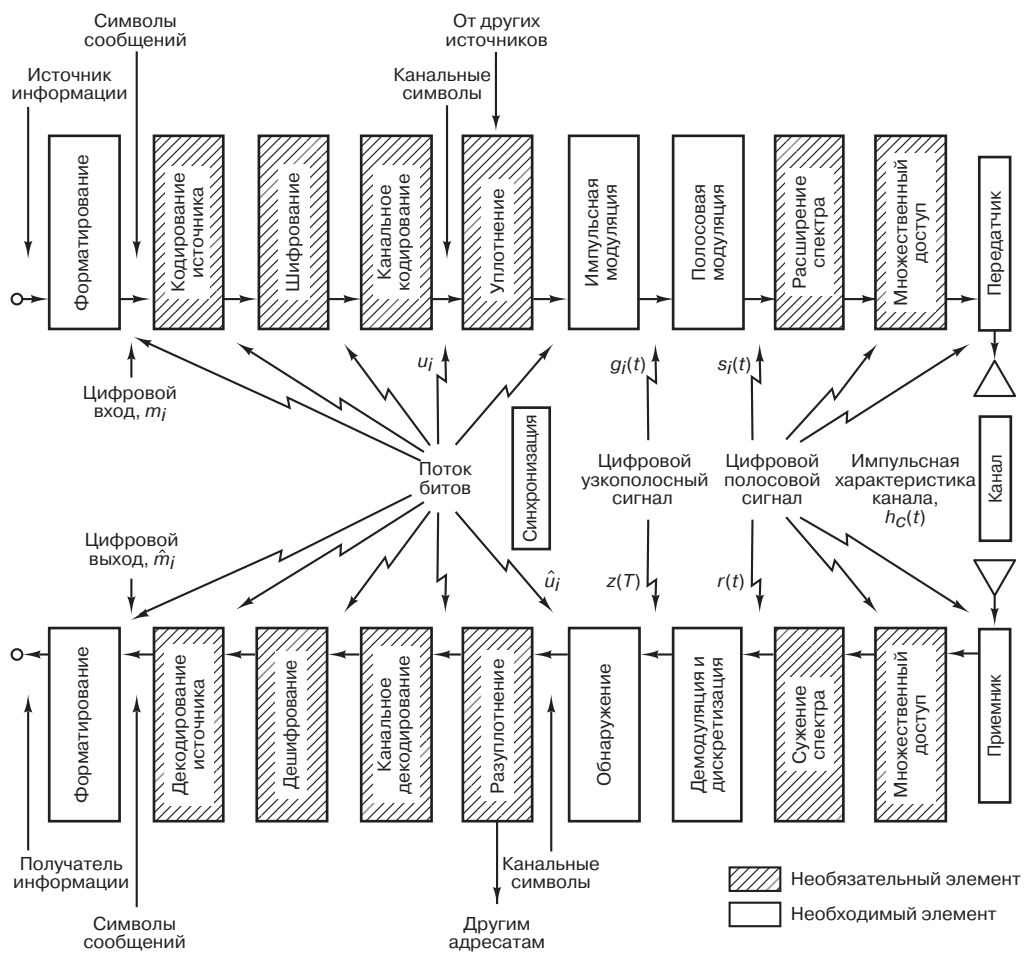


## Сигналы и спектры



В данной книге излагаются идеи и технологии, являющиеся фундаментальными для систем цифровой связи. Основное внимание обращается на вопросы проектирования систем и необходимость компромиссов между основными параметрами системы, такими как отношение сигнал/шум (signal-to-noise ratio — SNR), вероятность появления ошибки и эффективность использования полосы. Мы рассмотрим передачу информации (речь, видео или данные) по каналу связи, где средой передачи является проводник, волновод или окружающая среда.

Системы цифровой связи становятся все более привлекательными вследствие постоянно растущего спроса и из-за того, что цифровая передача предлагает возможности обработки информации, не доступные при использовании аналоговой передачи. В данной книге цифровые системы часто рассматриваются в контексте спутникового канала связи. Иногда это трактуется в контексте систем мобильной радиосвязи, в этом случае передача сигнала обычно ухудшается вследствие явления, называемого *замиранием*. Здесь стоит отметить, что спроектировать и описать систему связи, противостоящую замиранию, сложнее, чем выполнить то же для системы без замирания.

Отличительной особенностью систем цифровой связи (digital communication system — DCS) является то, что за конечный промежуток времени они посылают сигнал, состоящий из конечного набора элементарных сигналов (в отличие от систем аналоговой связи, где сигнал состоит из бесконечного множества элементарных сигналов). В системах DCS задачей приемника является *не* точное воспроизведение переданного сигнала, а определение на основе искаженного шумами сигнала, какой именно сигнал из конечного набора был послан передатчиком. Важным критерием производительности системы DCS является вероятность ошибки ( $P_E$ ).

## 1.1. Обработка сигналов в цифровой связи

### 1.1.1. Почему “цифровая”

Почему в военных и коммерческих системах связи используются “цифры”? Существует множество причин. Основным преимуществом такого подхода 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; и (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 со-

стояния 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.

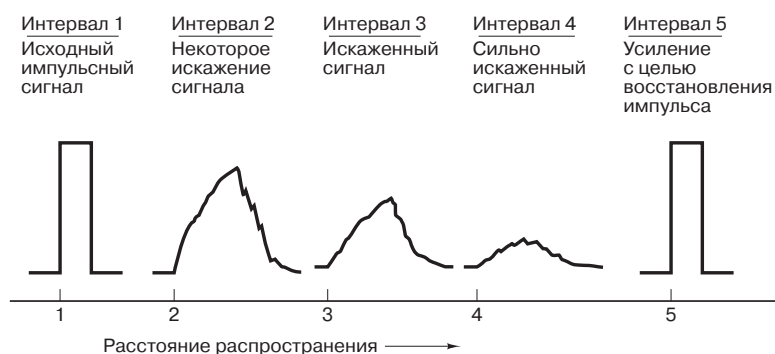


Рис. 1.1. Искажение и восстановление импульса

Существуют и другие важные 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 (large-scale integrated circuit — LSI)). 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. (Подобные технологии рассматриваются в главах 12 и 14.) 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 систем необходимо выделение значительной части ресурсов для синхронизации на различных уровнях (см. главу 10). Аналоговые системы, наоборот, легче синхронизировать. Еще одним недостатком систем цифровой связи является то, что *ухудше-*

ние 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.

### 1.1.2. Типичная блочная диаграмма и основные преобразования

Функциональная блочная диаграмма, приведенная на рис. 1.2, иллюстрирует распространение сигнала и этапы его обработки в типичной системе цифровой связи (DCS). Этот рисунок является чем-то вроде плана, направляющего читателя по главам данной книги. 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. 1.2; 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 (radio frequency — RF), усилителя мощности и антенны, а приемник — из антенны и малошумящего усилителя (low-noise amplifier — LNA). Обратное понижение частоты производится на выходе приемника и/или демодулятора.

На рис. 1.2 иллюстрируется соответствие блоков верхней (передающей) и нижней (принимающей) частей системы. Этапы обработки сигнала, имеющие место в передатчике, являются преимущественно обратными к этапам приемника. На рис. 1.2 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. (Сравните подобный конечный алфавит систем DCS с тем, что мы имеем в аналоговых системах, когда сигнал сообщения является элементом бесконечного множества возможных сигналов.) 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 *поток битов* (рис. 1.2).

Рассмотрим ключевые блоки обработки сигналов, изображенные на рис. 1.2; необходимыми для систем DCS 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 DCS. С этой точки рисунка и вплоть до блока импульсной модуляции информация остается в форме *потока битов*.

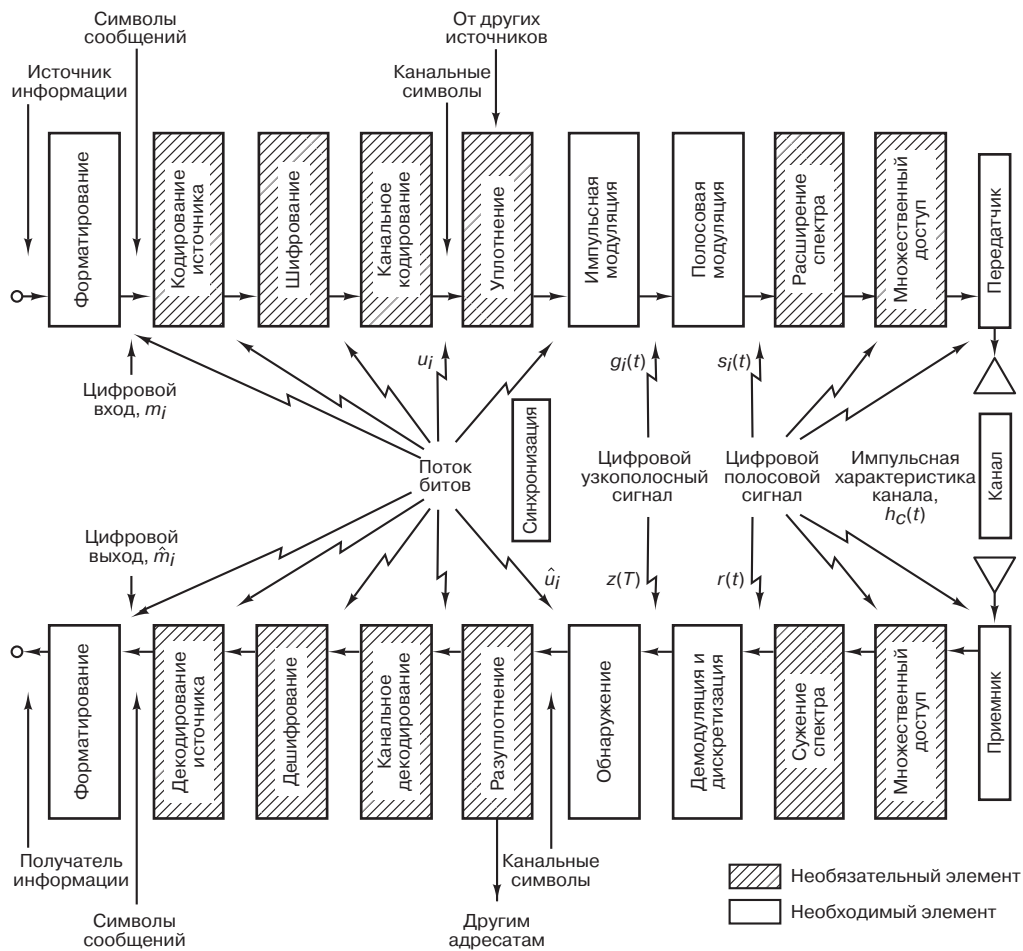


Рис. 1.2. Блочная диаграмма типичной системы цифровой связи

Модуляция — это процесс, посредством которого символы сообщений или каналные 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 (описанных в главе 2); в приложениях телефонной связи эти сигналы часто называются *кодами канала*. При применении импульсной модуляции к небинарным симво-

лам результирующий сигнал именуется  $M$ -арным импульсно-модулированным. Существует несколько типов подобных сигналов, которые также описаны в главе 2, где основное внимание уделяется *амплитудно-импульсной модуляции* (pulse-amplitude modulation — PAM). 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 фактически 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  $M$ -арные сигналы. В разделе 2.8.2 описаны различные полезные параметры этих типов сигналов. 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 “полосовой” (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*  $h_c(t)$  (см. раздел 1.6.1). Кроме того, в различных точках вдоль маршрута сигнала дополнительные случайные шумы искажают принятый сигнал  $r(t)$ , поэтому прием должен выражаться через поврежденную версию сигнала  $s_i(t)$ , поступающего от передатчика. Принятый сигнал  $r(t)$  можно выразить следующим образом:

$$r(t) = s_i(t) * h_c(t) + n(t) \quad i = 1, \dots, M, \quad (1.1)$$

где знак “\*” представляет собой операцию свертки (см. приложение А), а  $n(t)$  — процесс шума (см. раздел 1.5.5).

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 огибающего узкополосного сигнала  $z(t)$ . Обычно с приемником и демодулятором связано несколько фильтров — фильтрация производится для удаления нежелательных высокочастотных составляющих (в процессе преобразования полосового сигнала в узкополосный) и формирования импульса. Выравнивание можно описать как разновидность фильтрации, используемой в демодуляторе (или после демодулятора) для удаления всех эффектов ухудшения качества сигнала, причиной которых мог быть канал. Выравнивание (equalization) необходимо в том случае, если импульсная характеристика канала

$h_c(t)$  настолько плоха, что принимаемый сигнал сильно искажен. Эквалайзер (устройство выравнивания) реализуется для компенсации (т.е. для удаления или ослабления) всех искажений сигнала, вызванных неидеальной характеристикой  $h_c(t)$ . И последнее, этап дискретизации преобразовывает сформированный импульс  $z(t)$  в выборку  $z(T)$  для восстановления (приблизительно) символа канала  $\hat{u}_i$  или символа сообщения  $\hat{m}_i$  (если не используется канальное кодирование). Некоторые авторы используют термины “демодуляция” и “обнаружение” как синонимы. В данной книге под *демодуляцией* (demodulation) подразумевается восстановление сигнала (полосового импульса), а под *обнаружением* (detection) — принятие решения относительно цифрового значения этого сигнала.

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 (для аналоговых источников) и удаление избыточной (ненужной) информации. Отметим, что типичная система DCS может использовать либо *кодирование источника* (для оцифровывания и сжатия исходной информации), либо более простое преобразование *форматирование* (только для оцифровывания). Система не может одновременно применять и кодирование источника, и форматирование, поскольку первое уже включает необходимый этап оцифровывания информации. Шифрование, которое используется для обеспечения секретности связи, предотвращает понимание сообщения несанкционированным пользователем и введение в систему ложных сообщений. *Канальное кодирование* (channel coding) при данной скорости передачи данных может снизить вероятность ошибки  $P_E$  или уменьшить отношение сигнал/шум, необходимое для получения желаемой вероятности  $P_E$  за счет увеличения полосы передачи или усложнения декодера. Процедуры *уплотнения* (multiplexing) и *множественного доступа* (multiple access) объединяют сигналы, которые могут иметь различные характеристики или могут поступать от разных источников, с тем, чтобы они могли совместно использовать часть ресурсов связи (например, спектр, время). *Расширение частоты* (frequency spreading) может давать сигнал, относительно неустойчивый для интерференции (как естественной, так и умышленной), и может использоваться для повышения конфиденциальности общающихся сторон. Также оно является ценной технологией, используемой для множественно доступа.

Блоки обработки сигналов, показанные на рис. 1.2, представляют типичную схему системы цифровой связи; впрочем, эти блоки иногда реализуются в несколько ином порядке. Например, уплотнение может происходить до канального кодирования *или* модуляции *либо* — при двухэтапном процессе модуляции (поднесущая и несущая) — оно может выполняться между двумя этапами модуляции. Подобным образом блок расширения частоты может находиться в различных местах верхнего ряда рис. 1.2; точное его местонахождение зависит от конкретной используемой технологии. Синхронизация и ее ключевой элемент, синхронизирующий сигнал, задействованы во всех этапах обработки сигнала в системе DCS. Для простоты блок синхронизации на рис. 1.2 показан безотносительно к чему-либо, хотя фактически он участвует в регулировании операций практически в каждом блоке, приведенном на рисунке.

На рис. 1.3 показаны основные функции обработки сигналов (которые можно рассматривать как преобразования сигнала), разбитые на следующие девять групп.

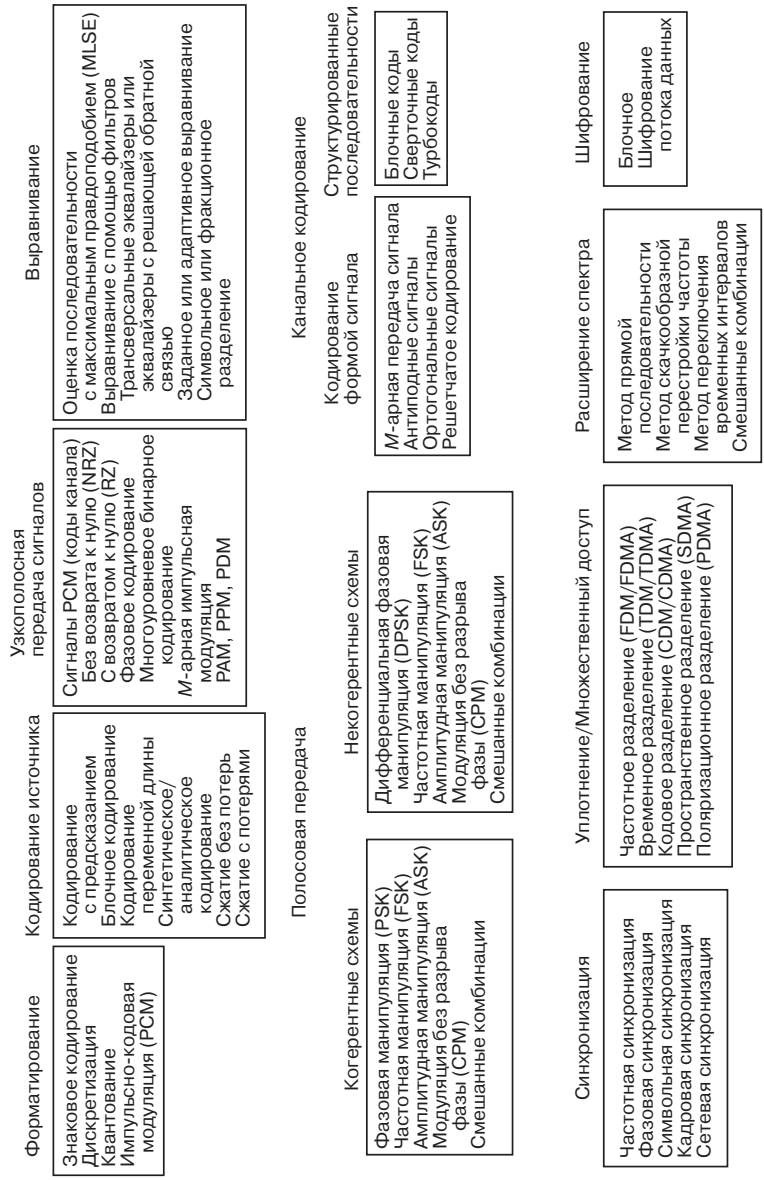


Рис. 1.3. Основные преобразования цифровой связи



1. Форматирование и кодирование источника
2. Узкополосная передача сигналов
3. Полосовая передача сигналов
4. Выравнивание
5. Канальное кодирование
6. Уплотнение и множественный доступ
7. Расширение спектра
8. Шифрование
9. Синхронизация

Хотя пункты такого разделения частично перекрываются, все же это позволяет удобно упорядочить материал книги. Начиная с главы 2 поочередно рассматриваются все девять основных преобразований. В главе 2 исследуются основные методы форматирования, используемые для преобразования исходной информации в символы сообщений. Кроме того, здесь описывается выбор узкополосного сигнала и фильтрация импульса, обеспечивающая совместимость символов сообщений с узкополосной передачей. Обратные этапы демодуляции, выравнивания, дискретизации и обнаружения представлены в главе 3. Форматирование и кодирование источника являются подобными процессами, поскольку оба включают оцифровывание данных. Впрочем, термин “кодирование источника” подразумевает дополнительное сжатие данных и рассматривается позднее (глава 13) как частный случай форматирования.

На рис. 1.3 блок *Узкополосная передача сигналов* содержит перечень бинарных альтернатив при использовании модуляции РСМ или линейных кодов. В этом блоке также указана небинарная категория сигналов, называемая *M*-арной импульсной модуляцией. Еще одно преобразование на рис. 1.3, помеченное как *Полосовая передача сигналов*, разделено на два основных блока, когерентный и некогерентный. Демодуляция обычно выполняется с помощью *опорных* сигналов. При использовании известных сигналов в качестве меры всех параметров сигнала (особенно фазы) процесс демодуляции называется *когерентным*; когда информация о фазе не используется, процесс именуется *некогерентным*. Обе технологии подробно описаны в главе 4.

Глава 5 посвящена *анализу канала связи*. Среди множества спецификаций, анализов и табличных представлений, поддерживающих разработку систем связи, анализ канала связи занимает особое место, поскольку позволяет представить общую картину системы. В главе 5 воедино сводятся все основные понятия канала связи, необходимые для анализа большинства систем связи.

Канальное кодирование связано с методами, используемыми для улучшения цифровых сигналов, которые в результате становятся менее уязвимыми к таким факторам ухудшения качества, как шум, замирание и подавление сигнала. На рис. 1.3 канальное кодирование разделено на два блока, блок кодирования формой сигнала и блок структурированных последовательностей. *Кодирование формой сигнала* включает использование новых сигналов, привносящих улучшенное качество обнаружения по сравнению с исходным сигналом. Структурированные последовательности включают применение дополнительных битов для определения наличия ошибки, вызванной шумом в канале. Одна из таких технологий, *автоматический запрос повторной передачи* (automatic repeat request — ARQ), просто распознает появление ошибки и запрашивает отправителя повторно передать сообщение; другая технология, известная как *прямая коррекция ошибок*

(forward error correction — FEC), позволяет автоматически исправлять ошибки (с определенными ограничениями). При рассмотрении структурированных последовательностей мы обсудим три распространенных метода — блочное, сверточное и турбокодирование. Вначале в главе 6 описывается *линейное блочное кодирование*. В главе 7 мы рассмотрим *сверточное кодирование*, декодирование Витерби (и другие алгоритмы декодирования) и сравним аппаратные и программные процедуры кодирования. В главе 8 представлено каскадное кодирование, которое привело к созданию класса кодов, известных как *турбокоды*, а также подробно рассмотрены коды *Рида-Соломона*.

В главе 9 обобщаются вопросы проектирования систем связи и представляются различные компромиссы из областей модуляции и кодировки, которые обязательно должны быть рассмотрены при проектировании системы. Обсуждаются теоретические ограничения, такие как критерий Найквиста и предел Шеннона. Также исследуются схемы модуляции, позволяющие эффективно использовать полосу, такие как решетчатое кодирование.

Глава 10 посвящена *синхронизации*. В цифровой связи синхронизация включает вычисление как времени, так и частоты. Как показано на рис. 1.3, синхронизация выполняется на пяти уровнях. Эталонные частоты когерентных систем требуется синхронизировать с несущей (и возможно, поднесущей) по частоте и фазе. Для некогерентных систем синхронизация фазы не обязательна. Основной процесс синхронизации по времени — это символьная синхронизация (или битовая синхронизация для бинарных символов). Демодулятор и детектор должны знать, когда начинать и заканчивать процесс обнаружения символа и бита; ошибка синхронизации приводит к снижению эффективности обнаружения. Следующий уровень синхронизации по времени, кадровая синхронизация, позволяет перестраивать сообщения. И последний уровень, сетевая синхронизация, позволяет скоординировать действия с другими пользователями с целью эффективного использования ресурсов. В главе 10 мы рассмотрим синхронизацию пространственно разделенных периодических процессов.

В главе 11 описаны *уплотнение* и *множественный доступ*. Значения этих двух терминов очень похожи; оба связаны с идеей совместного использования ресурсов. Основным отличием является то, что уплотнение реализуется локально (например, на печатной плате, в компоновочном узле или даже на аппаратном уровне), а множественный доступ — удаленно (например, несколькими пользователями требуется совместно использовать спутниковый транспондер). При уплотнении применяется алгоритм, известный априорно; обычно он внедрен непосредственно в систему. Множественный доступ, наоборот, обычно адаптивен и может требовать для работы некоторых служебных издержек. В главе 11 мы рассмотрим классические способы совместного использования ресурсов связи: частотное, временное и кодовое разделение. Кроме того, будут описаны некоторые технологии множественного доступа, возникшие в результате использования спутниковой связи.

В главе 12 вводится преобразование, изначально разработанное для военной связи и известное как *расширение* (spreading). Здесь рассмотрены методы расширения спектра, важные для получения защиты от интерференции и обеспечения секретности. Сигналы могут расширяться по частоте, времени или по частоте и времени. В основном в главе обсуждается расширение частоты. Также глава иллюстрирует применение методом расширения частоты для совместного использования ресурсов с ограниченной полосой в коммерческой переносной телефонии.

В главе 13 рассматривается *кодирование источника*, которое включает эффективное описание исходной информации. Оно связано с процессом компактного описания

сигнала согласно заданным критериям точности. Кодирование источника может применяться и к цифровым, и аналоговым сигналам; путем уменьшения избыточности информации коды источника могут снизить системную скорость передачи данных. Следовательно, основным преимуществом кодирования источника является возможность уменьшения объема требуемых ресурсов системы (например, ширины полосы).

Глава 14 посвящена *шифрованию* и *дешифрованию*, основными задачами которых является аутентификация и обеспечение конфиденциальности связи. Поддержание конфиденциальности означает предотвращение извлечения информации из канала несанкционированными лицами (“подслушивание”). Аутентификация подразумевает предотвращение ввода в канал ложных сигналов несанкционированными лицами. В этой главе значительное внимание уделяется стандарту шифрования данных (data encryption standard — DES) и основным идеям, относящимся к классу систем шифрования, называемых *системы с открытым ключом*. Кроме того, здесь рассмотрена новая схема, названная Pretty Good Privacy (“достаточно хорошая секретность”), которая позволяет эффективно шифровать файлы, предназначенные для отправки по электронной почте.

В последней главе 15 рассмотрены каналы с замираниями. Здесь мы обсудим замирание, которое воздействует на мобильные системы, такие как переносные и персональные системы связи (personal communication system — PCS). В главе перечисляются основные механизмы замирания, типы ухудшения качества и методы борьбы с этим ухудшением. Подробно исследуются два метода: эквалайзер Витерби, реализованный в системе GSM (Global Systems for Mobile Communication — глобальная система мобильной связи), и приемник Рейка, используемый в системах CDMA (Code Division Multiple Access — множественный доступ с кодовым разделением каналов).

### 1.1.3. Основная терминология области цифровой связи

Ниже приведены некоторые основные термины, часто используемые в области цифровой связи.

*Источник информации* (information source). Устройство, передающее информацию посредством системы DCS. Источник информации может быть *аналоговым* или *дискретным*. Выход аналогового источника может принимать любое значение из непрерывного диапазона амплитуд, тогда как выход дискретного источника информации — значения из конечного множества амплитуд. Аналоговые источники информации преобразуются в цифровые посредством *дискретизации* или *квантования*. Методы дискретизации и квантования, называемые форматированием и кодированием источника (рис. 1.3), описаны в главах 2 и 13.

*Текстовое сообщение* (textual message). Последовательность символов (рис. 1.4, *a*). При цифровой передаче данных сообщение представляет собой последовательность цифр или символов, принадлежащих конечному набору символов или алфавиту.

*Знак* (Character). Элемент алфавита или набора символов (рис. 1.4, *b*). Знаки могут отображаться в последовательность двоичных цифр. Существует несколько стандартизованных кодов, используемых для знакового кодирования, в том числе код ASCII (American Standard Code for Information Interchange — Американский стандартный код для обмена информацией), код EBCDIC (Extended Binary Coded Decimal Interchange Code — расширенный двоичный код обмена инфор-

мацией), код Холлерита (Hollerith code), код Бодо (Baudot code), код Муррея (Murray code) и код (азбука) Морзе (Morse code).

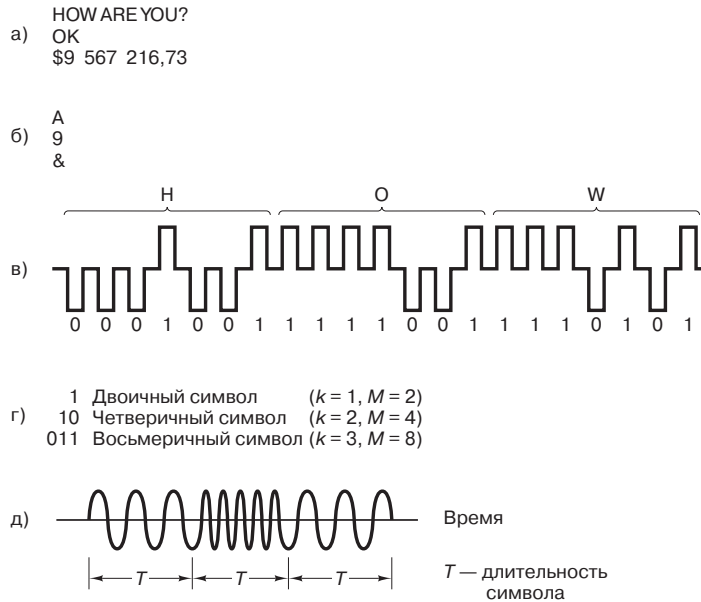


Рис. 1.4. Иллюстрация терминов: а) текстовые сообщения; б) символы; в) поток битов (7-битовый код ASCII); г) символы  $m_i$ ,  $i = 1, \dots, M$ ,  $M = 2^k$ ; д) полосовой цифровой сигнал  $s(t)$ ,  $i = 1, \dots, M$

**Двоичная цифра** (binary digit) (*бит*) (bit). Фундаментальная единица информации для всех цифровых систем. Термин “бит” также используется как единица объема информации, что описывается в главе 9.

**Поток битов** (bit stream). Последовательность двоичных цифр (нулей и единиц). Поток битов часто называют *узкополосным* (baseband) сигналом; это подразумевает, что его спектральные составляющие размещены от (или около) постоянной составляющей до некоторого конечного значения, обычно не превышающего несколько мегагерц. На рис. 1.4, в сообщение “HOW” представлено с использованием семибитового кода ASCII, а поток битов показан в форме двухуровневых импульсов. Последовательность импульсов изображена посредством крайне стилизованных (идеально прямоугольных) сигналов с промежутками между соседними импульсами. В реальной системе импульсы никогда не будут выглядеть так, поскольку подобные промежутки абсолютно бесполезны. При данной скорости передачи данных промежутки увеличат ширину полосы, необходимую для передачи; или, при данной ширине полосы, они увеличат временную задержку, необходимую для получения сообщения.

**Символ** (symbol) (*цифровое сообщение*) (digital message). Символ — это группа из  $k$  бит, рассматриваемых как единое целое. Далее мы будем называть этот блок *символом сообщения* (message symbol)  $m_i$  ( $i = 1, \dots, M$ ) из конечного набора символов или алфавита (рис. 1.4, г.) Размер алфавита  $M$  равен  $2^k$ , где  $k$  — число битов в символе. При *узкополосной* передаче каждый из символов  $m_i$  будет

представлен одним из набора узкополосных импульсных сигналов  $g_1(t)$ ,  $g_2(t)$ , ...,  $g_M(t)$ . Иногда при передаче последовательности таких импульсов для выражения скорости передачи импульсов (скорости передачи символов) используется единица *бод* (baud). Для типичной *полосовой* (bandpass) передачи каждый импульс  $g_i(t)$  будет представляться одним из набора полосовых импульсных сигналов  $s_1(t)$ ,  $s_2(t)$ , ...,  $s_M(t)$ . Таким образом, для беспроводных систем символ  $m_i$  посылается путем передачи цифрового сигнала  $s_i(t)$  в течение  $T$  секунд. Следующий символ посылается в течение следующего временного интервала,  $T$ . То, что набор символов, передаваемых системой DCS, является конечным, и есть главное отличие этих систем от систем аналоговой связи. Приемник DCS должен всего лишь определить, какой из  $M$  возможных сигналов был передан; тогда как аналоговый приемник должен точно определять значение, принадлежащее непрерывному диапазону сигналов.

*Цифровой сигнал* (digital waveform). Описываемый уровнем напряжения или тока, сигнал (импульс — для узкополосной передачи или синусоида — для полосовой передачи), представляющий цифровой символ. Характеристики сигнала (для импульсов — амплитуда, длительность и расположение или для синусоиды — амплитуда, частота и фаза) позволяют его идентифицировать как один из символов конечного алфавита. На рис. 1.4,  $d$  приведен пример полосового цифрового сигнала. Хотя сигнал является синусоидальным и, следовательно, имеет аналоговый вид, все же он именуется *цифровым*, поскольку кодирует цифровую информацию. На данном рисунке цифровое значение указывается посредством передачи в течение каждого интервала времени  $T$  сигнала определенной частоты.

*Скорость передачи данных* (data rate). Эта величина в битах в секунду (бит/с) дается формулой  $R = k/T = (1/T) \log_2 M$  (бит/с), где  $k$  бит определяют символ из  $M = 2^k$ -символьного алфавита, а  $T$  — это длительность  $k$ -битового символа.

#### 1.1.4. Цифровые и аналоговые критерии производительности

Принципиальное отличие систем аналоговой и цифровой связи связано со способом оценки их производительности. Сигналы аналоговых систем составляют континуум, так что приемник должен работать с бесконечным числом возможных сигналов. Критерием производительности аналоговых систем связи является точность, например отношение сигнал/шум, процент искажения или ожидаемая среднеквадратическая ошибка между переданным и принятым сигналами.

В отличие от аналоговых, цифровые системы связи передают сигналы, представляющие цифры. Эти цифры формируют конечный набор или алфавит, и этот набор известен приемнику априорно. Критерием качества цифровых систем связи является вероятность неверного обнаружения цифры или вероятность ошибки ( $P_E$ ).

## 1.2. Классификация сигналов

### 1.2.1. Детерминированные и случайные сигналы

Сигнал можно классифицировать как *детерминированный* (при отсутствии неопределенности относительно его значения в любой момент времени) или *случайный*, в про-

тивном случае. Детерминированные сигналы моделируются математическим выражением  $x(t) = 5 \cos 10t$ . Для случайного сигнала такое выражение написать *невозможно*. Впрочем, при наблюдении случайного сигнала (также называемого *случайным процессом*) в течение достаточно длительного периода времени, могут отмечаться некоторые закономерности, которые можно описать через вероятности и среднее статистическое. Такая модель, в форме вероятностного описания случайного процесса, особенно полезна для описания характеристик сигналов и шумов в системах связи.

### 1.2.2. Периодические и непериодические сигналы

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

$$x(t) = x(t + T_0) \quad \text{для } -\infty < t < \infty, \quad (1.2)$$

где через  $t$  обозначено время. Наименьшее значение  $T_0$ , удовлетворяющее это условие, называется *периодом* сигнала  $x(t)$ . Период  $T_0$  определяет длительность одного полного цикла функции  $x(t)$ . Сигнал, для которого не существует значения  $T_0$ , удовлетворяющего уравнению (1.2), именуется *непериодическим*.

### 1.2.3. Аналоговые и дискретные сигналы

Аналоговый сигнал  $x(t)$  является непрерывной функцией времени, т.е.  $x(t)$  однозначно определяется для всех  $t$ . Электрический аналоговый сигнал возникает тогда, когда физический сигнал (например, речь) некоторым устройством преобразовывается в электрический. Для сравнения, *дискретный сигнал*  $x(kT)$  является сигналом, существующим в дискретные промежутки времени; он характеризуется последовательностью чисел, определенных для каждого момента времени,  $kT$ , где  $k$  — целое число, а  $T$  — фиксированный промежуток времени.

### 1.2.4. Сигналы, выраженные через энергию или мощность

Электрический сигнал можно представить как изменение напряжения  $v(t)$  или тока  $i(t)$  с мгновенной мощностью  $p(t)$ , подаваемой на сопротивление  $\mathfrak{R}$ :

$$p(t) = \frac{v^2(t)}{\mathfrak{R}} \quad (1.3,а)$$

или

$$p(t) = i^2(t)\mathfrak{R} \quad (1.3,б)$$

В системах связи мощность часто нормируется (предполагается, что сопротивление  $\mathfrak{R}$  равно 1 Ом, хотя в реальном канале оно может быть любым). Если требуется определить действительное значение мощности, оно получается путем “денормирования” нормированного значения. В нормированном случае уравнения (1.3,а) и (1.3,б) имеют одинаковый вид. Следовательно, вне зависимости от того, представлен сигнал через напряжение или ток, нормированная форма позволяет нам выразить мгновенную мощность как

$$p(t) = x^2(t), \quad (1.4)$$

где  $x(t)$  — это либо напряжение, либо ток. Рассеивание энергии в течение промежутка времени  $(-T/2, T/2)$  реального сигнала с мгновенной мощностью, полученной с помощью уравнения (1.4), может быть записано следующим образом.

$$E_x^T = \int_{-T/2}^{T/2} x^2(t) dt \quad (1.5)$$

Средняя мощность, рассеиваемая сигналом в течение этого интервала, равна следующему.

$$P_x^T = \frac{1}{T} E_x^T = \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (1.6)$$

Производительность системы связи зависит от *энергии* принятого сигнала; сигналы с более высокой энергией обнаруживаются более достоверно (с меньшим числом ошибок) — *работу по обнаружению выполняет принятая энергия*. С другой стороны, *мощность* — это *скорость* поступления энергии. Этот момент важен по нескольким причинам. Мощность определяет напряжение, которое необходимо подать на передатчик, и напряженность электромагнитных полей, которые следует учитывать в радиосистемах (т.е. поля в волноводах, соединяющих передатчик с антенной, и поля вокруг излучающих элементов антенны).

При анализе сигналов связи желательно работать с *энергией сигнала*. Будем называть  $x(t)$  *энергетическим сигналом* тогда и только тогда, когда он в любой момент времени имеет ненулевую конечную энергию ( $0 < E_x < \infty$ ), где

$$E_x = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x^2(t) dt = \int_{-\infty}^{\infty} x^2(t) dt \quad (1.7)$$

В реальной ситуации мы всегда передаем сигналы с конечной энергией ( $0 < E_x < \infty$ ). Впрочем, для описания *периодических сигналов*, которые по определению (уравнение (1.2)) существуют всегда и, следовательно, имеют бесконечную энергию, и для работы со случайными сигналами, также имеющими неограниченную энергию, удобно определить класс сигналов, выражаемых через *мощность*. Итак, сигнал удобно представить с использованием мощности, если он является *периодическим* и в любой момент времени имеет ненулевую конечную мощность ( $0 < P_x < \infty$ ), где

$$P_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (1.8)$$

Определенный сигнал можно отнести *либо* к энергетическому, *либо* периодическому. Энергетический сигнал имеет конечную энергию, но *нулевую среднюю мощность*, тогда как периодический сигнал имеет нулевую среднюю мощность, но *бесконечную энергию*. Сигнал в системе может выражаться либо через его энергетические, либо периодические значения. *Общее правило*: периодические и случайные сигналы выражаются через мощность, а сигналы, являющиеся детерминированными и непериодическими, — через энергию [1, 2].

Энергия и мощность сигнала — это два важных параметра в описании системы связи. Классификация сигнала либо как энергетического, либо как периодического является удобной моделью, облегчающей математическую трактовку различных сигналов и шумов. В разделе 3.1.5 эти идеи развиваются в контексте цифровых систем связи.

### 1.2.5. Единичная импульсная функция

Полезной функцией в теории связи является единичный импульс, или *дельта-функция Дирака*  $\delta(t)$ . Импульсная функция — 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.

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (1.9)$$

$$\delta(t) = 0 \text{ для } t \neq 0 \quad (1.10)$$

$$\delta(t) \text{ не ограничена в точке } t = 0 \quad (1.11)$$

$$\int_{-\infty}^{\infty} x(t) \delta(t - t_0) dt = x(t_0) \quad (1.12)$$

Единичный импульс  $\delta(t)$  — это не функция в привычном смысле этого слова. 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. Графически  $\delta(t - t_0)$  можно изобразить как пик, расположенный в точке  $t = t_0$ , высота которого равна интегралу от него или его площади. Таким образом,  $A\delta(t - t_0)$  с постоянной  $A$  представляет импульсную функцию, площадь которой (или вес) равна  $A$ , а значение везде нулевое, за исключением точки  $t = t_0$ .

Уравнение (1.12) известно как *просеивающее* (или *квантующее*) *свойство* единичной импульсной функции; интеграл от единичного импульса и произвольной функции дает выборку функции  $x(t)$  в точке  $t = t_0$ .

## 1.3. Спектральная плотность

*Спектральная плотность* (spectral density) характеристик сигнала — это распределение энергии или мощности сигнала по диапазону частот. Особую важность это понятие приобретает при рассмотрении фильтрации в системах связи. Мы должны иметь возможность оценить сигнал и шум на выходе фильтра. При проведении подобной оценки используется спектральная плотность энергии (energy spectral density — ESD) или спектральная плотность мощности (power spectral density — PSD).

### 1.3.1. Спектральная плотность энергии

Общая энергия действительного энергетического сигнала  $x(t)$ , определенного в интервале  $(-\infty, \infty)$ , описывается уравнением (1.7). Используя теорему Парсеваля [1], мы мо-



жем связать энергию такого сигнала, выраженную во временной области, с энергией, выраженной в частотной области:

$$E_x = \int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} |X(f)|^2 df, \quad (1.13)$$

где  $X(f)$  — Фурье-образ непериодического сигнала  $x(t)$ . (Краткие сведения об анализе Фурье можно найти в приложении А.) Обозначим через  $\psi_x(f)$  прямоугольный амплитудный спектр, определенный как

$$\psi_x(f) = |X(f)|^2 \quad (1.14)$$

Величина  $\psi_x(f)$  является *спектральной плотностью энергии* (ESD) сигнала  $x(t)$ . Следовательно, из уравнения (1.13) можно выразить общую энергию  $x(t)$  путем интегрирования спектральной плотности по частоте.

$$E_x = \int_{-\infty}^{\infty} \psi_x(f) df \quad (1.15)$$

Данное уравнение показывает, что энергия сигнала равна площади под  $\psi_x(f)$  на графике в частотной области. Спектральная плотность энергии описывает энергию сигнала на единицу ширины полосы и измеряется в Дж/Гц. Положительные и отрицательные частотные компоненты дают равные энергетические вклады, поэтому, для реального сигнала  $x(t)$ , величина  $|X(f)|$  представляет собой четную функцию частоты. Следовательно, спектральная плотность энергии симметрична по частоте относительно начала координат, а общую энергию сигнала  $x(t)$  можно выразить следующим образом.

$$E_x = 2 \int_0^{\infty} \psi_x(f) df \quad (1.16)$$

### 1.3.2. Спектральная плотность мощности

Средняя мощность  $P_x$  действительного сигнала в периодическом представлении  $x(t)$  определяется уравнением (1.8). Если  $x(t)$  — это периодический сигнал с периодом  $T_0$ , он классифицируется как сигнал в периодическом представлении. Выражение для средней мощности периодического сигнала дается формулой (1.6), где среднее по времени берется за один период  $T_0$ .

$$P_x = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x^2(t) dt \quad (1.17,а)$$

Теорема Парсеваля для действительного периодического сигнала [1] имеет вид

$$P_x = \frac{1}{T} \int_{-T_0/2}^{T_0/2} x^2(t) dt = \sum_{n=-\infty}^{\infty} |c_n|^2, \quad (1.17,б)$$

где члены  $|c_n|$  являются комплексными коэффициентами ряда Фурье для периодического сигнала (см. приложение А).

Чтобы использовать уравнение (1.17,б), необходимо знать только значение коэффициентов  $|c_n|$ . Спектральная плотность мощности (PSD)  $G_x(f)$  периодического сигнала  $x(t)$ , которая является действительной, четной и неотрицательной функцией частоты и дает распределение мощности сигнала  $x(t)$  по диапазону частот, определяется следующим образом.

$$G_x(f) = \sum_{n=-\infty}^{\infty} |c_n|^2 \delta(f - nf_0) \quad (1.18)$$

Уравнение (1.18) определяет спектральную плотность мощности периодического сигнала  $x(t)$  как последовательность взвешенных дельта-функций. Следовательно, PSD периодического сигнала является дискретной функцией частоты. Используя PSD, определенную в уравнении (1.18), можно записать среднюю нормированную мощность действительного сигнала.

$$P_x = \int_{-\infty}^{\infty} G_x(f) df = 2 \int_0^{\infty} G_x(f) df \quad (1.19)$$

Уравнение (1.18) описывает PSD только периодических сигналов. Если  $x(t)$  — непериодический сигнал, он *не может быть* выражен через ряд Фурье; если он является непериодическим сигналом в периодическом представлении (имеющим бесконечную энергию), он *может не иметь* Фурье-образа. Впрочем, мы по-прежнему можем выразить спектральную плотность мощности таких сигналов в *пределе*. Если сформировать *усеченную версию*  $x_T(t)$  непериодического сигнала в периодическом представлении  $x(t)$ , взяв для этого только его значения из интервала  $(-T/2, T/2)$ , то  $x_T(t)$  будет иметь конечную энергию и соответствующий Фурье-образ  $X_T(f)$ . Можно показать [2], что спектральная плотность мощности непериодического сигнала  $x(t)$  определяется как предел.

$$G_x(f) = \lim_{T \rightarrow \infty} \frac{1}{T} |X_T(f)|^2 \quad (1.20)$$

### Пример 1.1. Средняя нормированная мощность

- Найдите среднюю нормированную мощность сигнала  $x(t) = A \cos \pi f_0 t$ , используя усреднение по времени.
- Выполните п. а путем суммирования спектральных коэффициентов.

*Решение*

- Используя уравнение (1.17,а), имеем следующее.

$$\begin{aligned} P_x &= \frac{1}{T} \int_{-T_0/2}^{T_0/2} (A^2 \cos^2 2\pi f_0 t) dt = \\ &= \frac{A^2}{2T_0} \int_{-T_0/2}^{T_0/2} (1 + \cos 4\pi f_0 t) dt = \end{aligned}$$

$$= \frac{A^2}{2T_0}(T_0) = \frac{A^2}{2}$$

б) Используя уравнения (1.18) и (1.19), получаем следующее.

$$G_x(f) = \sum_{n=-\infty}^{\infty} |c_n|^2 \delta(f - nf_0)$$

$$\left. \begin{array}{l} c_1 = c_{-1} = \frac{A}{2} \\ c_n = 0 \text{ для } n = 0, \pm 2, \pm 3, \dots \end{array} \right\} \text{(см. приложение А)}$$

$$G_x(f) = \left(\frac{A}{2}\right)^2 \delta(f - f_0) + \left(\frac{A}{2}\right)^2 \delta(f + f_0)$$

$$P_x = \int_{-\infty}^{\infty} G_x(f) df = \frac{A^2}{2}$$

## 1.4. Автокорреляция

### 1.4.1. Автокорреляция энергетического сигнала

Корреляция — это процесс согласования; *автокорреляцией* называется согласование сигнала с собственной запаздывающей версией. Автокорреляционная функция действительного энергетического сигнала  $x(t)$  определяется следующим образом.

$$R_x(\tau) = \int_{-\infty}^{\infty} x(t)x(t + \tau)dt \text{ для } -\infty < \tau < \infty \quad (1.21)$$

Автокорреляционная функция  $R_x(\tau)$  дает меру схожести сигнала с собственной копией, смещенной на  $\tau$  единиц времени. Переменная  $\tau$  играет роль параметра сканирования или поиска.  $R_x(\tau)$  — это не функция времени; это всего лишь функция разности времен  $\tau$  между сигналом и его смещенной копией.

Автокорреляционная функция действительного *энергетического* сигнала имеет следующие свойства.

- |   |   |
|---|---|
| 1. $R_x(\tau) = R_x(-\tau)$                     | симметрия по $\tau$ относительно нуля   |
| 2. $R_x(\tau) \leq R_x(0)$ для всех $\tau$      | максимальное значение в нуле  |
| 3. $R_x(\tau) \leftrightarrow \psi_x(f)$        | автокорреляция и ESD являются Фурье-образами друг друга, что обозначается двусторонней стрелкой |
| 4. $R_x(0) = \int_{-\infty}^{\infty} x^2(t) dt$ | значение в нуле равно энергии сигнала   |

При удовлетворении пп. 1–3  $R_x(\tau)$  является автокорреляционной функцией. Условие 4 — следствие условия 3, поэтому его не обязательно включать в основной набор для проверки на автокорреляционную функцию.

### 1.4.2. Автокорреляция периодического сигнала

Автокорреляция действительного периодического сигнала  $x(t)$  определяется следующим образом.

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t + \tau) dt \quad \text{для } -\infty < \tau < \infty \quad (1.22)$$

Если сигнал  $x(t)$  является периодическим с периодом  $T_0$ , среднее по времени в уравнении (1.22) можно брать по *одному периоду*  $T_0$ , а автокорреляцию выразить следующим образом.

$$R_x(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)x(t + \tau) dt \quad \text{для } -\infty < \tau < \infty \quad (1.23)$$

Автокорреляция *периодического* сигнала, принимающего действительные значения, имеет свойства, сходные со свойствами энергетического сигнала.

- |   |   |
|---|---|
| 1. $R_x(\tau) = R_x(-\tau)$                                 | симметрия по $\tau$ относительно нуля                   |
| 2. $R_x(\tau) \leq R_x(0)$ для всех $\tau$                  | максимальное значение в нуле                            |
| 3. $R_x(\tau) \leftrightarrow G_x(f)$                       | автокорреляция и PSD являются Фурье-образами друг друга |
| 4. $R_x(0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x^2(t) dt$ | значение в нуле равно средней мощности сигнала          |

## 1.5. Случайные сигналы

Основной задачей системы связи является передача информации по каналу связи. Все полезные сигналы сообщений появляются случайным образом, т.е. приемник не знает заранее, какой из возможных символов сообщений будет передан. Кроме того, вследствие различных электрических процессов возникают шумы, которые сопровождают информационные сигналы. Следовательно, нам нужен эффективный способ описания случайных сигналов.

### 1.5.1. Случайные переменные

Пусть *случайная переменная*  $X(A)$  представляет функциональное отношение между случайным событием  $A$  и действительным числом. Для удобства записи обозначим случайную переменную через  $X$ , а ее функциональную зависимость от  $A$  будем считать явной. Случайная переменная может быть дискретной или непрерывной. *Распределение*  $F_X(x)$  случайной переменной  $X$  находится выражением:

$$F_X(x) = P(X \leq x), \quad (1.24)$$

где  $P(X \leq x)$  — вероятность того, что значение принимаемой случайной переменной  $X$  меньше действительного числа  $x$  или равно ему. Функция распределения  $F_X(x)$  имеет следующие свойства.

1.  $0 \leq F_X(x) \leq 1$
2.  $F_X(x_1) \leq F_X(x_2)$ , если  $x_1 \leq x_2$
3.  $F_X(-\infty) = 0$
4.  $F_X(+\infty) = 1$

Еще одной полезной функцией, связанной со случайной переменной  $X$ , является *плотность вероятности*, которая записывается следующим образом.

$$p_X(x) = \frac{dF_X(x)}{dx} \quad (1.25,а)$$

Как и в случае функции распределения, плотность вероятности — это функция действительного числа  $x$ . Название “функция плотности” появилось вследствие того, что вероятность события  $x_1 \leq X \leq x_2$  равна следующему.

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(X \leq x_2) - P(X \leq x_1) = \\ &= F_X(x_2) - F_X(x_1) = \\ &= \int_{x_1}^{x_2} p_X(x) dx \end{aligned} \quad (1.25,б)$$

Используя уравнение (1.25,б), можно приближенно записать вероятность того, что случайная переменная  $X$  имеет значение, принадлежащее очень малому промежутку между  $x$  и  $x + \Delta x$ .

$$P(x \leq X \leq x + \Delta x) \approx p_X(x) \Delta x \quad (1.25,в)$$

Таким образом, в пределе при  $\Delta x$ , стремящемся к нулю, мы можем записать следующее.

$$P(X = x) = p_X(x) dx \quad (1.25,г)$$

Плотность вероятности имеет следующие свойства.

1.  $p_X(x) \geq 0$ .
2.  $\int_{-\infty}^{\infty} p_X(x) dx = F_X(+\infty) - F_X(-\infty) = 1$ .

Таким образом, плотность вероятности всегда неотрицательна и имеет единичную площадь. В тексте книги мы будем использовать запись  $p_X(x)$  для обозначения плотности вероятности для непрерывной случайной переменной. Для удобства записи мы часто будем опускать индекс  $X$  и писать просто  $p(x)$ . Если случайная переменная  $X$

может принимать только *дискретные* значения, для обозначения плотности вероятности мы будем использовать запись  $p(X = x_i)$ .

### 1.5.1.1. Среднее по ансамблю

*Среднее значение* (mean value)  $m_X$ , или *математическое ожидание* (expected value), случайной переменной  $X$  определяется выражением

$$m_X = \mathbf{E}\{X\} = \int_{-\infty}^{\infty} xp_X(x) dx, \quad (1.26)$$

где  $\mathbf{E}\{\cdot\}$  именуется *оператором математического ожидания* (expected value operator). *Моментом  $n$ -го порядка* распределения вероятностей случайной переменной  $X$  называется следующая величина.

$$\mathbf{E}\{X^n\} = \int_{-\infty}^{\infty} x^n p_X(x) dx \quad (1.27)$$

Для анализа систем связи важны первые два момента переменной  $X$ . Так, при  $n = 1$  уравнение (1.27) дает момент  $m_X$ , рассмотренный выше, а при  $n = 2$  — среднеквадратическое значение  $X$ .

$$\mathbf{E}\{X^2\} = \int_{-\infty}^{\infty} x^2 p_X(x) dx \quad (1.28)$$

Можно также определить *центральные моменты*, представляющие собой моменты разности  $X$  и  $m_X$ . Центральный момент второго порядка (называемый также *дисперсией*) равен следующему.

$$\text{var}(X) = \mathbf{E}\{(X - m_X)^2\} = \int_{-\infty}^{\infty} (x - m_X)^2 p_X(x) dx \quad (1.29)$$

Дисперсия  $X$  также записывается как  $\sigma_X^2$ , а квадратный корень из этой величины,  $\sigma_X$ , называется *среднеквадратическим отклонением*  $X$ . Дисперсия — это мера “разброса” случайной переменной  $X$ . Задание дисперсии случайной переменной ограничивает ширину функции плотности вероятности. Дисперсия и среднеквадратическое значение связаны следующим соотношением.

$$\begin{aligned} \sigma_X^2 &= \mathbf{E}\{X^2 - 2m_X X + m_X^2\} = \\ &= \mathbf{E}\{X^2\} - 2m_X \mathbf{E}\{X\} + m_X^2 = \\ &= \mathbf{E}\{X^2\} - m_X^2 \end{aligned}$$

Таким образом, дисперсия равна разности среднеквадратического значения и квадрата среднего значения.

### 1.5.2. Случайные процессы

Случайный процесс  $X(A, t)$  можно рассматривать как функцию двух переменных: *события*  $A$  и *времени*. На рис. 1.5 представлен пример случайного процесса. Показаны  $N$

выборочных функций времени  $\{X_j(t)\}$ . Каждую из выборочных функций можно рассматривать как выход отдельного генератора шума. Для каждого события  $A_j$  имеем единственную функцию времени  $X(A_j, t) = X_j(t)$  (т.е. выборочную функцию). Совокупность всех выборочных функций называется ансамблем. В любой определенный момент времени  $t_k$ ,  $X(A, t_k)$  — это случайная переменная  $X(t_k)$ , значение которой зависит от события. И последнее, для конкретного события  $A = A_j$  и для конкретного момента времени  $t = t_k$ ,  $X(A_j, t_k)$  — это обычное число. Для удобства записи будем обозначать случайный процесс через  $X(t)$ , а функциональную зависимость от  $A$  будем считать явной.

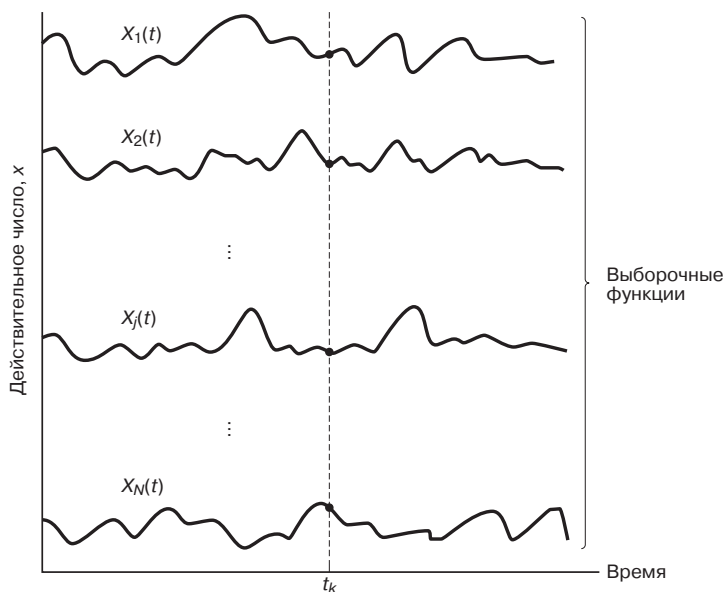


Рис. 1.5. Случайный процесс шума

### 1.5.2.1. Статистическое среднее случайного процесса

Поскольку значение случайного процесса в каждый последующий момент времени неизвестно, случайный процесс, функции распределения которого непрерывны, можно описать статистически через плотность вероятности. Вообще, в различные моменты времени эта функция для случайного процесса будет иметь разный вид. В большинстве случаев эмпирически определить распределение вероятностей случайного процесса нереально. В то же время для нужд систем связи часто достаточно частичного описания, включающего среднее и функцию автокорреляции. Итак, определим среднее случайного процесса  $X(t)$  как

$$\mathbf{E}\{X(t_k)\} = \int_{-\infty}^{\infty} xp_{X_k}(x) dx = m_X(t_k), \quad (1.30)$$

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

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

$$R_X(t_1, t_2) = \mathbf{E}\{X(t_1)X(t_2)\}, \quad (1.31)$$

где  $X(t_1)$  и  $X(t_2)$  — 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.

### 1.5.2.2. Стационарность

Случайный процесс  $X(t)$  называется *стационарным в строгом смысле*, если ни на одну из его статистик не влияет перенос начала отсчета времени. Случайный процесс именуется *стационарным в широком смысле*, если две его статистики, среднее и автокорреляционная функция, не меняются при переносе начала отсчета времени. Таким образом, процесс является стационарным в широком смысле, если

$$\mathbf{E}\{X(t)\} = m_X = \text{константа} \quad (1.32)$$

и

$$R_X(t_1, t_2) = R_X(t_1 - t_2) \quad (1.33)$$

Стационарность в строгом смысле подразумевает стационарность в широком смысле, но не наоборот. Большинство полезных результатов теории связи основывается на предположении, что случайные информационные сигналы и шум являются стационарными в широком смысле. С практической точки зрения случайный процесс не обязательно всегда должен быть стационарным, достаточно стационарности в некотором наблюдаемом интервале времени, представляющем практический интерес.

Для стационарных процессов автокорреляционная функция в уравнении (1.33) зависит не от времени, а только от разности  $t_1 - t_2$ . Иными словами, все пары значений  $X(t)$  в моменты времени, разделенные промежутком  $\tau = t_1 - t_2$ , имеют одинаковое корреляционное значение. Следовательно, для стационарных систем функцию  $R_X(t_1, t_2)$  можно записывать просто как  $R_X(\tau)$ .

### 1.5.2.3. Автокорреляция случайных процессов, стационарных в широком смысле

Как дисперсия предлагает меру случайности для случайных переменных, так и автокорреляционная функция предлагает подобную меру для случайных процессов. Для процессов, стационарных в широком смысле, автокорреляционная функция зависит только от разности времен  $\tau = t_1 - t_2$ .

$$R_X(\tau) = \mathbf{E}\{X(t)X(t + \tau)\} \text{ для } -\infty < \tau < \infty \quad (1.34)$$

Для стационарного в широком смысле процесса с нулевым средним, функция  $R_X(\tau)$  показывает, насколько статистически коррелируют случайные величины процесса, разделенные  $\tau$  секундами. Другими словами,  $R_X(\tau)$  дает информацию о частотной характеристике, связанной со случайным процессом. Если  $R_X(\tau)$  меняется медленно по мере увеличения  $\tau$  от нуля до некоторого значения, это показывает, что в среднем выборочные значения  $X(t)$ , взятые в моменты времени  $t = t_1$  и  $t = t_2$ , практически равны. Следовательно, мы вправе ожидать, что в частотном представлении  $X(t)$  будут преоб-



ладать низкие частоты. С другой стороны, если  $R_X(\tau)$  быстро уменьшается по мере увеличения  $\tau$ , стоит ожидать, что  $X(t)$  будет быстро меняться по времени и, следовательно, будет включать преимущественно высокие частоты.

Автокорреляционная функция стационарного в широком смысле процесса, принимающего действительные значения, имеет следующие свойства.

- |  |   |
|--|---|
| 1. $R_X(\tau) = R_X(-\tau)$                | симметрия по $\tau$ относительно нуля   |
| 2. $R_X(\tau) \leq R_X(0)$ для всех $\tau$ | максимальное значение в нуле  |
| 3. $R_X(\tau) \leftrightarrow G_X(f)$      | автокорреляция и спектральная плотность мощности являются Фурье-образами друг друга |
| 4. $R_X(0) = E\{X^2(t)\}$                  | значение в нуле равно средней мощности сигнала                                      |

### 1.5.3. Усреднение по времени и эргодичность

Для вычисления  $m_X$  и  $R_X(\tau)$  путем усреднения по ансамблю нам нужно усреднить их по всем выборочным функциям процесса, и, значит, нам потребуется полная информация о взаимном распределении функций плотности вероятности в первом и втором приближениях. 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. Чтобы случайный процесс был эргодическим, он должен быть стационарным в строгом смысле (обратное не обязательно). Впрочем, для систем связи, где нам достаточно стационарности в широком смысле, нас интересуют только среднее и автокорреляционная функция.

Говорят, что случайный процесс является *эргодическим по отношению к среднему значению*, если

$$m_X = \lim_{T \rightarrow \infty} 1/T \int_{-T/2}^{T/2} X(t) dt, \quad (1.35)$$

и *эргодическим по отношению к автокорреляционной функции*, если

$$R_X(\tau) = \lim_{T \rightarrow \infty} 1/T \int_{-T/2}^{T/2} X(t) X(t + \tau) dt. \quad (1.36)$$

Проверка случайного процесса на эргодичность обычно весьма непроста. На практике, как правило, используется интуитивное предположение о целесообразности замены средних по ансамблю средними по времени. При анализе большинства сигналов в каналах связи (при отсутствии импульсных эффектов) разумным будет предположение, что случайные сигналы являются эргодическими по отношению к автокорреляционной функции. Поскольку для эргодических процессов средние по времени равны средним по ансамблю, фундаментальные электротехнические параметры, такие как амплитуда постоянной составляющей, среднеквадратическое значение и средняя мощность, могут быть связаны с моментами эргодического случайного процесса.

1. Величина  $m_X = \mathbf{E}\{X(t)\}$  равна постоянной составляющей сигнала.
2. Величина  $m_X^2$  равна нормированной мощности постоянной составляющей.
3. Момент второго порядка  $X(t)$ ,  $\mathbf{E}\{X_2(t)\}$ , равен общей средней нормированной мощности.
4. Величина  $\sqrt{\mathbf{E}\{X^2(t)\}}$  равна среднеквадратическому значению сигнала, выраженного через ток или напряжение.
5. Дисперсия  $\sigma_X^2$  равна средней нормированной мощности переменного сигнала.
6. Если среднее процесса равно нулю (т.е.  $m_X = m_X^2 = 0$ ), то  $\sigma_X^2 = \mathbf{E}\{X^2\}$ , а дисперсия равна среднеквадратическому значению или (другая формулировка) дисперсия представляет общую мощность в нормированной нагрузке.
7. Среднеквадратическое отклонение  $\sigma^2$  является среднеквадратическим значением переменного сигнала.
8. Если  $m_X = 0$ , то  $\sigma_X$  — это среднеквадратическое значение сигнала.

#### 1.5.4. Спектральная плотность мощности и автокорреляция случайного процесса

Случайный процесс  $X(t)$  можно отнести к периодическому сигналу, имеющему такую спектральную плотность мощности  $G_X(f)$ , как указано в уравнении (1.20). Функция  $G_X(f)$  особенно полезна в системах связи, поскольку она описывает распределение мощности сигнала по диапазону частот. Спектральная плотность мощности позволяет оценить мощность сигнала, который будет передаваться через сеть с известными частотными характеристиками. Основные свойства функций спектральной плотности мощности можно сформулировать следующим образом.

- |  |   |
|--|---|
| 1. $G_X(f) \geq 0$                           | всегда принимает действительные значения  |
| 2. $G_X(f) = G_X(-f)$                        | для $X(t)$ , принимающих действительные значения                                    |
| 3. $G_X(f) \leftrightarrow R_X(\tau)$        | автокорреляция и спектральная плотность мощности являются Фурье-образами друг друга |
| 4. $P_X = \int_{-\infty}^{\infty} G_X(f) df$ | связь между средней нормированной мощностью и спектральной плотностью мощности      |

На рис. 1.6 приведено визуальное представление автокорреляционной функции и функции спектральной плотности мощности. Что означает термин “корреляция”? 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. Точная копия считается созданной и локализованной на минус бесконечности. Затем мы последовательно перемещаем копию в положительном направлении временной оси и задаем вопрос, насколько они (исходная версия и копия) соответствуют друг другу. Затем мы перемещаем копию еще на один шаг в положительном направлении и задаем вопрос, насколько они совпа-

дают теперь, и т.д. Корреляция между двумя сигналами изображается как функция времени, обозначаемого  $\tau$ ; при этом время  $\tau$  можно рассматривать как параметр сканирования.

На рис. 1.6, *a–г* изображена описанная выше ситуация в некоторые моменты времени. Рис. 1.6, *a* иллюстрирует отдельный сигнал стационарного в широком смысле случайного процесса  $X(t)$ . Сигнал представляет собой случайную двоичную последовательность с положительными и отрицательными (биполярными) импульсами единичной амплитуды. Положительные и отрицательные импульсы появляются с равной вероятностью. Длительность каждого импульса (двоичной цифры) равна  $T$  секунд, а среднее, или величина постоянной составляющей случайной последовательности, равно нулю. На рис. 1.6, *б* показана та же последовательность, смещенная во времени на  $\tau_1$  секунд. Согласно принятым обозначениям, эта последовательность обозначается  $X(t - \tau_1)$ . Предположим, что процесс  $X(t)$  является эргодическим по отношению к автокорреляционной функции, поэтому для нахождения  $R_X(\tau)$  мы можем использовать усреднение по времени вместо усреднения по ансамблю. Значение  $R_X(\tau)$  получается при перемножении двух последовательностей  $X(t)$  и  $X(t - \tau_1)$  с последующим нахождением среднего с помощью уравнения (1.36), которое справедливо для эргодических процессов *только в пределе*. Впрочем, интегрирование по целому числу периодов может дать нам некоторую оценку  $R_X(\tau)$ . Отметим, что  $R_X(\tau_1)$  может быть получено при смещении  $X(t)$  как в положительном, так и отрицательном направлении. Подобный случай иллюстрирует рис. 1.6, *в*, на котором использована исходная выборочная последовательность (рис. 1.6, *a*) и ее смещенная копия (рис. 1.6, *б*). Заштрихованные области под кривой произведения  $X(t)X(t - \tau_1)$  вносят положительный вклад в произведение, а серые области — отрицательный. Интегрирование  $X(t)X(t - \tau_1)$  по времени передачи импульсов дает точку  $R_X(\tau_1)$  на кривой  $R_X(\tau)$ . Последовательность может далее смещаться на  $\tau_2, \tau_3, \dots$ , и каждое такое смещение будет давать точку на общей автокорреляционной функции  $R_X(\tau)$ , показанной на рис. 1.6, *г*. Иными словами, каждой случайной последовательности биполярных импульсов соответствует автокорреляционная точка на общей кривой, приведенной на рис. 1.6, *г*. Максимум функции находится в точке  $R_X(0)$  (наилучшее соответствие имеет место при  $\tau$ , равном нулю, поскольку для всех  $\tau$   $R(\tau) \leq R(0)$ ), и функция спадает по мере роста  $\tau$ . На рис. 1.6, *г* показаны точки, соответствующие  $R_X(0)$  и  $R_X(\tau_1)$ .

Аналитическое выражение для автокорреляционной функции  $R_X(\tau)$ , приведенной на рис. 1.6, *г*, имеет следующий вид [1].

$$R_X(\tau) = \begin{cases} 1 - \frac{|\tau|}{T} & \text{для } |\tau| \leq T \\ 0 & \text{для } |\tau| > T \end{cases} \quad (1.37)$$

Отметим, что автокорреляционная функция дает нам информацию о частоте; она сообщает нам кое-что о полосе сигнала. В то же время автокорреляция — это временная функция; в формуле (1.37) отсутствуют члены, зависящие от частоты. Так как же она дает нам информацию о полосе сигнала?

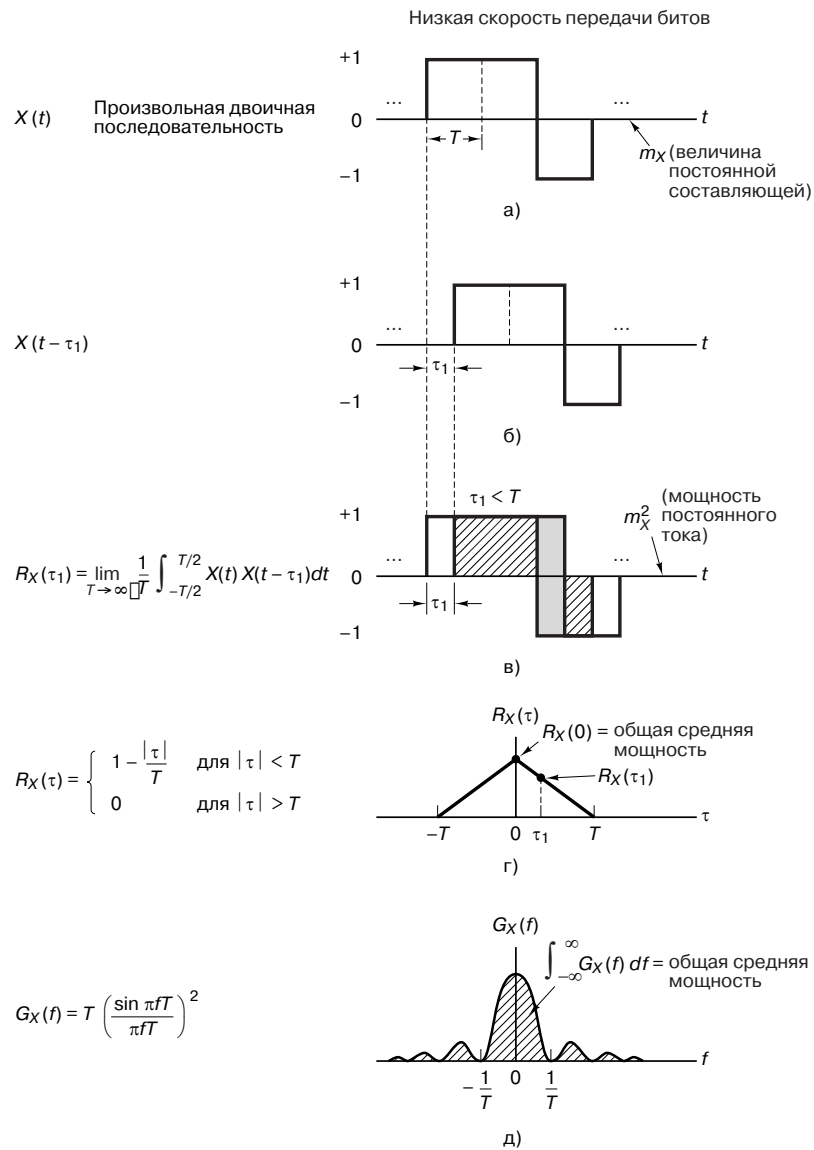


Рис. 1.6. Автокорреляция и спектральная плотность мощности

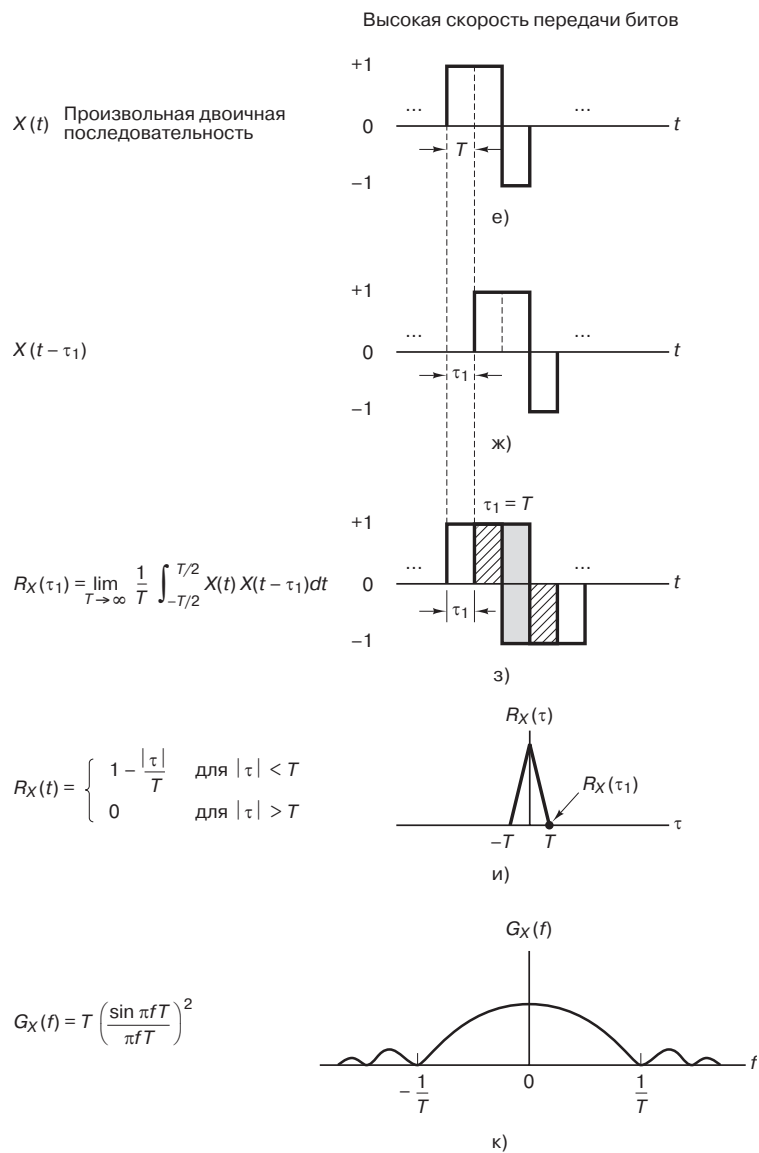


Рис. 1.6. Автокорреляция и спектральная плотность мощности (окончание)

Предположим, что сигнал перемещается очень медленно (сигнал имеет малую ширину полосы). Если мы будем смещать копию сигнала вдоль оси  $\tau$ , задавая на каждом этапе смещения вопрос, насколько соответствуют друг другу копия и оригинал, соответствие достаточно долго будет довольно сильным. Другими словами, треугольная автокорреляционная функция (рис. 1.6, з и формула 1.37) будет медленно спадать с ростом  $\tau$ . Предположим теперь, что сигнал меняется достаточно быстро (т.е. имеем большую полосу). В этом случае даже небольшое изменение  $\tau$  приведет к тому, что корреляция будет нулевой и автокорреляционная функция будет иметь очень узкую форму. Следовательно, сравнение автокорреляционных функций по форме дает нам некоторую инфор-

мацию о ширине полосы сигнала. Функция спадает постепенно? В этом случае имеем сигнал с узкой полосой. Форма функции напоминает узкий пик? Тогда сигнал имеет широкую полосу.

Автокорреляционная функция позволяет явно выражать спектральную плотность мощности случайного сигнала. Поскольку спектральная плотность мощности и автокорреляционная функция являются Фурье-образами друг друга, спектральную плотность мощности,  $G_X(f)$ , случайной последовательности биполярных импульсов можно найти как Фурье-преобразование функции  $R_X(\tau)$ , аналитическое выражение которой дано в уравнении (1.37). Для этого можно использовать табл. А.1. Заметим, что

$$G_X(f) = T \left( \frac{\sin \pi f T}{\pi f T} \right)^2 = T \operatorname{sinc}^2 f T, \quad (1.38)$$

где

$$\operatorname{sinc} y = \frac{\sin \pi y}{\pi y}. \quad (1.39)$$

Общий вид функции  $G_X(f)$  показан на рис. 1.6, *д*.

Отметим, что площадь под кривой спектральной плотности мощности представляет собой среднюю мощность сигнала. Одной из удобных мер *ширины полосы* является ширина основного спектрального лепестка (см. раздел 1.7.2). На рис. 1.6, *д* показано, что ширина полосы сигнала связана с обратной длительностью символа или шириной импульса. Рис. 1.6, *е–к* формально повторяют рис. 1.6, *а–д*, за исключением того, что на последующих рисунках длительность импульса меньше. Отметим, что для более коротких импульсов функция  $R_X(\tau)$  уже (рис. 1.6, *и*), чем для более длительных (рис. 1.6, *з*). На рис. 1.6, *и*  $R_X(\tau_1) = 0$ ; другими словами, в случае меньшей длительности импульса смещения на  $\tau_1$  достаточно для создания нулевого соответствия или для полной потери корреляции между смещенными последовательностями. Поскольку на рис. 1.6, *е* длительность импульса  $T$  меньше (выше скорость передачи импульса), чем на рис. 1.6, *а*, занятость полосы на рис. 1.6, *к* больше занятости полосы для более низкой частоты импульсов, показанной на рис. 1.6, *д*.

### 1.5.5. Шум в системах связи

Термин “шум” обозначает *нежелательные* электрические сигналы, которые всегда присутствуют в электрических системах. Наличие шума, наложенного на сигнал, “затеняет”, или маскирует, сигнал; это ограничивает способность приемника принимать точные решения о значении символов, а следовательно, ограничивает скорость передачи информации. Природа шумов различна и включает как естественные, так и искусственные источники. *Искусственные шумы* — это шумы искрового зажигания, коммутационные импульсные помехи и шумы от других родственных источников электромагнитного излучения. *Естественные шумы* исходят от атмосферы, солнца и других галактических источников.

Хорошее техническое проектирование может устранить большинство шумов или их нежелательные эффекты посредством фильтрации, экранирования, выбора модуляции и оптимального местоположения приемника. Например, чувствительные ра-