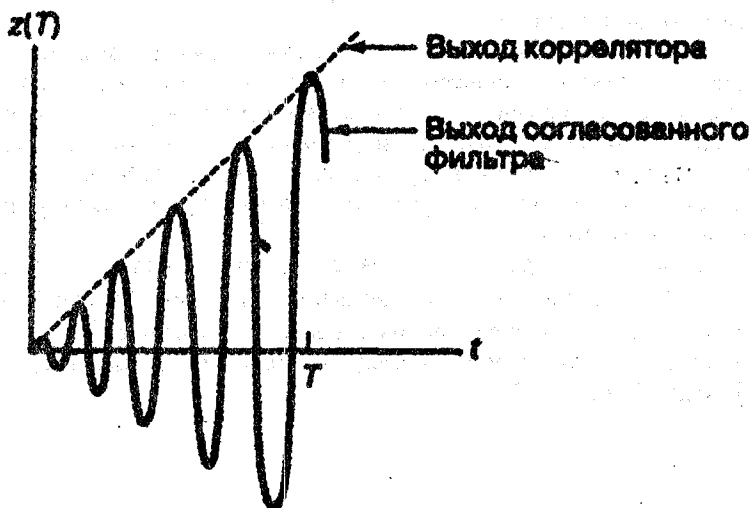


Рис. 10.8. Сравнение выходов коррелятора и согласованного фильтра.

Поскольку при соизмеримых входах выходы согласованного фильтра и коррелятора идентичны в момент взятия выборки $t=T$, функции согласованного фильтра и коррелятора часто используются как взаимозаменяемые.

Задача 10.5

А) Опишите импульсную характеристику согласованного фильтра, используемого для обнаружения дискретного сигнала, изображенного на рис. 10.9. Какой сигнал на выходе с фильтра получается при подаче данного сигнала на вход? Воздействием шума можно пренебречь. Чему равно значение на выходе?

Б) Предположим, что при реализации согласованного фильтра каналы случайно соединены так, что фильтр дает корреляцию (а не свертку) сигнала и его обращенной во времени копии. Покажите выход как функцию времени. Чему равно максимальное значение на выходе? При данных условиях максимальное значение на выходе появляется в другой момент времени, чем в пункте А.

В) С помощью значений на выходе неверного фильтра, описанного в пункте Б по сравнению с корректными значениями из пункта А, можно ли найти ключ, который поможет предсказать, появляется ли некото-

рая последовательность с выхода правильного или неправильного фильтра?

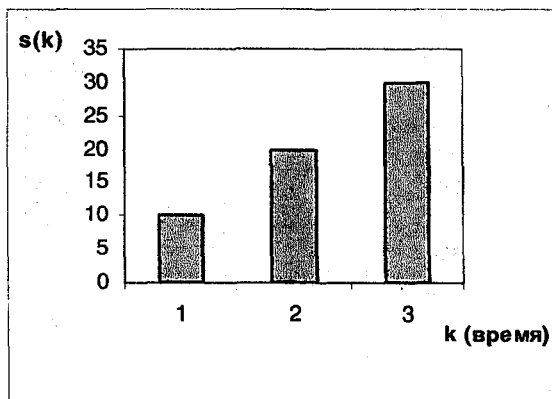


Рис. 10.9. Дискретный сигнал.

Решение

Импульсная характеристика согласованного фильтра (рис. 10.10) представляет собой зеркальное отображение сигнала.

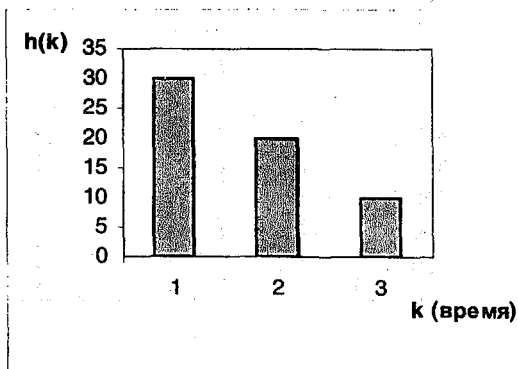


Рис. 10.10. Импульсная характеристика фильтра, согласованного с дискретным сигналом рисунка 10.9.

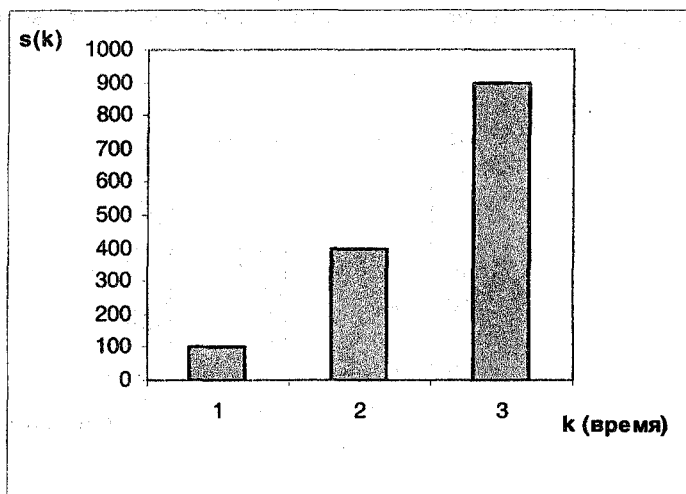


Рис. 10.11 Сигнал на выходе фильтра, согласованного с сигналом рисунка 10.9.

Максимальное значение на выходе «правильного» согласованного фильтра (рис. 10.11) равно 900 (появляется в момент времени $k=3$).

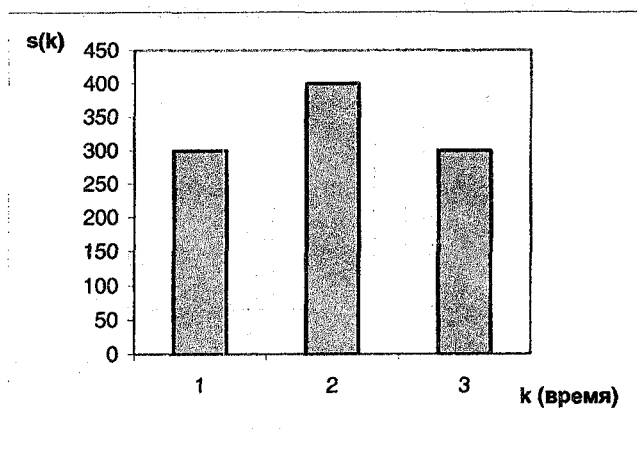


Рис. 10.12. Выходной сигнал «неправильного» фильтра (пункт Б).

Максимальное значение на выходе «неправильного» согласованного фильтра (рис. 10.12) равно 400 (появляется в момент времени $k=2$).

«Ключом» предсказания появления последовательности с выхода «правильного» фильтра служит сумма значений энергии импульсов. Для «правильного» фильтра (рис. 10.10) эта сумма равна $100+400+900=1400$, а для «неправильного» фильтра эта сумма равна $300+400+300=1000$.

ЛЕКЦИЯ 11

11.1. Формирование импульсов с целью снижения межсимвольной интерференции

Канал связи всегда имеет ограниченную полосу пропускания. Канал с ограниченной полосой всегда искажает или расширяет сигнал, проходящий через него. Если ширина полосы канала значительно больше ширины полосы сигнала, импульс искажается незначительно. Если ширина полосы канала приблизительно равна ширине полосы сигнала, то искажение будет превышать интервал передачи символа и приведет к наложению импульсов сигнала и возникновению межсимвольной интерференции (ISI). Исследованием проблемы задания формы принятого импульса с тем, чтобы предотвратить появление ISI на детекторе, долгое время занимался Найквист. Он показал, что минимальная теоретическая ширина полосы системы, требуемая для определения скорости передачи сим-

волов R_s сообщения без ISI, равна $\frac{R_s}{2}$ Гц. Это возможно, если пе-

редаточная функция системы $H(f)$ имеет прямоугольную форму, как показано на рис. 11.1, а. Импульсная характеристика функции $H(f)$, вычисляемая с помощью преобразования Фурье, имеет вид

$h(t) = \sin c \frac{t}{T}$. Она показана на рис. 11.1, б.

Импульс, описываемый функцией $h(t) = \sin c \frac{t}{T}$, имеет бесконечную длительность и состоит из множества лепестков: главного и боковых, называемых хвостами. На рис. 11.1,б показано, как можно обойти ISI. Итак, имеем два последовательных импульса $h(t)$ и $h(t-T)$. Несмотря на то что хвосты функции $h(t)$ имеют бесконечную длительность, в момент времени $t=T$ взятия выборки функции $h(t-T)$ хвост функции $h(t)$ проходит через точку нулевой амплитуды, и также он будет иметь нулевую амплитуду в моменты взятия выборок всех остальных импульсов последовательности $h(t-kT)$, $k=\pm 1, \pm 2, \dots$. Предполагая идеальную синхронизацию процесса взятия выборок, получаем, что межсимвольная интерференция не будет влиять на процесс обнаружения. При идеальной фильтрации Найквиста и нулевой ISI максимальная возможная скорость передачи символов на 1 Гц полосы, называемая уплотнением скорости передачи символов равна 2 (символа/с)/Гц. Фильтр Найквиста – это фильтр, передаточная функция которого может быть представлена прямоугольной функцией, свернутой с любой четно-симметричной частотной функцией. Импульс Найквиста – это импульс, форма которого может быть описана функцией $\sin c \frac{t}{T}$, умноженной на другую временную функцию. Следовательно, существует бесконечное множество фильтров Найквиста и соответствующих импульсов.

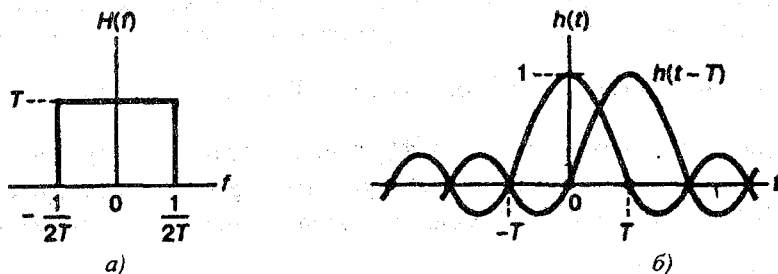


Рис. 11.1. Каналы Найквиста для нулевой ISI.
а) передаточная функция системы, б) принятый импульс.

В классе фильтров Найквиста наиболее популярными являются фильтры с характеристикой типа приподнятого косинуса или корня из приподнятого косинуса.

Основным параметром систем связи является эффективность использования полосы $\frac{R}{W} \frac{\text{бит/с}}{\text{Гц}}$. Она представляет собой меру скорости

переноса данных на единицу ширины полосы и показывает, насколько эффективно любой метод передачи сигналов использует ресурс полосы.

При нахождении $\frac{R}{W}$ для любой схемы передачи сигналов нужно знать,

сколько битов представляет каждый символ. Пусть сигналы кодируют с использованием М-уровневой кодировки РАМ. Каждый символ, включающий k бит представляется одной из М импульсных амплитуд. Для $k=6$ бит на символ размер набора символов составляет $M = 2^k = 64$ амплитуды, а теоретическая максимальная эффективность использования полосы, не допускающая ISI, равна $12 \frac{\text{бит/с}}{\text{Гц}}$. Таким образом, мак-

симальная эффективность использования полосы без ISI для сигнала с цифровым представлением k бит/символ равна $2k$.

Чем компактнее спектр передачи сигналов, тем выше разрешенная скорость передачи данных или больше число пользователей, которые могут обслуживаться одновременно. Это имеет большое значение для поставщиков услуг связи, так как более эффективное использование доступной ширины полосы приносит больший доход. Хотя компактный спектр и позволяет оптимальным образом использовать полосу, он оказывается очень чувствительным к ошибкам синхронизации, которые приводят к увеличению ISI.

11.2. Сложности связи по каналу с замираниями

Если характеристики радиоканала не заданы, то подразумевается, что сигнал затухает с расстоянием так же, как при распространении в идеальном свободном пространстве. В модели свободного пространства область между антеннами передатчика и приемника предполагается свободной от объектов, которые могли бы поглощать или отражать энергию на радиочастотах. Предполагается также, что внутри этой области атмосфера ведет себя как совершенно однородная непоглощаю-

щая среда. Кроме того, считается, что земля имеет пренебрежимо малый коэффициент отражения. В этой идеализированной модели свободного пространства ослабление радиочастотной энергии между передатчиком и приемником происходит по закону обратных квадратов. Мощность приемника, выраженная через переданную мощность, ослабляется в $L_s(d)$ раз. Данный параметр называется потерями в тракте. Если антенна приемника изотропна (одинаково принимает по всем направлениям), то этот коэффициент определяется как

$$L_s(d) = \left(\frac{4\pi d}{\lambda}\right)^2, \quad (11.1)$$

где d – расстояние между передатчиком и приемником, а λ – длина волны распространяемого сигнала.

Для большинства реальных каналов модель распространения в свободном пространстве неадекватно описывает поведение канала и не позволяет предсказывать характеристики системы. *В системах мобильной радиосвязи сигнал может передаваться от передатчика к приемнику по множеству отражающих путей. Это явление, называемое многолучевым распространением, может вызывать флуктуации амплитуды, фазы и угла прибытия полученного сигнала, что определило название замирание вследствие многолучевого распространения (multipath fading).*

Крупномасштабное замирание отражает среднее ослабление мощности сигнала или потери в тракте вследствие распространения на большое расстояние. Статистика крупномасштабного замирания позволяет приблизительно рассчитать потери в тракте как функцию расстояния. Это часто описывается через средние потери в тракте (степенной закон n -порядка) и логарифмически нормально распределенные отклонения от среднего.

Мелкомасштабное замирание – это значительные изменения амплитуды и фазы сигнала, которые на практике могут быть результатом небольших изменений (порядка половины длины волны) расстояния между передатчиком и приемником. *Мелкомасштабное замирание называется релеевским*, если имеется большое число многократно отражающихся путей и нет компонента сигнала вдоль луча обзора. Огибающая такого сигнала статистически описывается с помощью релевской функции плотности вероятности. Если преобладает незамирающий компонент сигнала такой, как путь распространения вдоль луча

обзора, огибающая мелкомасштабного замирания описывается функцией плотности вероятности Райса.

Мобильный радиоприемник, который перемещается по большому пространству, должен иметь возможность обрабатывать сигналы, подвергнувшиеся мелкомасштабному замиранию, наложенному на крупномасштабное. Крупномасштабное замирание (ослабление или потери в тракте) можно рассматривать как пространственное усреднение мелкомасштабных флуктуаций сигнала. Оно вычисляется, как правило, путем усреднения полученного сигнала по интервалу, превышающему 10 – 30 длин волн, чтобы отделить мелкомасштабные (главным образом, релейские) флуктуации от крупномасштабных эффектов затенения (обычно с логарифмически нормальным распределением).

Комплексный безразмерный множитель замирания $\alpha(t)$ можно выразить как

$$\alpha(t) = m(t) \cdot r_0(t), \quad (11.2)$$

где $m(t)$ – компонент крупномасштабного замирания;

$r_0(t)$ – компонент мелкомасштабного замирания.

На рис. 11.2 показано, что мощность полученного сигнала является функцией множителя $\alpha(t)$. Обычное изменение положения антенны, соответствующее переходу между соседними нулями изменения интенсивности сигнала вследствие мелкомасштабного замирания, равно приблизительно половине длины волны. Компонент $m(t)$ оценивается путем усреднения принятой огибающей по 10-30 длинам волн. Логарифмически нормально распределенное замирание является относительно медленно изменяющейся функцией местоположения. В приложениях, включающих движение, таких, например, как использование радио в движущейся машине, зависимость от местоположения равносильна зависимости от времени.

На рис. 11.3 показан основной механизм, приводящий к замиранию в каналах с многолучевым распространением. На рисунке отраженный сигнал запаздывает по фазе (из-за увеличения расстояния распространения) относительно ожидаемого сигнала. Отраженный сигнал также имеет меньшую амплитуду (функция коэффициента отражения препятствия). Отраженные сигналы можно также записать с помощью ортогональных компонентов $x_n(y)$, $y_n(t)$. Если количество стохастических

компонентов велико и ни один из них не преобладает, то в фиксированный момент времени переменные $x_r(y)$, $y_r(t)$, являющиеся результатом суммирования всех $x_n(y)$, $y_n(t)$ будут иметь гауссову функцию распределения вероятностей.

Мелкомасштабное замирание проявляется двумя способами:

- Путем расширения цифровых импульсов сигнала.
- Посредством переменного во времени поведения канала, вызванного движением (например, принимающая антенна находится на движущейся платформе).

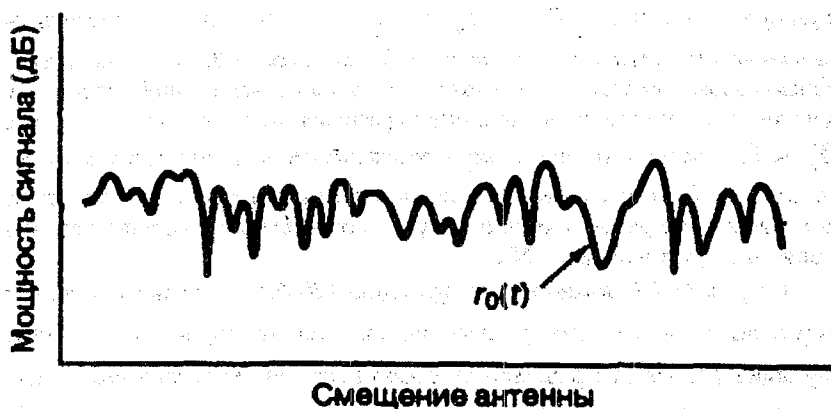
В этой связи важно различать задержку τ сигнала и время передачи сигнала t . Задержка – это следствие расширения сигнала во времени, являющегося результатом неоптимальной импульсной характеристики канала с замираниями. Время передачи связано с передвижением антенны или пространственными изменениями, учитывающими изменения пути распространения, которые определяют нестационарное поведение канала.

Механизм расширения по времени во временной области будет характеризоваться задержкой многолучевого распространения, а в частотной области – полосой когерентности канала. Нестационарный механизм во временной области будет характеризоваться временем когерентности канала, а в частотной области – скоростью замирания в канале или доплеровским расширением.

На рис. 11.4,а изображен профиль интенсивности многолучевого распространения (зависимость принятой мощности сигнала от задержки $S(\tau)$). Для типичного беспроводного сигнала полученный сигнал обычно состоит из нескольких дискретных многолучевых компонентов, приводящих к появлению изолированных пиков $S(\tau)$, называемых пальцами (fingers). Для единичного переданного импульса время T_m между приемом первого и последнего компонентов представляет собой максимальную избыточную задержку распространения, после которой мощность многолучевого сигнала падает ниже определенного порогового уровня относительно самого мощного компонента. Пороговый уровень можно выбрать на 10 или 20 дБ ниже уровня самого мощного луча.



а) Суперпозиция мелкомасштабных и крупномасштабных замираний



б) Мелкомасштабное замирание относительно средней мощности

Рис. 11.2. Крупномасштабное и мелкомасштабное замирание.

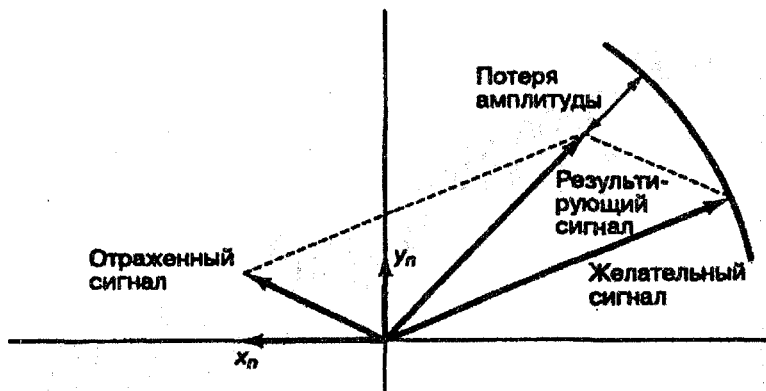


Рис. 11.3. Влияние многолучевого отражения сигнала на ожидаемый сигнал.

Взаимосвязь между максимальной избыточной задержкой распространения T_m и временем передачи символа T_s можно рассматривать с позиции двух различных категорий ухудшения качества передачи – частотно-селективного замирания и частотно-неселективного или амплитудного замирания. Если $T_m > T_s$, то говорят о частотно-селективном замирании и вводимой каналом ISI. При частотно-селективном замирании возможно уменьшение искажений, поскольку многие многолучевые компоненты разрешаются приемником. Если $T_m < T_s$, говорят о частотно-неселективном (или амплитудном) замирании. В этом случае ISI каналом не вводится, но многолучевые компоненты не разрешаются и могут деструктивно суммироваться, что приводит к уменьшению SNR.

На рис. 11.4,б можно видеть функцию $|R(\Delta f)|$, обозначенную как корреляционная функция разнесения частоты, это Фурье-образ $S(\tau)$. Функция $|R(\Delta f)|$ представляет собой корреляцию между реакциями канала на два сигнала как функцию разности частот этих сигналов. Ее можно рассматривать как частотную передаточную функцию канала. Следовательно, расширение сигнала во времени можно считать следствием процесса фильтрации. На рис. 11.5 изображен широкополосный фильтр (короткая импульсная характеристика) и его влияние на сигнал во временной и частотной областях. Этот фильтр похож на канал с амплитудным замиранием, выход которого относительно неискажен. На

рис. 11.6 показан узкополосный фильтр (широкая импульсная характеристика). Выходной сигнал претерпевает большее искажение как во временной, так и в частотной области. Данный процесс подобен происходящему в частотно-селективном канале.

Функцию $|R(\Delta f)|$ можно измерить, передав пару синусоид, разнесенных по частоте на Δf , и изучив взаимную корреляцию спектров двух полученных сигналов. Полоса когерентности f_0 представляет собой диапазон частот, в пределах которого частотные компоненты сигнала имеют большую вероятность амплитудной корреляции. Иными словами, на все спектральные компоненты этого диапазона канал влияет одинаково.

Разброс задержек и полоса когерентности канала связаны с характеристиками многолучевого распространения в канале и отличаются для разных путей распространения (городская черта, пригород, холмистая местность, помещения). Приблизленно полосу когерентности канала можно записать в виде:

$$f_0 \approx \frac{1}{5\sigma_\tau}, \quad (11.3)$$

где σ_τ – среднеквадратический разброс задержек.

Полоса когерентности устанавливает верхний предел скорости передачи, которую можно использовать, не включая в приемник эквалайзер.

Применяемые в мобильной радиосвязи каналы нестационарны, поскольку движение передатчика и приемника приводит к изменениям пути распространения. Для переданного непрерывного сигнала это вызовет изменение амплитуды и фазы в приемнике.

На рис. 11.4, в показана функция $R(\Delta t)$, обозначающая пространственно-временную корреляционную функцию. Эта функция определяет степень корреляции между откликом канала на синусоиду, отправленную в момент времени t_1 , и откликом на аналогичную синусоиду, отправленную в момент времени t_2 . $\Delta t = t_2 - t_1$. Для измерения $R(\Delta t)$ можно передать одну и ту же синусоиду ($\Delta f = 0$) в моменты времени t_1 и t_2 и определить функцию взаимной корреляции полученных сигналов.

Время когерентности канала T_0 – эта мера ожидаемого времени, на которое характеристика канала существенно инвариантна (это значение Δt соответствующее большому значению $R(\Delta t)$).

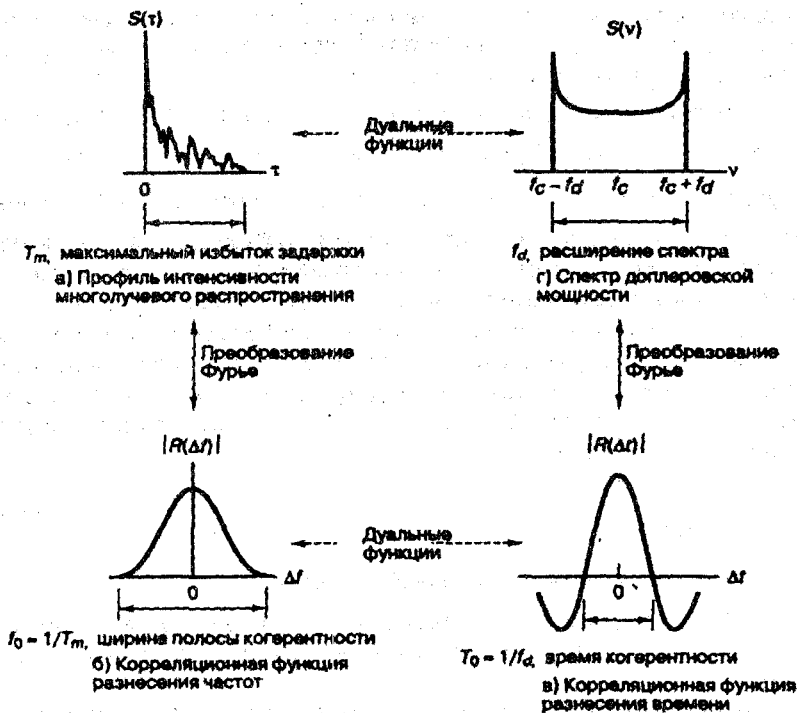


Рис. 11.4. Соотношения между корреляционными функциями канала и функциями плотности мощности.

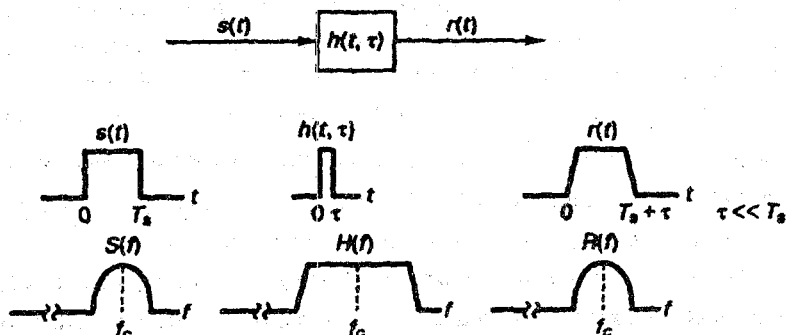


Рис. 11.5. Характеристики канала с амплитудными замираниями.

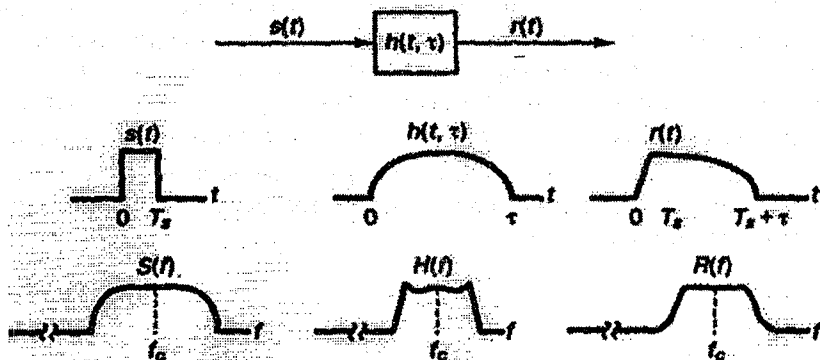


Рис. 11.6. Характеристики канала с частотно-селективным замиранием.

Параметры T_0 и $R(\Delta t)$ несут информацию о скорости замирания в канале. Поскольку малые расстояния (порядка 13 см для несущей 900 МГц) соответствуют статистической декорреляции принятых сигналов, основные проявления замирания, дисперсия сигнала и скорость замирания могут рассматриваться независимо друг от друга.

При движении мобильного радиопередатчика к новому пространственному положению, которое характеризуется иным профилем, будут происходить изменения в состоянии замирания канала, как обуславливает профиль в новом местоположении. Однако вследствие того, что один профиль декоррелирует с другим уже на расстоянии порядка 13 см. (для несущей 900 МГц), скорость таких изменений зависит только от скорости движения, но не от общей геометрии местности.

Нестационарную природу или механизм скорости замирания в канале можно рассматривать с позиций ухудшения качества передачи — быстрого и медленного замирания. Если $T_0 < T_s$, то имеет место быстрое замирание. Характер замирания в этом случае изменяется несколько раз за время передачи символа, что приводит к искажению вида узкополосного импульса. Чем выше скорость изменения состояния канала, тем большее расширение спектра испытывает сигнал. Данное искажение аналогично внесенной каналом ISI. Если $T_0 > T_s$, то канал вносит медленное замирание, т.е. состояние канала остается практически неизменным в течение времени передачи символа, и основное ухудшение качества передачи в этом случае связано, как и в случае с амплитудным замиранием, с уменьшением SNR.

На рис. 11.4,г заостренность и крутизна границ спектра доплеровских частот является следствием резкого верхнего предела доплеровского сдвига, вызванного перемещением подвижной антенны среди стационарных рассеивающих элементов в модели плотного размещения. Наибольшая величина (бесконечность) $S(\nu)$ соответствует случаю, когда рассеивающий элемент находится прямо перед движущейся платформой антенны или прямо позади нее. В этом случае величина сдвига частот описывается формулой:

$$f_d = \frac{V}{\lambda}, \quad (11.4)$$

где V — относительная скорость, а λ — длина волны сигнала.

Если передатчик и приемник движутся навстречу друг другу, то f_d положительна, а если они удаляются друг от друга, то f_d отрицательна. Что касается рассеивающих элементов, находящихся в направлении поперечного излучения движущейся платформы, то для них величина частотного сдвига равна нулю. Хотя доплеровские компоненты, поступившие точно под углами 0° и 180° , имеют бесконечно большую спектральную плотность мощности, это не представляет проблемы, так как угол имеет непрерывное распределение, а вероятность поступления компонентов точно под этими углами равна нулю.

$S(\nu)$ является Фурье-образом $R(\Delta t)$. Известно, что Фурье-образ автокорреляционной функции временного ряда равен квадрату амплитуды Фурье-образа исходного временного ряда. Следовательно, измерения могут проводиться просто путем передачи синусоиды и с использованием Фурье-анализа для получения спектра мощности полученной амплитуды. Этот доплеровский спектр мощности канала дает информацию о спектральном расширении переданной синусоиды. Знание $S(\nu)$ делает возможным приблизительное вычисление величины расширения спектра как функции скорости изменения состояния канала. Ширина доплеровского спектра мощности f_d в литературе называется по-разному: доплеровское расширение, скорость замирания, ширина полосы замирания, спектральное расширение.

Доплеровское расширение и время когерентности канала обратно пропорциональны:

$$T_0 \approx \frac{1}{f_d}. \quad (11.5)$$

Время, требуемое для прохождения расстояния $\lambda/2$, равное приблизительно времени когерентности, при движении с постоянной скоростью V запишется как

$$T_0 \approx \frac{\lambda/2}{V} = \frac{0,5}{f_d}. \quad (11.6)$$

Пример

Пусть объект мобильной связи движется со скоростью 120 км/ч и принимает сигнал с несущей частотой 900 МГц. Каково время когерентности канала и доплеровское расширение?

Решение

Длина волны принимаемого сигнала равна

$$\lambda = \frac{c}{f} = \frac{3 \cdot 10^8 \text{ м/с}}{9 \cdot 10^8 \text{ Гц}} \approx 0,33 \text{ м}.$$

Время когерентности канала $T_0 \approx \frac{\lambda/2}{V} = \frac{0,16 \text{ м}}{\frac{120 \cdot 10^3}{3600} \text{ м/с}} = 5 \text{ мс}$

Величина доплеровского сдвига $f_d = \frac{0,5}{T_0} = \frac{0,5}{5 \cdot 10^{-3}} = 100 \text{ Гц}$

11.3. Борьба с ухудшением характеристик, вызванным эффектами замирания

Если в результате замирания канал вносит искажения в сигнал, для системы может быть характерен неисправимый уровень ошибок, превышающий допустимую частоту появления ошибок. В этом случае никакое увеличение битового отношения сигнал/шум не поможет достичь желаемого уровня достоверности передачи, и единственно доступным подходом, допускающим улучшение, является использование каких-либо иных методов устранения или уменьшения искажений. Метод

борьбы зависит от того, вызвано ли искажение частотно-селективным или быстрым замиранием.

При наличии в канале быстрого замирания могут применяться коды коррекции ошибок и чередование для улучшения рабочих характеристик системы. В сигнал вводится достаточная избыточность, чтобы скорость передачи символов превышала скорость замирания в канале, но в то же время не превышала ширину полосы когерентности. Выбирается метод модуляции/демодуляции, наиболее устойчивый в условиях быстрого замирания. Это значит, например, что необходимо избегать схем, которые требуют контуров ФАПЧ для восстановления несущей, поскольку быстрое замирание может не позволить контурам ФАПЧ достичь синхронизации.

При наличии частотно-селективных замираний в канале связи для борьбы с вызванной каналом ISI используется выравнивание (эквалайзеры). Суть выравнивания заключается в использовании методов, собирающих рассеянную энергию сигнала в его исходный временной интервал (по сути, это создание обратного фильтра канала). Поскольку в мобильной связи характеристика канала меняется во времени, эквалайзер должен быть адаптивным.

Техника расширения спектра сигнала позволяет сделать длительность импульса сигнальной последовательности намного меньше задержки многолучевого распространения. В этом случае каждый отдельный многолучевой компонент выделяется отдельным коррелятором, так называемым Rake-приемником. При этом для выяснения характеристик канала используется контрольный сигнал.

Ортогональное уплотнение с частотным разделением может использоваться для увеличения периода передачи символа, что позволит избежать применения эквалайзера. Исходная сигнальная цифровая последовательность с высокой скоростью передачи разуплотняется на N групп символов так, чтобы каждая группа содержала последовательность с более низкой скоростью передачи (в N раз меньшую), чем у исходной последовательности символов. Полоса сигнала состоит из N ортогональных несущих, каждая из которых модулируется отличной от других группой символов. Целью является снижение скорости передачи сигналов (и соответственно полосы пропускания), чтобы она была меньше полосы когерентности канала.

Литература:

В.Ю.Бабков, М.А.Вознюк, П.А.Михайлов «Сети мобильной связи», «Передача информации в системах подвижной связи», «Системы связи с кодовым разделением каналов» – СПб, СПбГУТ, 2000

Передача дискретных сообщений: Учебник для ВУЗов/В.П.Шувалов и др.; под ред.В.П.Шувалова.-М.:Радио и связь,-1990.-464с.

Прокис Дж. Цифровая связь. Пер. с англ – М.: Радио и связь, 2000. – 800с.

Рид Р. Основы теории передачи информации.: Пер. с англ. – М.:Издательский дом «Вильямс», 2005

Скляр Б. Цифровая связь. Теоретические основы и практическое применение.- 2-е изд. М., издательский дом «Вильямс», 2003

Филиппов Л.И. Теория передачи дискретных сообщений – М., 1981

Цимер Р., Петерсон Р. Цифровая связь.- в 2-х томах.- М., Техносфера, 2003 .

Приложение 1

Гауссовский интеграл ошибок $Q(x)$

x	$Q(x)$									
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2168	0,2148
0,8	0,2169	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
3,1	0,0010	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002

Пример вычисления аппроксимации функции

$$Q(x) = \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \text{ для } x > 3 \text{ в пакете MATLAB:}$$

`x=4,27;`

`p1=1/(x*sqrt(2*pi));`

`p2=exp((-x^2/2));`

`q=p1*p2`

`q =`

`1,0263e-005`

`x =`

`4,8550`

`p1=1/(x*sqrt(2*pi)); p2=exp((-x^2/2)); q=p1*p2`

`q =`

`6,2566e-007`

Содержание

Лекция 1

1.1 Причина замены аналоговой связи на цифровую	3
1.2 Основные понятия и определения цифровой связи	5
1.3 Типичная блочная диаграмма и основные преобразования в цифровой системе связи	8

Лекция 2

2.1 Классификация сигналов	12
2.2 Преобразование Фурье	15
2.3 Спектральная плотность	21
2.4 Случайные процессы	22
2.5 Связь между скоростью передачи данных и шириной полосы	24
2.6 Автокорреляционная функция сигнала	28
2.7 Импульсная характеристика	30

Лекция 3

3.1 Децибелы	34
3.2 Передача сигнала без искажений по каналу связи	36
3.3 Реализуемые фильтры	38
3.4 Дискретизация аналоговой информации	42

Лекция 4

4.1 Цифровые фильтры	50
4.2 Выборка с запасом	57
4.3 Устойчивость линейных дискретных систем (ЛДС)	60

Лекция 5

5.1 Сообщения, знаки и символы	65
5.2 Квантование аналогового сигнала	66
5.3 Импульсно-кодовая модуляция (ИКМ)	74
5.4 Дифференциальная импульсно-кодовая модуляция (ДИКМ, Differential PCM – DPCM)	77

Лекция 6

6.1 Особенности передачи речи по цифровым каналам связи	82
6.2 Процесс речеобразования	86

Лекция 7

7.1 Типы сигналов импульсно-кодовой модуляции.....	91
7.2 Характеристики кодов канала	95
7.3 Логическое кодирование для улучшения характеристик потенциальных кодов	100
7.4 Спектральные параметры сигналов PCM	103

Лекция 8

8.1 Причины применения полосовой модуляции.....	106
8.2 Амплитудная модуляция	108
8.3 Демодуляция амплитудно-модулированных сигналов	110
8.4 Балансная модуляция	112
8.5 Угловая модуляция.....	116
8.6 Квадратурная амплитудная модуляция.....	119

Лекция 9

9.1 Важнейший параметр цифровой связи – отношение сигнал/шум.....	121
9.2 Векторное представление синусоиды	124
9.3 Расстояние между тонами цифровой частотной манипуляции.....	126
9.4 Вероятность ошибки при использовании для связи наборов ортogonalных и неортogonalных сигналов	129

Лекция 10

10.1 Демодуляция и обнаружение	132
10.2 Вероятность ошибки	137
10.3 Согласованный фильтр	142

Лекция 11

11.1 Формирование импульсов с целью снижения межсимвольной интерференции	147
11.2 Сложность связи по каналу с замираниями.....	149
11.3 Борьба с ухудшением характеристик, вызванным эффектами замирания	159
Литература.....	161
Приложение 1.....	162
Приложение 2.....	163

Учебное издание

Чернецова Елена Анатольевна

ТЕОРИЯ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

Конспект лекций

Редактор Л.В. Ковель

Заказ № 185. Подписано в печать 19.06.2006. Бумага офсетная.

Формат 60x84¹/₁₆. Объем 10,37 п. л. Тираж 100 экз.

Издательство РГГМУ

195196 Санкт-Петербург, Малоохтинский пр., 98

Отпечатано с готового оригинал-макета в ЦНИТ «АСТЕРИОН»

Санкт-Петербург, 191015, а/я 83,

тел. /факс (812) 275-73-00, 275-53-92, тел. 970-35-70

www.asterion.ru E-mail: asterion@asterion.ru