Министерство образования и науки Российской Федерации
Национальный исследовательский ядерный университет «МИФИ»

Е.Ф. Березкин

ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ И КОДИРОВАНИЯ

Рекомендовано УМО «Ядерные физика и технологии»
в качестве учебного пособия
для студентов высших учебных заведений

Москва 2010

УДК 519.72(076.5)
ББК 22.18я 7
Б48

Березкин Е.Ф. **Основы теории информации и кодирования**: Учебное пособие. – М.: НИЯУ МИФИ, 2010. – 312 с.

Материал учебного пособия представляет собой теоретическую предпрофилирующую подготовку специалистов по проектированию информационных систем и цифровых комплексов обработки данных. Первые пять разделов посвящены исследованию математических моделей непрерывных сигналов, следующие пять – исследованию информационных моделей дискретных сигналов. Каждый раздел сопровождается задачами с методическими указаниями по их решению и ответами.

Учебное пособие предназначено для студентов, обучающихся по специальности 230102 «Автоматизированные системы обработки информации и управления» и направлению подготовки бакалавров и магистров 230100 «Информатика и вычислительная техника».

Подготовлено в рамках Программы создания и развития НИЯУ МИФИ.

Рецензент канд. техн. наук, доцент С.В. Синицын

***ISBN* 978-5-7262-1294-4**

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Курс «Основы теории информации и кодирования» читается студентам III курса (5-й и 6-й семестры) кафедры «Управляющие интеллектуальные системы». Предлагаемое учебное пособие фактически завершает издание учебно-методических материалов для данного курса. Ранее были изданы учебные пособия (Березкин Е.Ф. **Сборник задач по курсу "Основы теории информации и кодирования".** – М.: МИФИ, 2002; Березкин Е.Ф. **Основы теории информации и кодирования. Лабораторный практикум: Учебно-методическое пособие**. – 2-е изд., перераб. и доп. – М.: МИФИ, 2009), которые предназначены для самостоятельной подготовки студентов к практическим занятиям и лабораторным работам. Кроме теоретического материала в новое учебное пособие частично вошли задачи из указанного выше сборника задач.

Помимо учебных пособий на бумажном носителе для данного курса разработан специализированный компьютерный учебник ОТИК 4.15 нового поколения, который интенсифицирует учебный процесс и обеспечивает формирование знаний, умений и навыков на уровне применения, а также на уровне творчества.

В целом весь комплекс учебных материалов представляет системно-деятельностный подход, который акцентирует внимание на результате образования, причем в качестве результата рассматривается не сумма усвоенной информации (математических моделей), а способность действовать в определенных ситуациях.

В отличие от опубликованных пособий (Панин В.В. **Основы теории информации. Ч.1.** – М.: МИФИ, 2001; Панин В.В. **Основы теории информации. Ч.2. Введение в теорию кодирования.** – М.: МИФИ, 2004), адресованных в первую очередь специалистам в области разработки и эксплуатации электронных измерительных систем, данное пособие ориентировано на будущих проектировщиков информационных систем и цифровых комплексов обработки данных. Учебное пособие содержит более простые объяснения, излагающиеся, по возможности, на языке, понятном бакалаврам и инженерам. Математические рассуждения проводятся на возможно более упрощенном уровне, хотя в некоторых вопросах уровень остается достаточно высоким.

ВВЕДЕНИЕ

Теория информации была создана и развивалась как наука, направленная на решение проблем связи. В настоящее время теория информации является составной частью кибернетики, которая, по выражению А.Н. Колмогорова, «занимается изучением систем любой природы, способных воспринимать, хранить, перерабатывать информацию и использовать ее для управления и регулирования».

Теория информации началась с работ, написанных в конце двадцатых годов XX столетия. Еще в 1928 г. американский ученый Р. Хартли предложил логарифмическую меру оценки количества информации. В 1933 г. была опубликована работа советского ученого В.А. Котельникова, в которой было фактически заложено начало общей теории передачи сообщений.

Наиболее бурное развитие теория информации получила после опубликования в 1947-1948 гг. классических работ американского математика и инженера К. Шеннона. Большой вклад внесли в развитие теории информации американский ученый Н. Винер, советские ученые А.Я. Хинчин, А.Н. Колмогоров.

На сегодняшний день достаточно четко определилась прикладная сторона теории информации – информационная техника, направленная на использование основных положений теории при создании конкретных технических устройств. К информационной технике относятся средства, служащие для восприятия, подготовки, кодирования, передачи, переработки, хранения и представления информации, поступающей от объекта наблюдения.

Теория информации и информационная техника являются сравнительно новыми отраслями, получающими наибольшее развитие на этапе применения вычислительной техники и систем управления, ибо без информации нет управления.

Существуют различные определения информации. Комитет научно-технической терминологии РАН рекомендует следующее определение: ***информация*** – это сведения, являющиеся объектом хранения, передачи и преобразования.

Философия уже много веков бьется над пониманием природы информации. Идеалисты рассматривают ее как духовную субстан-

цию или как комплекс ощущений субъекта. Материалисты считают, что информация есть свойство материи.

Можно стоять на разных позициях в этом вопросе, но необходимо помнить следующее. Информация возникает всегда, когда устанавливаются некоторые общие свойства конкретных вещей и явлений. Под информацией понимают не сами явления, а выделенную сущность, существенные характеристики вещей и явлений. Сама по себе информация может быть отнесена к абстрактной категории идеального, но проявляется она всегда в материально-энергетической форме – в виде сигналов (рис. 1).

Рис. 1. Материально-энергетическая форма информации

Выделим основные аспекты понятия информации:

– семантический, связанный с содержанием и значением сообщения;

– семиотический, связанный с обозначением в знаковой системе;

– коммуникативный, отражающий свойство информации быть средством связи и общения;

– физический, связанный с материальным воплощением информации;

– гносеологический, в котором информация выступает как средство познания;

– аксиологический, указывающий на роль информации для самоуправления;

– казуальный, причинный;

– количественный, наиболее важный для технических приложений.

Теория информации, математическим аппаратом которой являются теория вероятностей и математическая статистика, превратилась к настоящему времени в строгую и достаточно универсальную науку. Положения этой теории все шире и шире используются

при исследовании процессов мышления, познания, психологии, а также общественных явлений.

Теория информации представляет собой общую теорию связи и управления, применимую к любой системе независимо от ее физической природы.

Создание и внедрение информационно-управляющих систем (ИУС) позволяет успешно решать задачи управления объектами, технологическими процессами, координации работ и планирования производства, что, в конечном счете, ускоряет научно-технический прогресс.

В состав сложной ИУС могут входить отдельные самостоятельные автоматизированные комплексы. Неотъемлемыми элементами таких комплексов являются системы восприятия и передачи информации, в которых, в свою очередь, находят применение различные датчики, предназначенные для представления входной информации в виде эквивалентного электрического сигнала, анализаторы случайных процессов, предназначенные для получения обобщенных характеристик источника информации, и каналы связи, различаемые по используемым линиям связи и по физической природе сигналов.

Теоретической основой техники восприятия и передачи информации является статистическая теория связи, которая определяет принципы и основные пути построения оптимальных систем передачи информации.

Основными проблемами теории связи являются проблемы эффективности и достоверности. Проблема эффективности состоит в определении возможных путей, обеспечивающих передачу наибольшего количества информации наиболее экономным способом. Проблема достоверности состоит в определении путей, обеспечивающих при наличии помех максимальное соответствие принятых сообщений переданным.

При таких видах связи, как телефония и телеграфия, информация воспринимается непосредственно человеком и благодаря свойственной человеческому языку избыточности небольшие искажения предаваемого текста не приводят к нарушению достоверности принятых сообщений. Совершенно иначе обстоит дело при передаче данных, где даже незначительное количество ошибок может полностью исказить результаты обработки информации в ЭВМ.

Вопрос о международных нормах на допустимое количество ошибок для различных систем передачи находится в стадии изучения. Однако уже сейчас ясно, что речь может идти о вероятности искажения одного элемента порядка $10^{-5} \div 10^{-9}$. Для обеспечения высокой достоверности применяются системы передачи данных, использующие специальные методы кодирования, и виды модуляции, обеспечивающие максимальную помехоустойчивость.

Итак, задачей системы связи является передача сообщения о каком-либо событии на расстоянии. Расстояние разделяет отправителя и адресата, датчик команд и исполнительное устройство, исследуемый процесс и измерительный механизм, источник излучения и регистрирующий прибор, различные блоки ЭВМ – словом, источник и потребителя информации.

Расстояние, на которое передается сигнал, может быть очень незначительным (передача команд в ЭВМ от одного блока к другому) или огромным (космическая связь). Передача сообщений осуществляется с помощью проводных, кабельных, волноводных линий или в свободном пространстве. Естественно, что для передачи сигналов целесообразно использовать те физические процессы, которые имеют свойство перемещаться в пространстве без существенного затухания. К числу таких процессов относятся применяемые в радиотехнике и радиолокации, например, электромагнитные колебания – радиоволны.

Рассмотрим общую схему ***системы передачи информации***, которая в зависимости от характеристик пары источник – потребитель может претерпевать существенные изменения (рис. 2).

Источник информации представляет собой некоторый физический процесс. Физическому процессу необходимо поставить в соответствие эквивалентный электрический сигнал.

Целью кодирования, как правило, является согласование источника информации с каналом связи, обеспечивающее либо максимально возможную скорость передачи информации, либо заданную помехоустойчивость.

Под сигналом понимается изменяющаяся физическая величина, отображающая физический процесс. Другими словами ***сигнал*** – это материальный переносчик информации.

Физическая среда, по которой происходит передача сигналов от передатчика к приемнику, называется ***линией связи***.

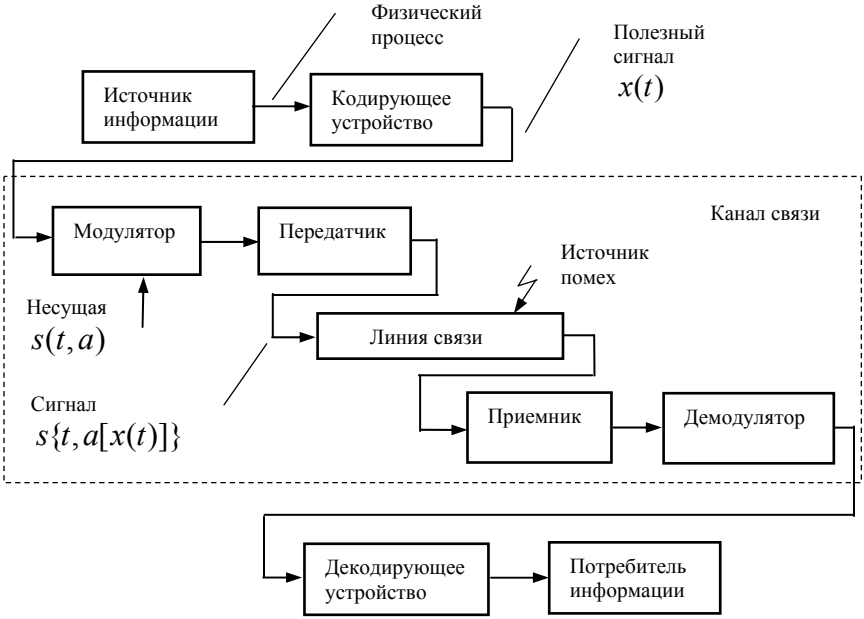


Рис. 2. Общая схема передачи информации

Для передачи исследуемого физического процесса необходимо применять тот переносчик, который способен эффективно распространяться по используемой в системе линии связи. Например, по проводной линии связи наиболее легко проходят постоянный ток и переменные токи невысоких частот (не более нескольких десятков килогерц). По радиолинии эффективно распространяются только электромагнитные колебания высоких частот (до десятков тысяч мегагерц).

Процесс ***модуляции*** заключается в том, что высокочастотное колебание $s(t,a)$, способное распространяться на большие расстояния, наделяются признаками, характеризующими полезное колебание $x(t)$. Таким образом, $s\{t,a[x(t)]\}$ используется как переносчик сообщения, подлежащего передаче. Для этого один или несколько параметров высокочастотного колебания изменяются по закону, совпадающему с законом изменения передаваемого сообщения (рис. 3). В зависимости от материальных носителей и изменяемых параметров различают основные виды модуляции.

Обратное преобразование электромагнитных колебаний в исходный сигнал, осуществляемое на приемной стороне, называется ***демодуляцией***.

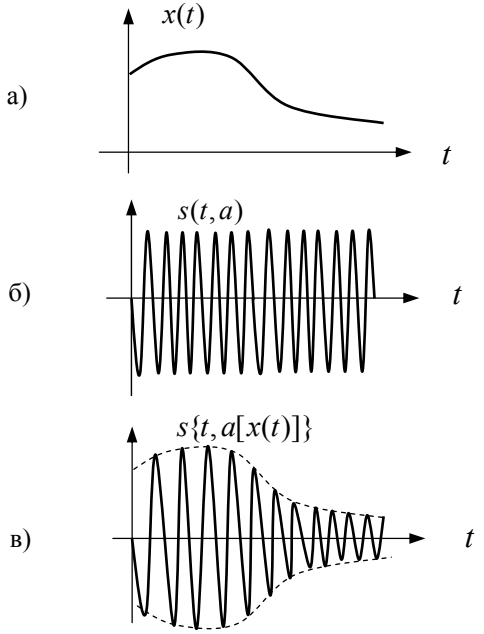


Рис. 3. Пример амплитудной модуляции:
а – полезный сигнал; б – несущая частота; в – модулированный сигнал

Модулятор, выполняя функции генератора несущей частоты, начинает процесс преобразования полезного сигнала в высокочастотный сигнал. Передатчик завершает процесс преобразования, например, выполняя усилительные функции. Приемник, демодулятор и декодирующее устройство реализуют процесс восстановления сформированного источником физического процесса по принятому сигналу.

Для пары источник излучения – регистрирующий прибор схема передачи информации может вырождаться в набор технических средств и сигналов, представленных на рис. 4.

Входной сигнал $s(t)$ подвергается сначала дискретизации по времени с помощью электронного ключа (ЭК), работающего с

шагом T. Дискретизированный сигнал $s_1(t)$ имеет вид последовательности равноотстоящих коротких импульсов, являющихся отсчетами сигнала $s(t)$. Каждый из отсчетов запоминается в интегрирующей RC-цепи на время, необходимое для срабатывания аналого-цифрового преобразователя (АЦП). В результате, на выходе RC-цепи получается ступенчатое колебание $s_2(t)$.

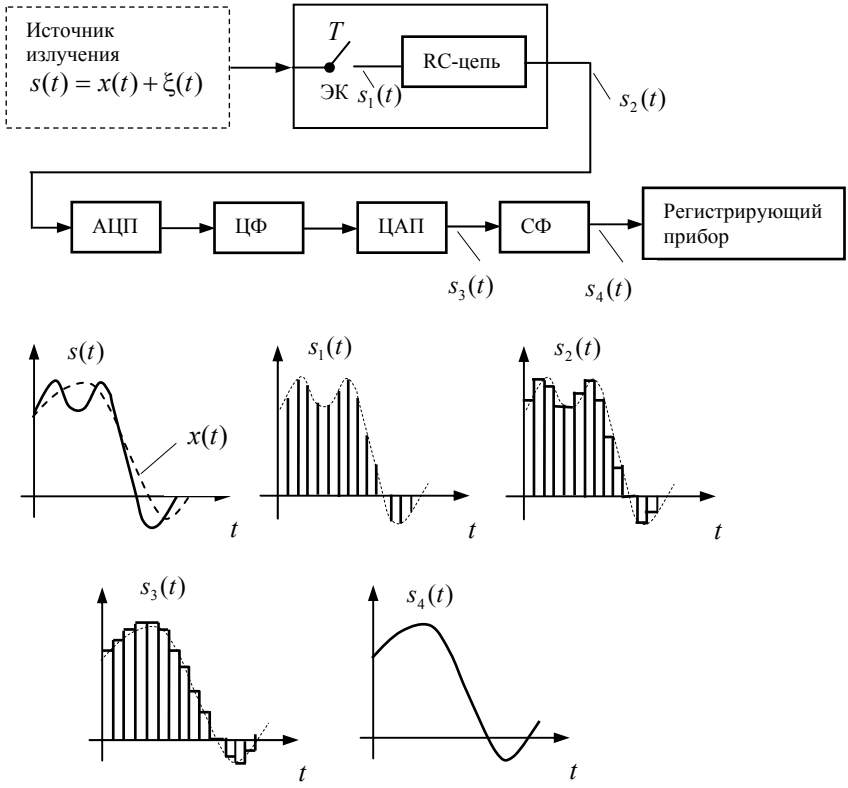


Рис. 4. Схема передачи информации для пары
источник излучения – регистрирующий прибор

В АЦП каждый отсчет квантуется по уровню и преобразуется в кодовое слово. Последовательность кодовых слов обрабатывается в цифровом фильтре (ЦФ). Цифро-аналоговый преобразователь

(ЦАП) осуществляет суммирование эталонных напряжений ($s_3(t)$), соответствующих каждому из разрядов.

Наконец, синтезирующий фильтр (СФ) формирует выходной сигнал $s_4(t)$ максимально похожий на $x(t)$.

Применяемые в современной технике связи сигналы можно разделить на классы, приведенные на рис. 5.

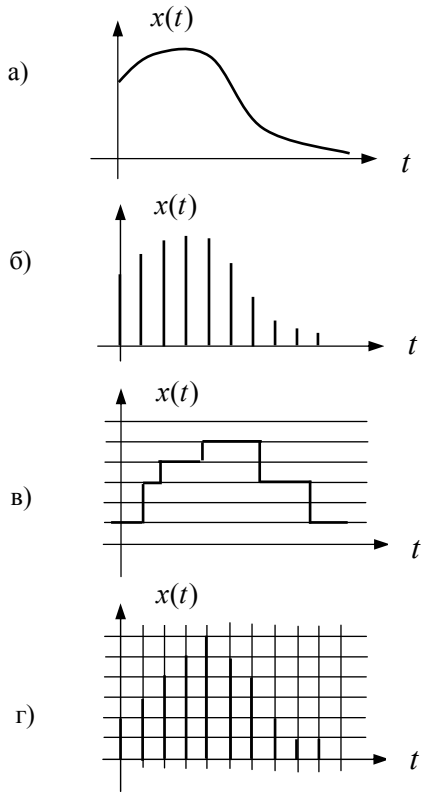


Рис. 5. Классы сигналов:
а – континуальный; б – дискретный; в – квантованный; г – цифровой

В дальнейшем термин «дискретный» будет применяться только по отношению к дискретизации по времени. Дискретность же по уровню будет обозначаться термином «квантование».

Каждому из этих классов сигналов можно поставить в соответствие аналоговую, дискретную или цифровую цепи (рис. 6).

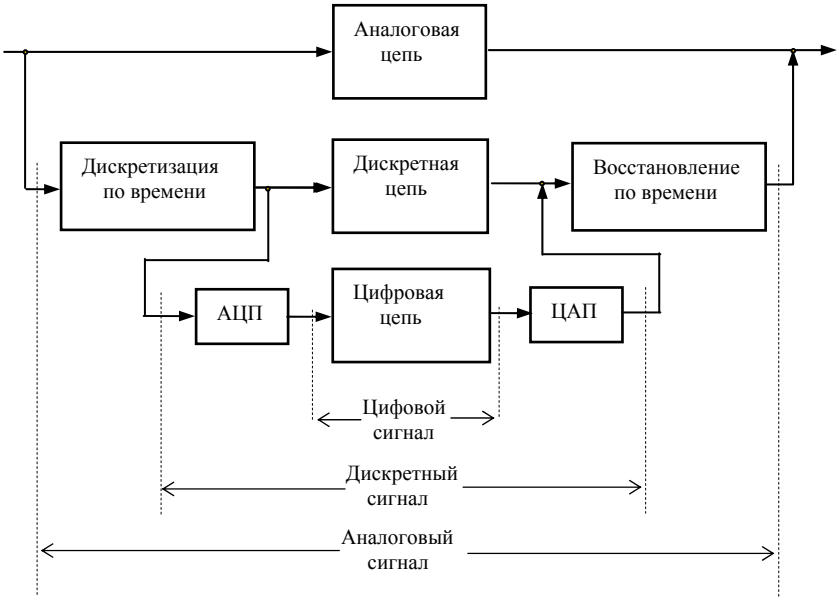


Рис. 6. Классификация электрических цепей

Следует отметить, что в настоящее время цифровая обработка сигналов получает все более широкое применение, что связано не только с ее универсальностью, точностью, но и с возможностями достижений микроэлектроники.

Как отмечалось, ***сигнал*** – изменяющаяся физическая величина, обеспечивающая передачу информации по линии связи. В технических системах используются в большинстве случаев электрические сигналы.

Все многообразие сигналов можно по своим особенностям разделить на две группы: детерминированные и случайные сигналы. Детерминированные сигналы характеризуются тем, что в любые моменты времени их значения являются известными величинами, а случайные – тем, что их значения в любые моменты времени – случайные величины.

Деление сигналов на детерминированные и случайные условно, так как детерминированных сигналов в точном их понимании в

природе нет. На практике нельзя точно предсказать значение сигнала в любые моменты времени, в противном случае сигнал не нес бы полезной информации. Кроме того, любой реальный сигнал случаен в силу воздействия на него многочисленных случайных факторов.

Несмотря на это, исследование детерминированных сигналов весьма важно, так как выводы, полученные в результате анализа математических моделей именно таких сигналов, во многих случаях можно использовать для исследования и анализа случайных сигналов.

Детерминированные сигналы можно подразделить на периодические и непериодические. В реальных условиях периодические сигналы не существуют, так как идеальный периодический сигнал бесконечен во времени, в то время как всякий реальный сигнал имеет начало и конец. Однако во многих случаях конечностью времени действия сигнала можно пренебречь и для его анализа допустимо использовать аппарат, пригодный для идеальных периодических сигналов.

Предлагаемое учебное пособие представляет собой единую научную дисциплину, основу которой составляют теория сигналов, теория случайных процессов, теория информации и теория помехоустойчивого кодирования.

1. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЕТЕРМИНИРОВАННЫХ ПЕРИОДИЧЕСКИХ СИГНАЛОВ

Периодическим называется любой сигнал (рис. 1.1), для которого выполняется условие $x(t) = x(t + kT)$, где период T является конечным отрезком, а k – любое целое число ($k = 0,\pm1,\pm2,...$).

Простейшим периодическим детерминированным сигналом является гармоническое колебание (рис. 1.2), определяемое законом

$$x(t) = A\cos(\frac{2\pi}{T}t + \Theta) = A\cos(\omega t + \Theta), \quad -\infty < t < \infty,$$

где A, T, ω и Θ – постоянные амплитуда, период, угловая частота и начальная фаза колебания.

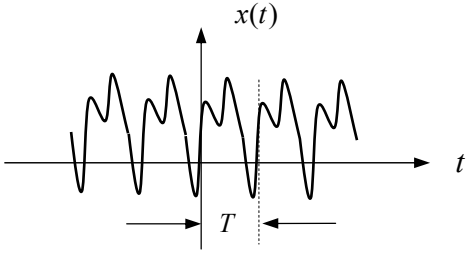


Рис. 1.1. Произвольный периодический сигнал

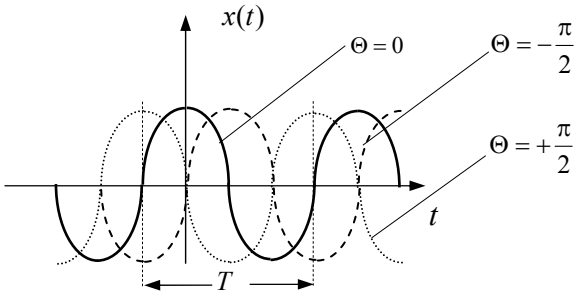


Рис. 1.2. Гармонический сигнал:
$\Theta < 0$ – сдвиг в сторону запаздывания; $\Theta > 0$ – сдвиг в сторону опережения

Любой сложный периодический сигнал, как правило, можно представить в виде суммы гармонических колебаний с частотами, кратными основной частоте $\omega = \frac{2\pi}{T}$. Основной характеристикой сложного периодического сигнала является его спектральная функция, содержащая информацию об амплитудах и фазах отдельных гармоник.

1.1. РАЗЛОЖЕНИЕ ПРОИЗВОЛЬНОГО СИГНАЛА ПО ЗАДАННОЙ СИСТЕМЕ ФУНКЦИЙ

Для техники формирования и обработки сигналов особое значение имеет разложение заданной функции по разным ортогональным системам функций [14]. Напомним основные определения, относящиеся к свойствам ортогональных систем.

Бесконечная система действительных непрерывных функций

$$\varphi_0(x), \varphi_1(x), \varphi_2(x), ..., \varphi_k(x), ..., \varphi_m(x), ...$$

называется ***ортогональной*** на отрезке $[a, b]$, если

$$\int_a^b \varphi_k(x)\varphi_m(x)dx = 0 \text{ при } k \neq m. \qquad (1.1)$$

Отрезок $[a, b]$, на котором выполняется это условие, называется ***интервалом ортогональности***.

При этом предполагается, что $\int_a^b \varphi_k^2(x)dx \neq 0$, т.е. никакая из функций рассматриваемой системы не равна тождественно нулю.

Условие (1.1) выражает попарную ортогональность функций. Величина $\|\varphi_k(x)\| = \sqrt{\int_a^b \varphi_k^2(x)dx}$ называется ***нормой функции*** $\varphi_k(x)$.

Функция $\varphi_k(x)$, для которой выполняется условие $\|\varphi_k(x)\|^2 = \int_a^b \varphi_k^2(x)dx = 1$, называется ***нормированной функцией***, а

система нормированных функций $\varphi_0(x), \varphi_1(x), \varphi_2(x), \ldots$, в которой две различные функции взаимно ортогональны, называется ***ортонормированной*** системой.

В математике доказывается, что если функции $\varphi_k(x)$ непрерывны, то произвольная кусочно-непрерывная функция, для которой выполняется условие

$$\int_L |f(x)| dx < \infty, \tag{1.2}$$

может быть представлена в виде ряда

$$f(x) = \sum_{k=0}^{\infty} C_k \varphi_k(x) = C_0 \varphi_0(x) + C_1 \varphi_1(x) + \ldots \tag{1.3}$$

Интеграл в выражении (1.2) вычисляется по области L определения $f(x)$.

Умножим обе части уравнения (1.3) на $\varphi_k(x)$ и проинтегрируем в пределах $[a,b]$:

$$\int_a^b f(x)\varphi_k(x)dx = C_k \int_a^b \varphi_k^2(x)dx = C_k \|\varphi_k\|^2 .$$

Откуда следует важное соотношение

$$C_k = \frac{1}{\|\varphi_k\|^2} \int_a^b f(x)\varphi_k(x)dx . \tag{1.4}$$

Ряд (1.3), в котором коэффициенты C_k определены по формуле (1.4), называется обобщенным ***рядом Фурье*** по данной системе $\varphi_k(x)$, а сами C_k – ***коэффициентами Фурье***. Обобщенный ряд Фурье обладает следующим важным свойством: при заданной системе функции $\varphi_k(x)$ и при фиксированном числе слагаемых ряда (1.3), он обеспечивает наилучшую аппроксимацию данной функции $f(x)$. Это означает, что среднеквадратическая ошибка, под которой подразумевается величина

$$E = \int_a^b \left[f(x) - \sum_{k=0}^{N} a_k \varphi_k(x) \right]^2 dx ,$$

достигает минимума, когда коэффициенты $a_k = C_k$.

Таким образом,

$$E_{\min} = \int_a^b f^2(x)dx - 2\sum_{k=0}^{N} C_k \int_a^b f(x)\varphi_k(x)dx + \sum_{k=0}^{N} C_k^2 \int_a^b \varphi_k^2(x)dx =$$

$$= \int_a^b f^2(x)dx - 2\sum_{k=0}^{N} C_k^2 \|\varphi_k\|^2 + \sum_{k=0}^{N} C_k^2 \|\varphi_k\|^2 = \int_a^b f^2(x)dx - \sum_{k=0}^{N} C_k^2 \|\varphi_k\|^2.$$

Ортогональная система называется полной, если увеличением числа членов в ряде среднеквадратическую ошибку $E_{\min}$ можно сделать сколь угодно малой.

Условие полноты

$$\int_a^b f^2(x)dx = \sum_{k=0}^{\infty} C_k^2 \|\varphi_k\|^2$$

приобретает энергетический смысл. Действительно, если под $I(t)$ подразумевается электрическое колебание (ток), то

$$Э = \int_{t_1}^{t_2} I^2(t)dt = \sum_{k=0}^{\infty} C_k^2 \|\varphi_k\|^2$$

есть не что иное, как ***энергия сигнала*** в промежутке $t_2 - t_1$ (при условии, что сопротивление, в котором выделяется энергия, равно 1 Ом). При этом имеется в виду, что промежуток времени $t_2 - t_1$, в котором определяется энергия, является интервалом ортогональности.

Очевидно, что средняя за время $t_2 - t_1$ ***мощность сигнала*** равна

$$\overline{I^2(t)} = \frac{Э}{t_2 - t_1} = \frac{1}{t_2 - t_1} \sum_{k=0}^{\infty} C_k^2 \|\varphi_k\|^2.$$

Для бесконечной системы функций $\ldots, \varphi_{-m}(x), \ldots, \varphi_{-k}(x), \ldots, \varphi_{-1}(x), \varphi_0(x), \varphi_1(x), \ldots, \varphi_k(x), \ldots, \varphi_m(x), \ldots$, принимающих комплексные значения, приведенные выше определения обобщаются следующим образом:

– условие ортогональности

$$\int_a^b \varphi_k(x)\varphi_m^*(x)dx = 0 \text{ при } k \neq m;$$

– квадрат нормы функции

$$\left\|\varphi_k(x)\right\|^2 = \int_a^b \varphi_k(x)\varphi_k^*(x)dx = \int_a^b \left|\varphi_k(x)\right|^2 dx;$$

– коэффициент Фурье

$$C_k = \frac{1}{\left\|\varphi_k\right\|^2}\int_a^b f(x)\varphi_k^*(x)dx;$$

– средняя мощность

$$\overline{I^2(t)} = \frac{1}{t_2 - t_1}\sum_{k=-\infty}^{\infty}\left|C_k\right|^2\left\|\varphi_k\right\|^2.$$

В этих выражениях $\varphi_k^*(x)$ обозначает функцию, комплексно-сопряженную функции $\varphi_k(x)$.

1.2. ЧАСТОТНОЕ ПРЕДСТАВЛЕНИЕ ДЕТЕРМИНИРОВАННЫХ ПЕРИОДИЧЕСКИХ СИГНАЛОВ

При разложении периодического колебания $x(t)$ в ряд Фурье в качестве ортогональной системы берут [3]

$$1, \cos\omega_0 t, \sin\omega_0 t, \cos 2\omega_0 t, \sin 2\omega_0 t, ..., \cos k\omega_0 t, \sin k\omega_0 t, ... \qquad (1.5)$$

или

$$..., e^{-j2\omega_0 t}, e^{-j\omega_0 t}, 1, e^{j\omega_0 t}, e^{j2\omega_0 t}, ... \qquad (1.6)$$

Интервал ортогональности в обоих случаях совпадает с периодом $T = 2\pi/\omega_0$ функции $x(t)$.

Система функций (1.5) приводит к тригонометрической форме ряда Фурье, а система (1.6) – к комплексной форме. Между этими двумя формами существует простая связь.

Воспользуемся сначала ортогональной системой (1.6). Тогда ряд Фурье должен быть записан в форме

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\omega_0 t}.$$

Коэффициенты Фурье определяются с помощью формул, приведенных в предыдущем подразделе ($\|\varphi_k\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{jk\omega_0 t} e^{-jk\omega_0 t} dt = T$),

$$C_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) e^{-jk\omega_0 t} dt \,. \qquad (1.7)$$

В этих выражениях учтено, что функции $e^{jk\omega_0 t}$ соответствует комплексно-сопряженная функция $e^{-jk\omega_0 t}$. Коэффициенты C_k в общем случае являются комплексными величинами. Подставив в (1.7) $e^{-jk\omega_0 t} = \cos k\omega_0 t - j \sin k\omega_0 t$, получим

$$C_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cos k\omega_0 t dt - j \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \sin k\omega_0 t dt = C_{kc} - jC_{ks} \,.$$

Часто бывает, что коэффициенты C_k удобно записывать в форме $C_k = |C_k| e^{j\Theta_k}$, где $|C_k| = \sqrt{C_{kc}^2 + C_{ks}^2}$, $\Theta_k = -\operatorname{arctg} \frac{C_{ks}}{C_{kc}}$.

Модуль $|C_k|$ является функцией четной относительно k, аргумент Θ_k – нечетной (последнее вытекает из того, что C_{kc} является четной, а C_{ks} – нечетной функциями k).

Общее выражение ряда можно привести к виду

$$x(t) = \sum_{k=-\infty}^{\infty} |C_k| e^{j(k\omega_0 t + \Theta_k)} \,. \qquad (1.8)$$

Теперь нетрудно перейти к тригонометрической форме ряда Фурье. Выделив из ряда (1.8) пару слагаемых, соответствующую какому-либо заданному значению $|k|$, и учтя соотношения $\Theta_{-k} = -\Theta_k$, $|C_{-k}| = |C_k|$, получим для суммы этих слагаемых

$$|C_{-k}|e^{j(-k\omega_0 t+\Theta_{-k})}+|C_k|e^{j(k\omega_0 t+\Theta_k)}=|C_k|\left[e^{-j(k\omega_0 t+\Theta_k)}+e^{j(k\omega_0 t+\Theta_k)}\right]=$$

$$=2|C_k|\cos(k\omega_0 t+\Theta_k).$$

Отсюда видно, что при переходе к тригонометрической форме ряд (1.8) необходимо записать так:

$$x(t)=C_0+\sum_{k=1}^{\infty}2|C_k|\cos(k\omega_0 t+\Theta_k). \qquad (1.9)$$

Смысл удвоения коэффициентов Фурье C_k в тригонометрическом ряду при $k\geq 1$ становится ясным из рассмотрения векторной диаграммы, приведенной на рис. 1.3.

Вещественная функция $2|C_k|\cos(k\omega_0 t+\Theta_k)$ получается как сумма проекций на горизонтальную ось ОВ двух векторов длиной $|C_k|$, вращающихся с угловой частотой $k\omega_0$ во взаимно противоположных направлениях.

Вектор, вращающийся против часовой стрелки, соответствует положительной частоте, а вектор, вращающийся по часовой стрелке, – отрицательной.

После перехода к тригонометрической форме (1.9) понятие ”отрицательная частота” теряет смысл. Коэффициент C_0 не удваивается, так как в спектре периодического сигнала составляющая с нулевой частотой не имеет ”дублера”.

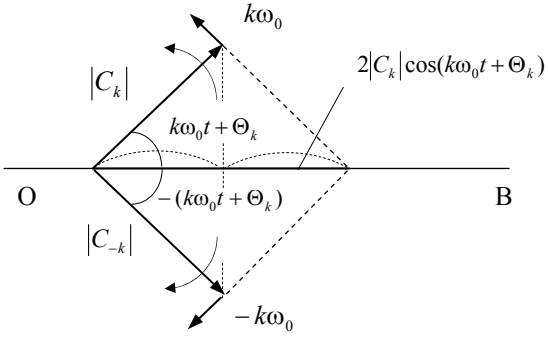


Рис. 1.3. Векторная диаграмма

Вместо выражения (1.9) в технической литературе чаще используется следующая форма записи

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} |A_k| \cos(k\omega_0 t + \Theta_k) = \\ = \frac{A_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega_0 t - b_k \sin k\omega_0 t), \tag{1.10}$$

где $\frac{A_0}{2}$ – ***постоянная составляющая*** функции $x(t)$; $|A_k| \cos(k\omega_0 t + \Theta_k)$ – k-я ***гармоническая составляющая***; $|A_k|, k\omega_0, \Theta_k$ – ***амплитуда***, ***частота*** и ***начальная фаза*** k-й гармонической составляющей; $\omega_0 = \frac{2\pi}{T}$ – частота основной гармоники (T – период колебаний).

Из сопоставления (1.9) и (1.10) видно, что модуль амплитуды k-й гармоники $|A_k|$ связан с модулем коэффициента $|C_k|$ соотношениями

$$|A_k| = 2|C_k| \quad (\frac{A_0}{2} = C_0).$$

Ряд Фурье (1.10) с учетом свойств периодической функции $x(t)$ приобретает еще более простой вид:

$$x(t) = \sum_{k=1}^{\infty} b_k \sin k\omega_0 t \qquad \text{– нечетная функция,}$$

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} a_k \cos k\omega_0 t \qquad \text{– четная функция.}$$

Коэффициенты a_k и b_k вычисляются в соответствии с выражениями

$$a_k = 2C_{kc} = |A_k| \cos\Theta_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos(k\omega_0 t) dt,$$

$$b_k = 2C_{ks} = |A_k| \sin\Theta_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin(k\omega_0 t) dt$$

и связаны между собой формулой $|A_k| = \sqrt{a_k^2 + b_k^2}$.

Соответственно комплексная форма ряда Фурье принимает вид

$$x(t) = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{jk\omega_0 t} = \frac{1}{2}\sum_{k=-\infty}^{\infty} |A_k| e^{j(k\omega_0 t + \Theta_k)}, \qquad (1.11)$$

где $A_k = a_k - jb_k = |A_k| e^{j\Theta_k}$ – комплексная амплитуда гармонической составляющей частоты $k\omega_0$, вычисляемая по формуле

$$A_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t) e^{-jk\omega_0 t} dt = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos(k\omega_0 t)dt - j\frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin(k\omega_0 t)dt ;$$

$\Theta_k = -\text{arctg}\dfrac{b_k}{a_k}$ – начальная фаза, определяемая на интервале $[-\pi, +\pi]$; $\dfrac{A_0}{2} = \dfrac{1}{T}\displaystyle\int_{-T/2}^{T/2} x(t)dt$ – постоянная составляющая.

Совокупность модулей амплитуд и соответствующих частот гармоник называют ***спектром амплитуд*** – $|A_k(\omega)|$, совокупность начальных фаз и соответствующих частот гармоник – ***спектром фаз*** $\Theta_k(\omega)$.

Спектр амплитуд и спектр фаз однозначно определяют сигнал. На рис. 1.4 даны графические изображения спектра амплитуд и спектра фаз гипотетического периодического сигнала.

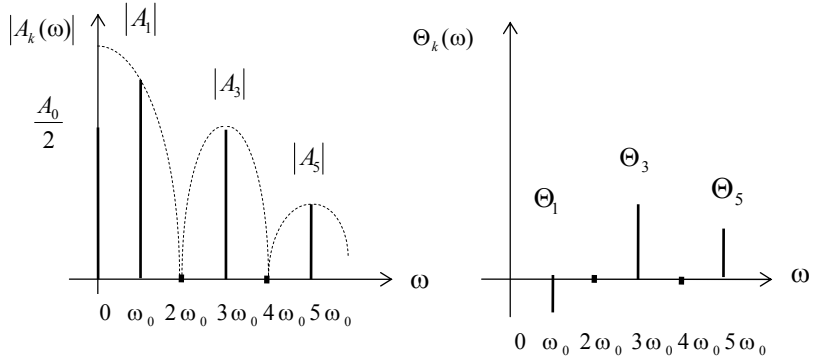


Рис. 1.4. Спектральные характеристики периодического сигнала

Характерной особенностью спектров периодического сигнала является его дискретность. Расстояние между соседними спектральными линиями одинаковое и равно частоте основной гармоники.

1.3. НОСИТЕЛИ ИНФОРМАЦИИ И ВИДЫ МОДУЛЯЦИИ

Нанесение информации на ***материальный носитель*** путем изменения одного из параметров физического процесса в соответствии с законом изменения полезного сигнала называется ***модуляцией***. Модулированный материальный носитель можно трактовать как сигнал, параметр которого содержит информацию [5].

Для образования таких сигналов используются (рис. 1.5) фиксированный уровень, колебание и импульсы любой физической природы. В исходном состоянии эти носители представляют собой как бы чистую поверхность, подготовленную к нанесению необходимых данных – модуляции. Последняя заключается в том, что изменяется какой-либо параметр носителя в соответствии с передаваемой информацией. Эти параметры будем называть информационными.

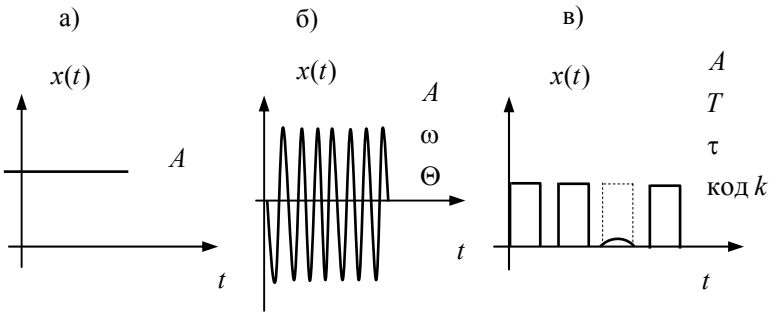


Рис. 1.5. Материальные носители:
а – фиксированный уровень; б – колебание; в – импульсы

Первый тип носителя – фиксированный уровень имеет только один информационный параметр – амплитуду A. Модуляция сводится к такому изменению амплитуды, что она в определенном масштабе представляет передаваемые данные.

Второй тип носителя – колебание содержит три параметра: ам-

плитуду A, фазу Θ и частоту $\omega = \dfrac{2\pi}{T}$.

Третий тип носителя – последовательность импульсов представляет еще большие возможности. Здесь параметрами модуляции могут быть: амплитуда A, частота $f = \dfrac{1}{T}$, длительность импульсов τ и комбинация импульсов и пауз, определяющая код k.

Перечень основных видов модуляции приведен на рис. 1.6.

1.4. АМПЛИТУДНО-МОДУЛИРОВАННЫЙ ГАРМОНИЧЕСКИЙ СИГНАЛ

Амплитудно-модулированный гармонический сигнал можно представить в следующем виде:

$$x(t) = A(t)\cos(\omega_0 t + \Theta_0).$$

Амплитуда такого сигнала изменяется по определенному закону:

$$A(t) = A_0 + \delta A f(t) = A_0[1 + m_A f(t)],$$

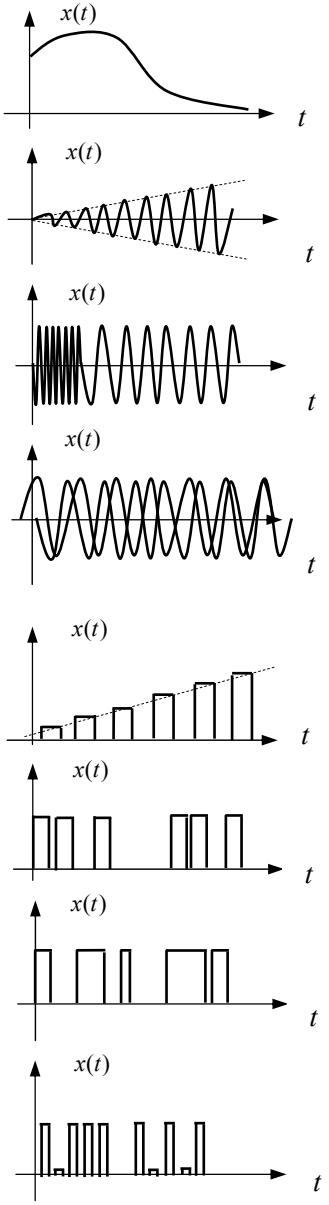
где A_0 – постоянная составляющая амплитуд; δA – наибольшее изменение амплитуды при модуляции; $f(t)$ – нормированная функция (изменяющаяся в пределах от –1 до +1); $m_A = \dfrac{\delta A}{A_0}$ – ***глубина амплитудной модуляции***. Так как модулируемый параметр сигнала является непосредственным переносчиком информации, то функция $f(t)$ выражает закон изменения во времени передаваемого полезного сигнала.

Амплитудно-модулированный гармонический сигнал как функция времени в общем случае имеет вид

$$x(t) = A_0[1 + m_A f(t)]\cos(\omega_0 t + \Theta_0).$$

Рассмотрим частный случай, когда функция $f(t)$ изменяется по закону $f(t) = \cos\Omega t$, $\Omega << \omega_0$ (рис. 1.7). Тогда $x(t)$ примет вид

$$x(t) = A_0[1 + m_A \cos\Omega t]\cos(\omega_0 t + \Theta_0) = A_0\cos(\omega_0 t + \Theta_0) +$$
$$+ \frac{A_0 m_A}{2}\cos[(\omega_0 - \Omega)t + \Theta_0] + \frac{A_0 m_A}{2}\cos[(\omega_0 + \Omega)t + \Theta_0]\ . \qquad (1.12)$$



Носитель первого типа:

ПМ – прямая модуляция

Носитель второго типа:

АМ – амплитудная модуляция;

ЧМ – частотная модуляция;

ФМ – фазовая модуляция;

Носитель третьего типа:

АИМ – амплитудно-импульсная модуляция;

ЧИМ – частотно-импульсная модуляция ($\tau = \text{const}$);

ВИМ – времяимпульсная модуляция ($T = \text{const}$);

КИМ – кодоимпульсная модуляция

Рис. 1.6. Основные виды модуляции

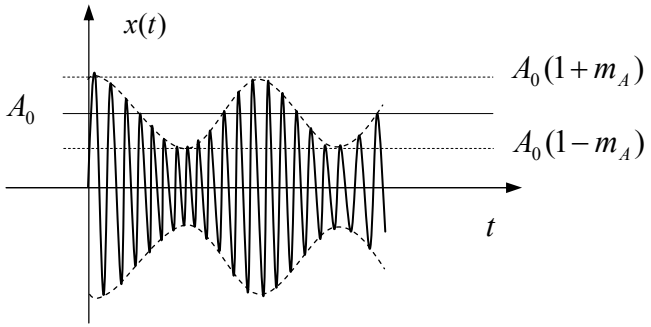


Рис. 1.7. Амплитудно-модулированный гармонический сигнал

Как видно из выражения (1.12), спектр сигнала, изображенного на рис. 1.8, состоит из трех гармонических составляющих: несущей с частотой ω_0 и двух боковых – нижней с частотой $\omega_0 - \Omega$ и верхней с частотой $\omega_0 + \Omega$. Ширина спектра сигнала $\Delta\omega = 2\Omega$.

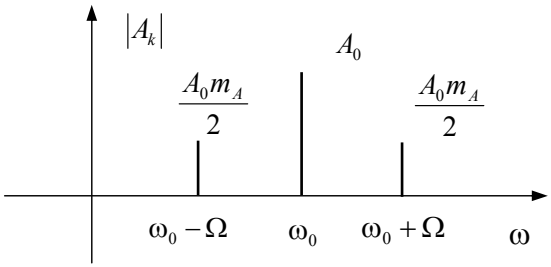


Рис. 1.8. Спектр амплитудно-модулированного гармонического сигнала

Если модуляция гармонического сигнала производится по более сложному закону, причем спектр огибающей амплитуды находится в диапазоне частот от $\Omega_{\min}$ до $\Omega_{\max}$, то можно показать, что в спектре амплитудно-модулированного сигнала вместо двух боковых частот будут две боковые полосы частот (рис. 1.9). Причем нижняя боковая полоса частот будет находиться в пределах $\omega_0 - \Omega_{\max} \div \omega_0 - \Omega_{\min}$, а верхняя боковая полоса частот – в пределах $\omega_0 + \Omega_{\min} \div \omega_0 + \Omega_{\max}$.

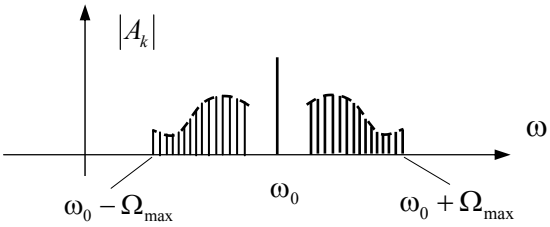


Рис. 1.9. Спектр сигнала, модулируемого по более сложному закону

Для неискаженной передачи такого сигнала канал связи должен обладать полосой пропускания частот равной $2\Omega_{max}$, т.е. полоса пропускания канала связи должна быть вдвое больше наивысшей частоты спектра модулирующего сигнала.

Иногда с помощью фильтров производится подавление несущей частоты и частот одной из боковых полос, т.е. передача производится в полосе частот $\Omega_{max} - \Omega_{min}$.

1.5. ЧАСТОТНО-МОДУЛИРОВАННЫЙ ГАРМОНИЧЕСКИЙ СИГНАЛ

При частотной модуляции частота сигнала изменяется в общем виде по закону

$$\omega(t) = \omega_0[1 + m_\omega f(t)],$$

где ω_0 – постоянная составляющая частоты; $m_\omega = \dfrac{\delta\omega}{\omega_0}$ – ***глубина частотной модуляции***; $\delta\omega$ – наибольшее изменение частоты при модуляции (***девиация частот***); $f(t)$ – закон модуляции. При изменении частоты всегда меняется фаза колебаний и наоборот. Этим определяется общий характер частотной (ЧМ) и фазовой (ФМ) модуляции. Иногда их объединяют под общим названием угловой модуляции. ЧМ осуществляется прямым воздействием датчика на генератор для изменения частоты его колебаний, хотя при этом меняется и фаза.

Здесь уместно напомнить некоторые соотношения для угловой

частоты колебаний, частоты в периодах и полной фазы колебания:

$$\omega(t) = \frac{d\varphi_{\text{п}}}{dt}; \quad f_0 = \frac{\omega_0}{2\pi} = \frac{1}{T}; \quad \varphi_{\text{п}}(t) = \int_0^t \omega(\xi) d\xi + \Theta_0 \ .$$

Из этих соотношений видно, что частоту можно оценивать по скорости изменения фазы, а полную фазу – по интегральному значению угловой частоты.

Итак, частотно-модулированный сигнал можно представить в виде

$$x(t) = A_0 \cos\varphi_{\text{п}}(t) = A_0 \cos[\varphi(t) + \Theta_0] = A_0 \cos[\int_0^t \omega(\xi) d\xi + \Theta_0] =$$

$$= A_0 \cos[\omega_0 t + \omega_0 m_\omega \int_0^t f(\xi) d\xi + \Theta_0] \ ,$$

где $\varphi(t)$ – текущая фаза сигнала; Θ_0 – начальная фаза.

В частном случае, когда модуляция частоты осуществляется по гармоническому закону $f(t) = \cos\Omega t, \ \Omega << \omega_0$, частотно-модулированный сигнал может быть записан в виде

$$x(t) = A_0 \cos[\omega_0 t + \omega_0 m_\omega \int_0^t \cos\Omega\xi d\xi + \Theta_0] =$$

$$= A_0 \cos[\omega_0 t + \beta \sin\Omega t + \Theta_0] \ ,$$

где $\beta = \dfrac{\delta\omega}{\Omega}$ – ***индекс модуляции***. Сигнал такого вида представляется следующим образом:

$$x(t) = A_0 \sum_{k=-\infty}^{\infty} J_k(\beta) \cos[(\omega_0 + k\Omega)t + \Theta_0] \ , \qquad (1.13)$$

где $J_k(\beta)$ – функция Бесселя первого рода k-го порядка от аргумента β.

Распределение амплитуд модулированного по частоте сигнала – дискретное, которое состоит из колебаний с несущей частотой ω_0 и бесконечного количества верхних и нижних боковых составляющих с частотами $\omega_0 \pm k\Omega$ (рис. 1.10).

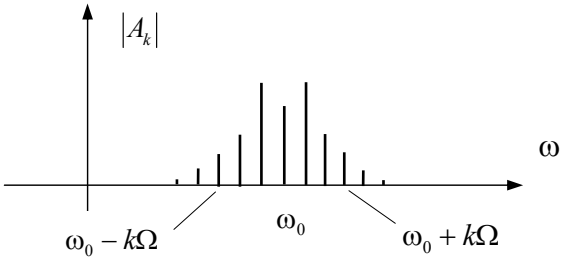


Рис. 1.10. Спектр частотно-модулированного гармонического сигнала

При очень малом индексе модуляции ($\beta << 1$) спектр частотно-модулированного сигнала практически не отличается от спектра амплитудно-модулированного сигнала, так как он состоит всего из трех гармонических составляющих: колебаний с несущей частотой ω_0 и двух боковых с частотами $\omega_0 - \Omega$ и $\omega_0 + \Omega$. В этом случае полоса пропускания канала связи должна быть равна 2Ω.

С ростом β увеличивается вес боковых составляющих и соответственно требуемая полоса пропускания. При больших индексах модуляции ($\beta >> 1$) ширина спектра сигнала практически равна удвоенной величине девиации частоты ($2\delta\omega$) и не зависит от частоты модулирующего сигнала Ω.

1.6. ФАЗОМОДУЛИРОВАННЫЙ ГАРМОНИЧЕСКИЙ СИГНАЛ

При этом виде модуляции фаза гармонического сигнала изменяется по закону модулирующего сигнала $f(t)$

$$\varphi(t) = \Theta_0[1 + m_\varphi f(t)] \ ,$$

где $m_\varphi = \dfrac{\delta\varphi}{\Theta_0}$ – ***глубина фазовой модуляции***; $\delta\varphi$ – наибольшее изменение фазы при модуляции.

Сигнал с фазовой модуляцией можно представить в следующем виде:

$$x(t) = A_0 \cos \varphi_{\text{п}}(t) = A_0 \cos[\omega_0 t + \varphi(t)] =$$
$$= A_0 \cos[\omega_0 t + \Theta_0 m_\varphi f(t) + \Theta_0].$$

Так как мгновенное значение частоты является производной фазы по времени, то при фазовой модуляции оно может быть представлено в виде

$$\omega(t) = \omega_0 + \Theta_0 m_\varphi \frac{df(t)}{dt}.$$

Таким образом, фазомодулированный сигнал эквивалентен сигналу с частотной модуляцией с модулирующей функцией $\frac{df(t)}{dt}$.

Для частного случая, когда модулирующий сигнал является гармоническим, полная фаза фазомодулированного сигнала определяется равенством

$$\varphi_{\text{п}}(t) = \omega_0 t + \Theta_0 m_\varphi \cos \Omega t + \Theta_0,$$

а фазомодулированное колебание – выражением

$$x(t) = A_0 \cos[\omega_0 t + \Theta_0 m_\varphi \cos \Omega t + \Theta_0].$$

Мгновенное значение частоты фазомодулированного колебания будет иметь вид

$$\omega(t) = \frac{d\varphi_{\text{п}}}{dt} = \omega_0 - \delta\varphi\Omega \sin \Omega t = \omega_0 - \omega_B \sin \Omega t,$$

где $\omega_B = \delta\varphi\Omega$ – девиация частоты фазомодулированного колебания.

Сравнение ЧМ и ФМ показывает, что при гармоническом модулирующем сигнале выражение, описывающее ЧМ колебание, отличается от такого же для ФМ колебания только фазой гармонической функции, определяющей изменение полной фазы высокочастотного колебания. В связи с этим по внешнему виду сигнала нельзя определить, как модулирован сигнал – по частоте или по фазе.

Однако при изменении частоты модуляции Ω проявляется различие между ЧМ и ФМ (рис. 1.11).

В общем случае спектр колебания, модулированного по фазе одной частотой Ω, выражается аналитически так же, как и спектр частотно-модулированного сигнала с такой же частотой Ω (1.13).

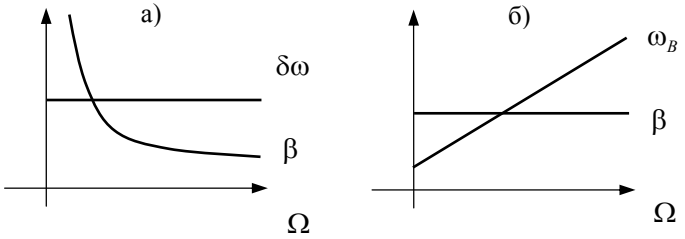


Рис. 1.11. Зависимость девиации и индекса модуляции от модулирующей частоты:

а – ЧМ ($\beta = \frac{\delta\omega}{\Omega}$); б – ФМ ($\omega_B = \delta\varphi\Omega,\ \beta = \frac{\omega_B}{\Omega} = \delta\varphi$)

Помимо различия в структуре колебания, ЧМ и ФМ различаются по способу осуществления. При ЧМ обычно применяется прямое воздействие на частоту колебаний генератора. При ФМ генератор дает стабильную частоту, а фаза колебания модулируется в одном из последующих элементов устройства.

1.7. ПЕРИОДИЧЕСКАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ПРЯМОУГОЛЬНЫХ ИМПУЛЬСОВ

Графическое представление периодической последовательности прямоугольных импульсов приведено на рис. 1.12. Аналитическое выражение такой последовательности как функции времени имеет вид

$$x(t) = \begin{cases} h, & -\frac{\tau}{2} + iT \le t \le \frac{\tau}{2} + iT; \\ 0, & \frac{\tau}{2} + iT < t < -\frac{\tau}{2} + (i+1)T, \quad i = 0, \pm 1, \pm 2, \ldots . \end{cases}$$

С другой стороны, функция $x(t)$ может быть представлена рядом Фурье

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} |A_k| \cos(k\omega_0 t + \Theta_k) =$$

$$= \frac{A_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega_0 t - b_k \sin k\omega_0 t).$$

Так как $x(t)$ – четная функция, то $b_k = 0$ и $\Theta_k = -\text{arctg}\frac{b_k}{a_k} = \begin{bmatrix} \pi \\ 0 \\ -\pi \end{bmatrix}$.

Следовательно, $x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} a_k \cos k\omega_0 t$.

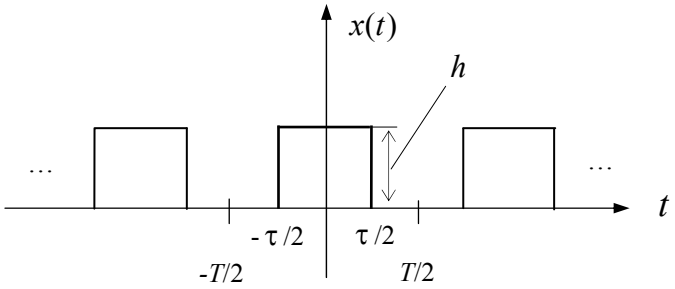


Рис. 1.12. Периодическая последовательность прямоугольных импульсов

Постоянная составляющая периодической последовательности прямоугольных импульсов принимает вид $\frac{A_0}{2} = \frac{1}{T}\int_{-\tau/2}^{\tau/2} h dt = \frac{\tau h}{T}$, а амплитуда k - й гармонической составляющей

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos(k\omega_0 t)dt = \frac{2h}{T}\frac{1}{k\omega_0}\int_{-\tau/2}^{\tau/2} d\sin k\omega_0 t = \\ = \frac{2h}{T}\frac{1}{k\omega_0}\sin k\omega_0 t\Big|_{-\tau/2}^{\tau/2} = 2\frac{\tau h}{T}\frac{\sin\frac{k\omega_0\tau}{2}}{\frac{k\omega_0\tau}{2}}. \qquad (1.14)$$

Таким образом, разложение периодической последовательности прямоугольных импульсов в ряд Фурье представляется в следующем виде:

$$x(t) = \frac{\tau h}{T}\left[1 + 2\sum_{k=1}^{\infty}\frac{\sin\frac{k\omega_0\tau}{2}}{\frac{k\omega_0\tau}{2}}\cos(k\omega_0 t)\right]. \qquad (1.15)$$

Как видно из (1.14), при очередном приращении частоты на величину $\frac{2\pi}{\tau}$ фаза гармоник изменяется на величину π.

Спектры амплитуд и фаз последовательности прямоугольных импульсов представлены на рис. 1.13.

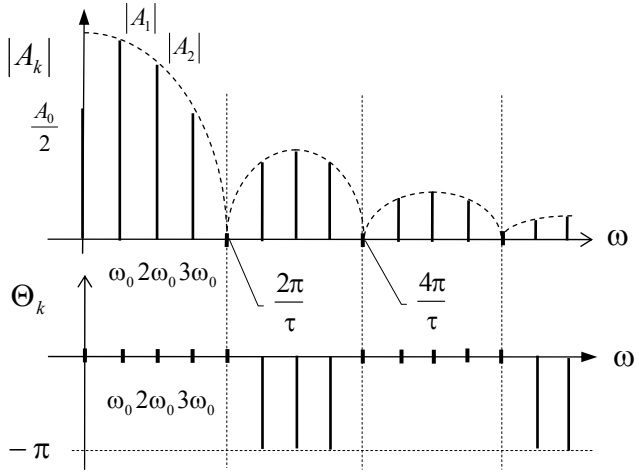


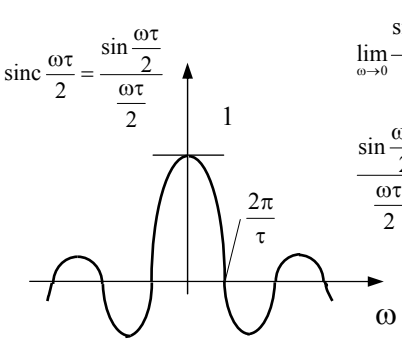
Рис. 1.13. Спектры амплитуд и фаз периодической последовательности прямоугольных импульсов

Причем огибающая спектра амплитуд определяется уравнением

$$A(\omega) = 2\frac{\tau h}{T}\left|\frac{\sin\frac{k\omega\tau}{2}}{\frac{k\omega\tau}{2}}\right|,$$

а форма огибающей – видом функции, приведенной на рис. 1.14.

Номер k-й гармонической составляющей, попадающей в первый нуль функции $\operatorname{sinc}\frac{\omega\tau}{2}$, определяется из соотношения $k\omega_0 = \frac{2\pi}{\tau}$, откуда получаем $k = \frac{T}{\tau}$ [15]. Например, при $T = 2\tau$ спектр амплитуд имеет вид, приведенный на рис. 1.15.



$$\lim_{\omega\to 0}\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}=1\ ,$$

$$\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}=0,\ \text{при}\ \omega=\frac{n\pi}{\tau},\ n-\text{четное}.$$

Рис. 1.14. Функция $\operatorname{sinc}\dfrac{\omega\tau}{2}$

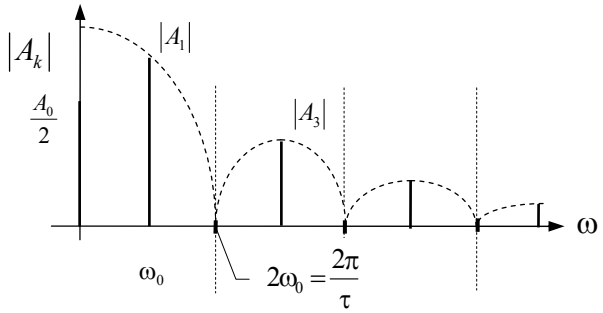


Рис. 1.15. Спектр амплитуд последовательности $T=2\tau$

1.8. АМПЛИТУДНО-МОДУЛИРОВАННАЯ ПЕРИОДИЧЕСКАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ИМПУЛЬСОВ

Пусть амплитуда прямоугольных импульсов модулируется по гармоническому закону (рис. 1.16) $h(t)=h_0[1+m_h\cos\Omega t]$.

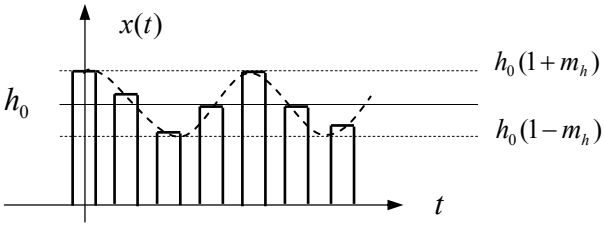


Рис. 1.16. Амплитудно-модулированная последовательность импульсов

Подставив $h(t)$ в выражение (1.15) для немодулированной последовательности прямоугольных импульсов, получим выражение для амплитудно-модулированной последовательности

$$x(t) = \frac{\tau h_0}{T}\left[1 + m_h \cos \Omega t\right]\left[1 + 2\sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_0 \tau}{2}}{\frac{k\omega_0 \tau}{2}} \cos(k\omega_0 t)\right].$$

После тригонометрических преобразований будем иметь

$$x(t) = \frac{\tau h_0}{T}\left[1 + 2\sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_0 \tau}{2}}{\frac{k\omega_0 \tau}{2}} \cos(k\omega_0 t)\right] + m_h \frac{\tau h_0}{T} \cos \Omega t +$$

$$+ m_h \frac{\tau h_0}{T} \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_0 \tau}{2}}{\frac{k\omega_0 \tau}{2}} \cos\left[(k\omega_0 + \Omega)t\right] +$$

$$+ m_h \frac{\tau h_0}{T} \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_0 \tau}{2}}{\frac{k\omega_0 \tau}{2}} \cos\left[(k\omega_0 - \Omega)t\right].$$

Сравнивая полученное выражение с разложением для периодической последовательности немодулированных импульсов, видим, что в результате амплитудной модуляции импульсов в спектре, кроме постоянной составляющей и составляющих с частотами, кратными основной частоте ω_0, появились составляющая с частотой модуляции Ω, а также по две боковые составляющие (нижняя $k\omega_0 - \Omega$ и верхняя $k\omega_0 + \Omega$) около каждой k - й гармоники основной частоты (рис. 1.17).

Появление боковых частот $k\omega_0 \pm \Omega$ практически не сказывается на ширине спектра сигнала, так как частота модуляции $\Omega << \omega_0$. В связи с этим практическая ширина спектра амплитудно-модулированной последовательности импульсов такая же, как и соответствующей последовательности немодулированных импуль-

сов, т.е. определяется длительностью импульсов.

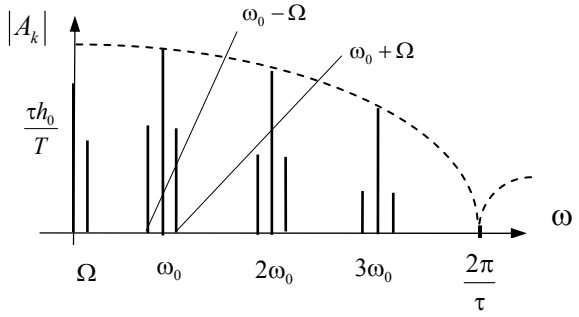


Рис. 1.17. Спектр амплитуд модулированной последовательности

ЗАДАЧИ

1.1. Доказать тождественность комплексной и тригонометрической форм обобщенного ряда Фурье

$$x(t) = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{jk\omega_0 t},$$

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} |A_k| \cos(k\omega_0 t + \Theta_k).$$

1.2. Найти спектр периодической последовательности прямоугольных импульсов (рис. 1.18)

$$x(t) = \begin{cases} +h, \; iT \le t < \dfrac{T}{2} + iT; \\ -h, -\dfrac{T}{2} + iT \le t < iT, \quad i = 0, \pm 1, \pm 2, \ldots, \end{cases}$$

называемой меандром.

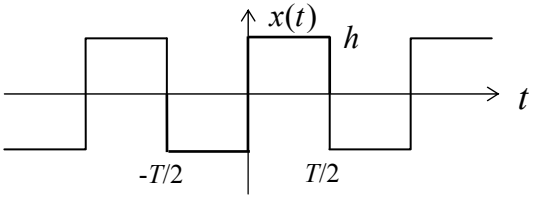


Рис. 1.18. Меандр

1.3. Найти спектр периодической последовательности прямоугольных импульсов для $T = 2\tau$ (рис. 1.19)

$$x(t) = \begin{cases} h, & -\frac{\tau}{2} + iT \le t < \frac{\tau}{2} + iT; \\ 0, & \frac{\tau}{2} + iT \le t < -\frac{\tau}{2} + (i+1)T, \quad i = 0, \pm 1, \pm 2, \ldots . \end{cases}$$

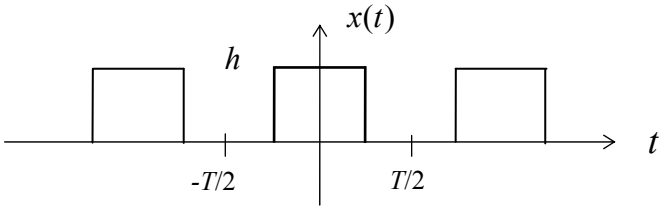


Рис. 1.19. Периодическая последовательность прямоугольных импульсов

1.4. Построить обобщенный ряд Фурье для последовательности униполярных треугольных импульсов (рис. 1.20)

$$x(t) = \frac{2h}{T}\left|t - iT\right|, \; -\frac{T}{2} + iT \le t < \frac{T}{2} + iT, i = 0, \pm 1, \pm 2, \ldots .$$

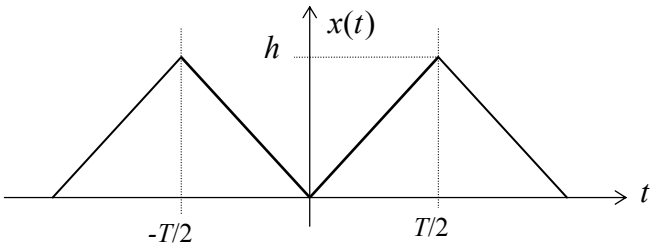


Рис. 1.20. Периодическая последовательность униполярных треугольных импульсов

1.5. Построить обобщенный ряд Фурье периодического колебания пилообразной формы (рис. 1.21)

$$x(t) = \frac{2h}{T}t - i2h, \; -\frac{T}{2} + iT \le t < \frac{T}{2} + iT, i = 0, \pm 1, \pm 2, \ldots .$$

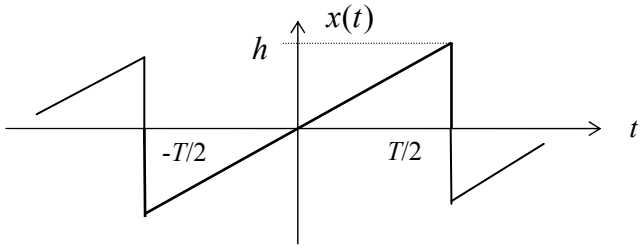


Рис. 1.21. Периодическое колебание пилообразной формы

1.6. Найти спектральные характеристики последовательности полукосинусоидальных импульсов (рис. 1.22)

$$x(t)=\begin{cases} h\cos\omega_0 t, & -\dfrac{\tau}{2}+iT\le t<\dfrac{\tau}{2}+iT; \\ 0, & \dfrac{\tau}{2}+iT\le t<-\dfrac{3\tau}{2}+iT, \quad i=0,\pm1,\pm2,\ldots . \end{cases}$$

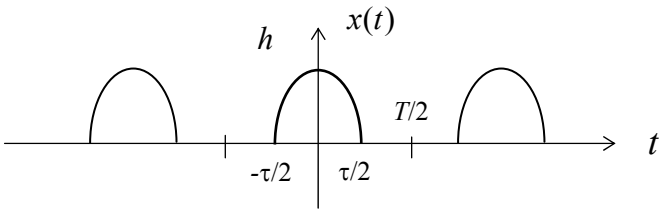


Рис. 1.22. Периодическая последовательность полукосинусоидальных импульсов

1.7. Построить обобщенный ряд Фурье периодической последовательности «выпрямленных» косинусоидальных импульсов (рис.1. 23).

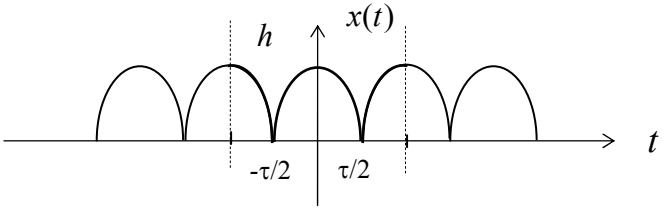


Рис. 1.23. Периодическая последовательность «выпрямленных» косинусоидальных импульсов

1.8. Определить, какая часть средней мощности сосредоточена в пределах практической ширины спектра периодической последовательности прямоугольных импульсов (см. рис. 1.19), учитывающей первые пять гармоник.

1.9. Определить практическую ширину спектра периодической последовательности прямоугольных импульсов при длительности импульсов $\tau = \frac{T}{2}$, если требуется учесть все гармонические составляющие сигнала, амплитуды которых не менее 0,1 от амплитуды первой гармоники.

1.10. Провести графическое суммирование составляющих спектра меандра (см. рис. 1.18) до трех гармоник включительно.

1.11. Найти спектр детерминированного периодического сигнала (рис. 1.24)

$$x(t) = \begin{cases} \frac{2h}{T}t - 2(i-1)h\ , & -\frac{T}{2} + iT \le t < 0 + iT; \\ \frac{2h}{T}t - 2ih\ , & 0 + iT \le t < \frac{T}{2} + iT, \quad i = 0, \pm 1, \pm 2, \ldots . \end{cases}$$

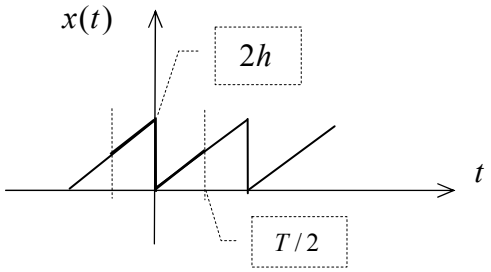


Рис. 1.24. Периодическое колебание

2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЕТЕРМИНИРОВАННЫХ НЕПЕРИОДИЧЕСКИХ СИГНАЛОВ

Непериодическим сигналом называется любой детерминированный сигнал, для которого не выполняется условие $x(t) = x(t + kT),\ k = \pm 1, \pm 2, \ldots$. Как правило, непериодический сигнал ограничен по времени. Примерами таких сигналов могут служить уже упоминавшиеся импульсы, пачки импульсов, "обрывки" гармонических колебаний и т. д. Непериодические сигналы представляют основной интерес, так как именно они преимущественно используются на практике.

Основой математической модели непериодического, как и периодического, сигнала является его спектральная характеристика. Однако структура спектра непериодического сигнала имеет некоторые особенности, которые будут подробно рассмотрены ниже.

2.1. ГАРМОНИЧЕСКИЙ АНАЛИЗ НЕПЕРИОДИЧЕСКИХ КОЛЕБАНИЙ

Гармонический анализ периодических колебаний можно распространить на непериодические колебания. Пусть такое колебание $x(t)$ задано в виде некоторой функции, отличной от нуля в промежутке (t_1, t_2) (рис. 2.1). Выделив произвольный отрезок времени T, включающий в себя промежуток (t_1, t_2), можно представить заданное колебание в виде ряда Фурье, предположив его периодичность,

$$x(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} A_k e^{jk\omega_0 t}, \qquad (2.1)$$

где $\omega_0 = \frac{2\pi}{T}$, а коэффициенты

$$A_k = \frac{2}{T} \int_{t_1}^{t_2} x(t) e^{-jk\omega_0 t} dt \quad .$$

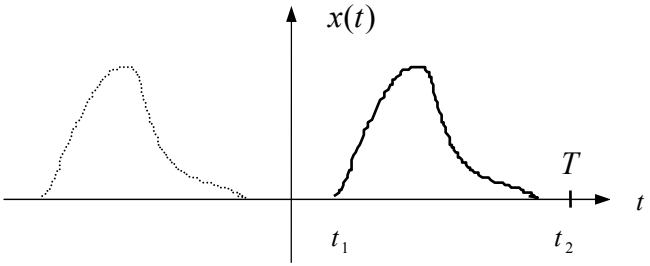


Рис. 2.1. Одиночный непериодический сигнал

Подставив A_k в (2.1), получим

$$x(t) = \sum_{k=-\infty}^{\infty} \left[\int_{t_1}^{t_2} x(\xi) e^{-jk\omega_0 \xi} d\xi \right] e^{jk\omega_0 t} \frac{\omega_0}{2\pi}. \qquad (2.2)$$

Вне отрезка $[0,T]$ ряд (2.1) определяет функцию $x(t) = x(t \pm kT)$, где k – целое число, т.е. периодическую функцию, полученную повторением $x(t)$ вправо и влево с периодом T. Для того чтобы вне отрезка $[0,T]$ функция равнялась нулю, величина T должна быть бесконечно большой. Но чем больше отрезок T, выбранный в качестве периода, тем меньше коэффициенты A_k. Устремляя T к бесконечности, в пределе получаем бесконечно малые амплитуды гармонических составляющих, сумма которых изображает исходную непериодическую функцию $x(t)$, заданную в интервале $t_1 \le t \le t_2$. Число гармонических составляющих, входящих в ряд Фурье, будет при этом бесконечно большим, так как при $T \to \infty$ основная частота $\omega_0 = \dfrac{2\pi}{T} \to 0$.

Иными словами, расстояние между спектральными линиями, равное основной частоте ω_0, становится бесконечно малым, а спектр сплошным.

Поэтому можно в выражении (2.2) заменить ω_0 на $d\omega$, $k\omega_0$ на текущую частоту ω, а операцию суммирования – операцией интегрирования

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} \left[\int_{t_1}^{t_2} x(\xi) e^{-j\omega\xi} d\xi \right] d\omega . \qquad (2.3)$$

Внутренний интеграл, являющийся функцией ω,

$$S(j\omega) = \int_{t_1}^{t_2} x(t) e^{-j\omega t} dt$$

называется ***спектральной плотностью*** или спектральной характеристикой функции $x(t)$.

В общем случае, когда пределы t_1 и t_2 в (2.3) не уточнены, спектральная плотность записывается в форме

$$S(j\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt . \qquad (2.4)$$

После подстановки (2.4) в (2.3) получаем

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega . \qquad (2.5)$$

Выражения (2.4) и (2.5) называются соответственно ***прямым и обратным преобразованиями Фурье***.

Спектральная плотность $S(j\omega)$ обладает всеми основными свойствами коэффициентов A_k комплексного ряда Фурье. Следовательно, по аналогии можно написать

$$S(j\omega) = \int_{-\infty}^{\infty} x(t) \cos(\omega t) dt - j \int_{-\infty}^{\infty} x(t) \sin(\omega t) dt =$$

$$= A(\omega) - jB(\omega) = \left| S(j\omega) \right| e^{j\Theta(\omega)} = S(\omega) e^{j\Theta(\omega)} \quad .$$

Модуль и аргумент (фазовая характеристика) ***спектральной плотности*** определяются выражениями:

$$S(\omega) = \sqrt{[A(\omega)]^2 + [B(\omega)]^2} ,$$

$$\Theta(\omega) = -\operatorname{arctg} \frac{B(\omega)}{A(\omega)} \quad .$$

Эти понятия не следует путать с амплитудно-частотной (АЧХ) и

фазочастотной характеристиками (ФЧХ), которые относятся к электрическим цепям, а не к сигналам.

Как и в случае ряда Фурье, $S(\omega)$ является четной $[S(\omega)=S(-\omega)]$, а $\Theta(\omega)$ – нечетной $[\Theta(\omega)=-\Theta(-\omega)]$ функциями частоты ω.

Интегральное преобразование (2.5) можно привести к тригонометрической форме

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)\cos(\omega t+\Theta)d\omega+j\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)\sin(\omega t+\Theta)d\omega.$$

Из четности модуля спектральной плотности и нечетности фазы следует, что подынтегральная функция в первом интеграле является четной, а во втором – нечетной относительно ω. Следовательно, второй интеграл равен нулю, и окончательно

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)\cos(\omega t+\Theta)d\omega=\frac{1}{\pi}\int_{0}^{\infty}S(\omega)\cos(\omega t+\Theta)d\omega.$$

Переход от комплексной формы к тригонометрической обычно целесообразен в конце анализа. Все промежуточные выкладки при применении интеграла Фурье удобнее и проще производить в комплексной форме.

Из сопоставления комплексных выражений

$$x(t)=\frac{1}{2}\sum_{k=-\infty}^{\infty}A_k e^{jk\omega_0 t},$$

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

видно, что величина $\frac{1}{\pi}S(j\omega)d\omega=2S(\omega)df$ имеет смысл коэффициента A_k (бесконечно малого) комплексного ряда Фурье при частоте $\omega=2\pi f$.

Соответственно, из сопоставления тригонометрических выражений

$$x(t)=\frac{A_0}{2}+\sum_{k=1}^{\infty}|A_k|\cos(k\omega_0 t+\Theta_k),$$

$$x(t) = \frac{1}{\pi}\int_0^{\infty} S(\omega)\cos(\omega t + \Theta)d\omega$$

видно, что величина $\frac{1}{\pi}S(\omega)d\omega = 2S(\omega)df$ имеет смысл модуля амплитуды $|A_k|$ (бесконечно малой) гармонической составляющей частоты $\omega = 2\pi f$.

Из этих сопоставлений становится ясным смысл термина "спектральная плотность": $2S(\omega)$ есть амплитуда колебания, приходящаяся на 1 Гц в бесконечно узкой полосе частот, включающей в себя рассматриваемую частоту $\omega = 2\pi f$.

2.2. СОПОСТАВЛЕНИЕ СПЕКТРОВ ПЕРИОДИЧЕСКИХ И СООТВЕТСТВУЮЩИХ НЕПЕРИОДИЧЕСКИХ СИГНАЛОВ

Рассмотрим периодический сигнал $x_k(t)$ с периодом T и непериодический сигнал $x(t)$, который в интервале $[t_1, t_1 + T]$ совпадает с периодическим сигналом, а вне этого интервала равен нулю (рис. 2.2).

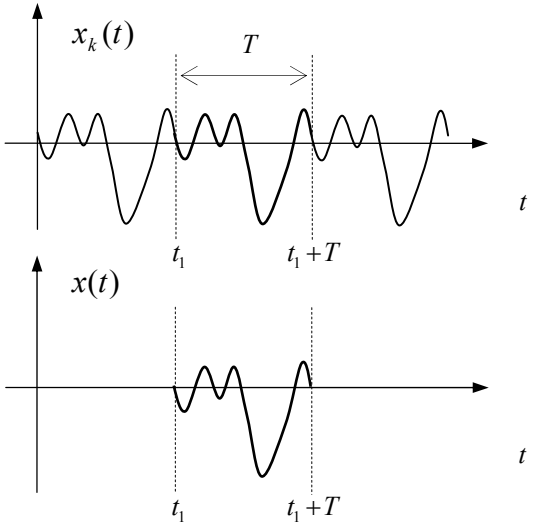


Рис. 2.2. Периодический и соответствующий непериодический сигналы

Комплексная амплитуда гармоник периодического и спектральная плотность непериодического сигналов соответственно равны

$$A_k = \frac{2}{T}\int_{t_1}^{t_1+T} x_k(t)e^{-jk\omega_0 t}dt \ ,$$

$$S(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt = \int_{t_1}^{t_1+T} x_k(t)e^{-j\omega t}dt \ .$$

Из сравнения этих выражений можно заключить, что для каждой частоты спектра периодического сигнала справедливы равенства

$$A_k = \frac{2}{T}S(jk\omega_0),\ |A_k| = \frac{2}{T}S(k\omega_0),\ \frac{A_0}{2} = \frac{1}{T}S(0) \ .$$

Отсюда следует, что спектр непериодического сигнала, заданного в интервале $[t_1, t_1+T]$, с точностью до постоянного множителя совпадает с огибающей спектра периодического сигнала, в интервале $[t_1, t_1+T]$ описываемого функцией такого же вида, что и непериодический сигнал (рис. 2.3).

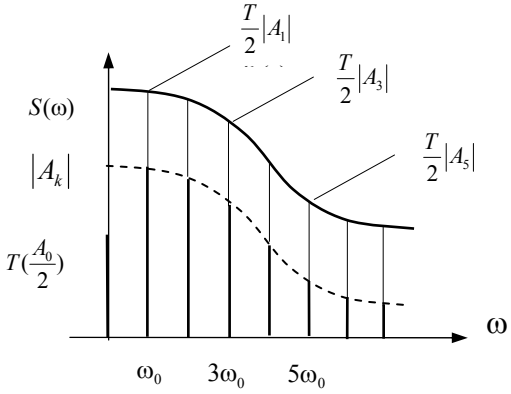


Рис. 2.3. Сопоставление спектров сигналов

Таким образом, спектр непериодического сигнала может быть определен по известному спектру соответствующего периодического сигнала и, наоборот, спектр периодического сигнала может

быть найден по известному спектру соответствующего непериодического сигнала.

2.3. СВОЙСТВА ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Между колебанием $x(t)$ и спектром $S(j\omega)$ существует однозначное соответствие. Для практических приложений важно установить связь между преобразованием колебания и соответствующим этому преобразованию изменением спектра.

2.3.1. Сдвиг колебания во времени

Пусть колебание $x_1(t)$ произвольной формы существует на интервале времени от t_1 до t_2 и обладает спектральной плотностью $S_1(j\omega)$. При задержке этого колебания на величину t_0 (при сохранении его формы) получим новую функцию времени $x_2(t) = x_1(t - t_0)$, существующую на интервале от $t_1 + t_0$ до $t_2 + t_0$ (рис. 2.4).

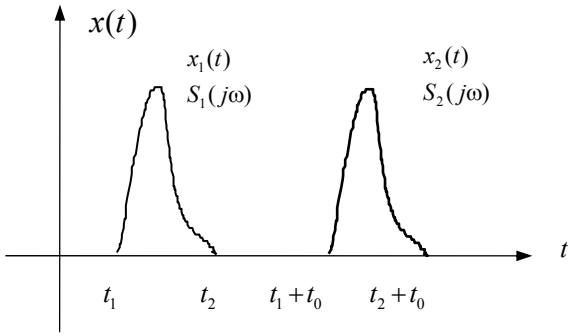


Рис. 2.4. Сдвиг сигнала

Спектральная плотность колебания $x_2(t)$ имеет вид

$$S_2(j\omega) = \int_{-\infty}^{\infty} x_2(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} x_1(t - t_0) e^{-j\omega t} dt .$$

Вводя новую переменную $\tau = t - t_0$, получаем

$$S_2(j\omega) = e^{-j\omega t_0} \int_{-\infty}^{\infty} x_1(\tau) e^{-j\omega\tau} d\tau = e^{-j\omega t_0} S_1(j\omega) = S_1(\omega) e^{j(\Theta_1 - \omega t_0)}.$$

Сдвиг во времени функции $x(t)$ на величину $\pm t_0$ приводит к изменению фазовой характеристики спектра $S(j\omega)$ на величину $\pm \omega t_0$. Модуль $S(\omega)$ от положения колебания на оси времени не зависит.

2.3.2. Инверсия сигнала

Пусть колебание $x_1(t)$ произвольной формы существует на интервале времени от t_1 до t_2 и обладает спектральной плотностью $S_1(j\omega)$. При инверсии этого колебания (при сохранении его формы) получим новую функцию времени $x_2(t) = -x_1(t)$, существующую на интервале от t_1 до t_2 (рис. 2.5).

Спектральная плотность колебания $x_2(t)$ имеет вид

$$S_2(j\omega) = \int_{-\infty}^{\infty} x_2(t) e^{-j\omega t} dt = -\int_{-\infty}^{\infty} x_1(t) e^{-j\omega t} dt = -S_1(j\omega).$$

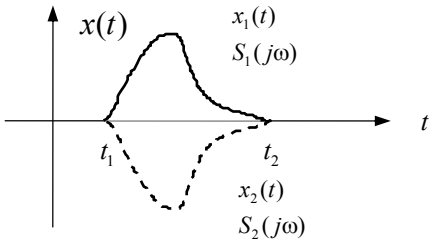


Рис. 2.5. Инверсия сигнала

Учитывая, что $-1 = e^{-j\pi} = \cos\pi - j\sin\pi$, спектральную плотность $S_2(j\omega)$ можно представить в виде $S_2(j\omega) = S_1(\omega) e^{j(\Theta_1 - \pi)}$.

Итак, инверсия сигнала $x_1(t)$ приводит к изменению фазовой характеристики спектра $S_1(j\omega)$ на величину $-\pi$.

2.3.3. Изменение масштаба времени

Пусть колебание $x_1(t)$ подверглось сжатию во времени. Новое сжатое колебание $x_2(t)$ связано с исходным колебанием соотношением $x_2(t) = x_1(nt),\ n > 1$ (рис. 2.6).

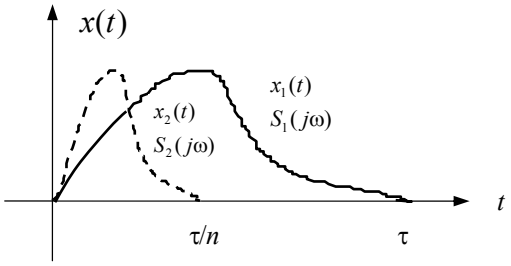


Рис. 2.6. Сжатие сигнала

Спектральная плотность колебания $x_2(t)$ имеет вид $S_2(j\omega) = \int\limits_{-\infty}^{\infty} x_2(t)e^{-j\omega t}dt = \int\limits_{-\infty}^{\infty} x_1(nt)e^{-j\omega t}dt$. Вводя новую переменную $\tau = nt$, получаем $S_2(j\omega) = \frac{1}{n}\int\limits_{-\infty}^{\infty} x_1(\tau)e^{-j\frac{\omega}{n}\tau}d\tau$. Но полученный интеграл есть не что иное, как спектральная плотность исходного колебания $x_1(t)$ при частоте $\frac{\omega}{n}$, т.е. $S_1(j\frac{\omega}{n})$. Таким образом, $S_2(j\omega) = \frac{1}{n}S_1(j\frac{\omega}{n})$.

Итак, при сжатии колебания в n раз на временной оси во столько же раз расширяется его спектр на оси частот. Модуль спектральной плотности при этом уменьшается в n раз. Очевидно, что при растягивании колебания ($n < 1$) имеет место сужение спектра и увеличение модуля.

2.3.4. Сдвиг спектра колебания по частоте

Применим преобразование Фурье к произведению следующего

вида $x_2(t) = x_1(t)\cos\omega_0 t$:

$$S_2(j\omega) = \int_{-\infty}^{\infty} x_1(t)\cos\omega_0 t e^{-j\omega t}dt = \int_{-\infty}^{\infty} x_1(t)\left[\frac{1}{2}e^{j\omega_0 t} + \frac{1}{2}e^{-j\omega_0 t}\right]e^{-j\omega t}dt =$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x_1(t)e^{-j(\omega-\omega_0)t}dt + \frac{1}{2}\int_{-\infty}^{\infty} x_1(t)e^{-j(\omega+\omega_0)t}dt \ .$$

Первый интеграл в правой части есть не что иное, как спектральная плотность функции $x_1(t)$ при частоте $\omega - \omega_0$, а второй интеграл – при $\omega + \omega_0$.

Таким образом, $S_2(j\omega) = \frac{1}{2}\{S_1[j(\omega-\omega_0)] + S_1[j(\omega+\omega_0)]\}$. Расщепление спектра $S_1(j\omega)$ на две части, смещенные соответственно на $+\omega_0$ и $-\omega_0$, эквивалентно умножению функции $x_1(t)$ на гармоническое колебание $\cos\omega_0 t$.

2.3.5. Сложение колебаний

Так как преобразование Фурье, определяющее спектральную плотность заданной функции времени, является линейным преобразованием, то очевидно, что при сложении колебаний $x_1(t), x_2(t), \ldots$, обладающих спектрами $S_1(j\omega), S_2(j\omega), \ldots$, суммарному колебанию $x(t) = x_1(t) + x_2(t) + \ldots$ соответствует спектр $S(j\omega) = S_1(j\omega) + S_2(j\omega) + \ldots$.

2.3.6. Дифференцирование и интегрирование колебания

Дифференцирование (интегрирование) колебания $x_1(t)$ можно рассматривать как почленное дифференцирование (интегрирование) всех гармонических составляющих, входящих в его спектр. Гармоническую составляющую колебания $x_1(t)$ при частоте ω можно представить в следующем виде: $\left[\frac{1}{2\pi}S_1(j\omega)d\omega\right]e^{j\omega t}$. Заключенная в квадратные скобки величина представляет собой ампли-

туду колебания в полосе $d\omega$. Следовательно, колебанию $x_2(t)=\dfrac{dx_1(t)}{dt}$ соответствует спектральная плотность $S_2(j\omega)=j\omega S_1(j\omega)$, а $x_2(t)=\int\limits_L x_1(t)dt$ – $S_2(j\omega)=\dfrac{1}{j\omega}S_1(j\omega)$.

2.3.7. Произведение колебаний

Пусть рассматриваемое колебание $x(t)$ является произведением двух функций времени $x(t)=f(t)g(t)$.

Определим спектр колебания $x(t)$:

$$S(j\omega)=\int\limits_{-\infty}^{\infty}x(t)e^{-j\omega t}dt=\int\limits_{-\infty}^{\infty}f(t)g(t)e^{-j\omega t}dt\,.$$

Каждую из функций $f(t)$ и $g(t)$ можно представить в виде интеграла Фурье

$$f(t)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}F(j\omega)e^{j\omega t}d\omega,\;\; g(t)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}G(j\omega)e^{j\omega t}d\omega\,.$$

Подставляя второй из них в $S(j\omega)$, получаем

$$S(j\omega)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}f(t)\left[\int\limits_{-\infty}^{\infty}G(j\xi)e^{j\xi t}d\xi\right]e^{-j\omega t}dt=$$

$$=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}G(j\xi)\left[\int\limits_{-\infty}^{\infty}f(t)e^{-j(\omega-\xi)t}dt\right]d\xi.$$

Заключенный в квадратные скобки интеграл представляет собой спектральную плотность функции $f(t)$ при частоте $\omega-\xi$, т.е. $F\left[j(\omega-\xi)\right]$. Следовательно, $S(j\omega)=\dfrac{1}{2\pi}\int\limits_{-\infty}^{\infty}G(j\xi)F\left[j(\omega-\xi)\right]d\xi.$

Итак, спектр произведения двух функций $f(t)$ и $g(t)$ равен (с коэффициентом $\dfrac{1}{2\pi}$) свертке их спектров

$$S(j\omega)=\frac{1}{2\pi}\left[G(j\omega)*F(j\omega)\right].$$

2.3.8. Взаимная заменяемость ω и t в преобразовании Фурье

Для четной функции $x(t)$ можно произвольно выбирать знак перед t в обратном преобразовании Фурье. Выберем знак минус и запишем формулу в виде $x(t)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}S(j\omega)e^{-j\omega t}d\omega$. Произведем теперь в последнем интеграле замену переменной интегрирования ω на t и параметра t на ω. Тогда $x(\omega)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}S(jt)e^{-j\omega t}dt$. Но этот интеграл можно рассматривать как спектральную плотность новой функции $S(jt)$, полученной путем замены ω на t в выражении спектральной плотности колебания $x(t)$.

Обозначим эту новую спектральную плотность через $S'(j\omega)$. Тогда $S'(j\omega)=2\pi x(\omega)$.

Итак, переменные ω и t в преобразовании Фурье взаимно заменимы: если четному колебанию $x(t)$ соответствует спектр $S(j\omega)$, то колебанию $S(jt)$ соответствует спектр $2\pi x(\omega)$.

2.4. СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ ОДИНОЧНОГО ПРЯМОУГОЛЬНОГО ИМПУЛЬСА

Простейшее колебание (рис. 2.7), определяемое выражением

$$x(t)=\begin{cases}h, & -\frac{\tau}{2}\le t\le\frac{\tau}{2};\\ 0, & |t|>\frac{\tau}{2},\end{cases}$$

получило широкое распространение как в технике, так и в теории сигналов.

Применяя преобразование (2.4), находим

$$S(j\omega)=\int_{-\infty}^{\infty}x(t)e^{-j\omega t}dt=h\int_{-\tau/2}^{\tau/2}e^{-j\omega t}dt=\frac{h}{-j\omega}\left(e^{-\frac{j\omega\tau}{2}}-e^{\frac{j\omega\tau}{2}}\right)=$$

$$=\frac{2h}{\omega}\sin\frac{\omega\tau}{2}=h\tau\left[\operatorname{sinc}\frac{\omega\tau}{2}\right]=h\tau\left|\operatorname{sinc}\frac{\omega\tau}{2}\right|e^{j\Theta(\omega)}.$$

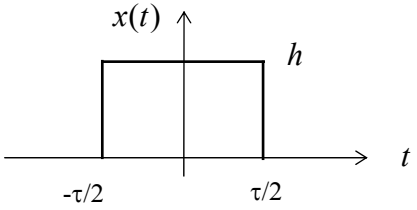


Рис. 2.7. Одиночный прямоугольный импульс

Заметим, что произведение $h\tau$, равное площади импульса, определяет значение спектральной плотности импульса при $\omega=0$, т.е. $S(0)=h\tau$. Этот вывод можно распространить на импульс произвольной формы. Таким образом $S(j\omega)=S(0)\operatorname{sinc}\left(\frac{\omega\tau}{2}\right)$.

При удлинении (растягивании) импульса расстояние между нулями функции $S(j\omega)$ сокращается, что равносильно сужению спектра. Значение $S(0)$ при этом возрастает. При укорочении (сжатии) импульса, наоборот, расстояние между нулями функции $S(j\omega)$ увеличивается (расширение спектра), а значение $S(0)$ уменьшается. В пределе при $\tau\to 0$ точки $\omega=\pm\frac{2\pi}{\tau}$, соответствующие двум первым нулям функции $S(j\omega)$, удаляются в бесконечность, и спектральная плотность, бесконечно малая по величине, становится равномерной в полосе частот от $-\infty$ до $+\infty$.

Модуль $S(\omega)$ и аргумент $\Theta(\omega)$ спектральной плотности прямоугольного импульса приведены на рис. 2.8. Отметим, что $S(\omega)=\frac{T}{2}A(\omega)$, где $A(\omega)$ – огибающая спектра периодической последовательности прямоугольных импульсов.

При отсчете времени не от середины импульса, а от фронта (рис. 2.9,а), фазовая характеристика спектра импульса должна быть дополнена слагаемым $-\frac{\omega\tau}{2}$, т.е. $S(j\omega) = h\tau\left|\text{sinc}\frac{\omega\tau}{2}\right| e^{j\left[\Theta(\omega)-\frac{\omega\tau}{2}\right]}$. Результирующая фазовая характеристика изображена на рис. 2.9,б.

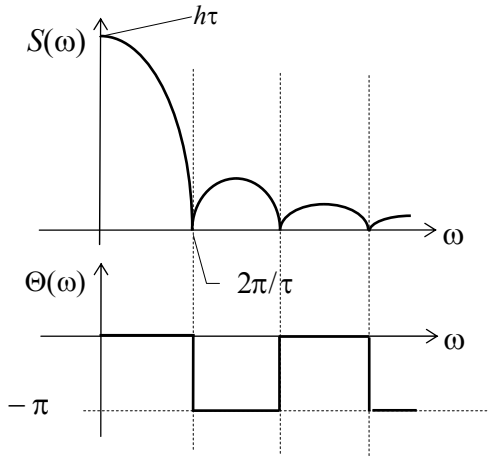


Рис. 2.8. Модуль и аргумент $S(j\omega)$ одиночного прямоугольного импульса

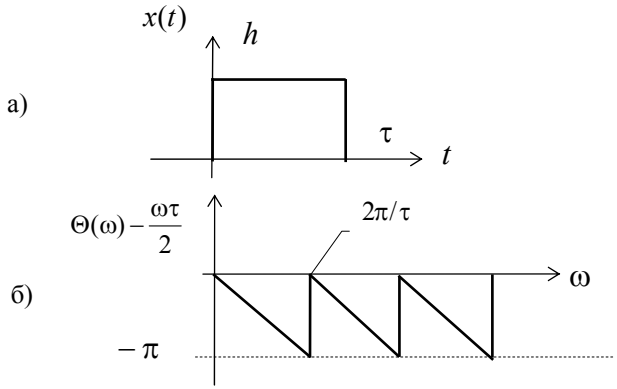


Рис. 2.9. Аргумент $S(j\omega)$ сдвинутого прямоугольного импульса:
а – сдвинутый прямоугольный импульс; б – аргумент

2.5. СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ ПАЧКИ ПРЯМОУГОЛЬНЫХ ИМПУЛЬСОВ

Рассмотрим группу одинаковых и равноотстоящих импульсов (рис. 2.10).

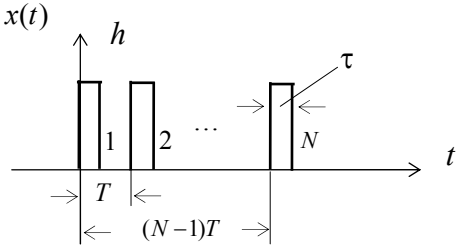


Рис. 2.10. Пачка прямоугольных импульсов

Спектральную плотность первого импульса в пачке обозначим $S_1(j\omega)$. Тогда для второго импульса, сдвинутого относительно первого на время T (в сторону запаздывания), спектральную функцию можно представить выражением

$$S_2(j\omega) = S_1(j\omega)e^{-j\omega T},$$

для третьего импульса –

$$S_3(j\omega) = S_1(j\omega)e^{-j2\omega T},$$

и так далее:

$$S_N(j\omega) = S_1(j\omega)e^{-j(N-1)\omega T}.$$

Таким образом, для группы из N импульсов в соответствии с принципом линейного суммирования спектров при сложении сигналов получим спектральную плотность

$$S(j\omega) = S_1(j\omega)\left[1 + e^{-j\omega T} + e^{-j2\omega T} + ... + e^{-j(N-1)\omega T}\right].$$

При частотах, отвечающих условию $\omega = k\dfrac{2\pi}{T}$ (k – целое), каждое из слагаемых в квадратных скобках равно единице и, следовательно, $S(jk\dfrac{2\pi}{T}) = NS_1(jk\dfrac{2\pi}{T})$. При частотах же $\omega = \dfrac{1}{N}\dfrac{2\pi}{T}$, а также при некоторых других частотах, для которых сумма векторов $e^{-j\omega T}$ обращается в нуль, суммарная спектральная плотность равна нулю.

При промежуточных значениях частот модуль $S(\omega)$ определяется как геометрическая сумма спектральных плотностей отдельных импульсов (рис. 2.11).

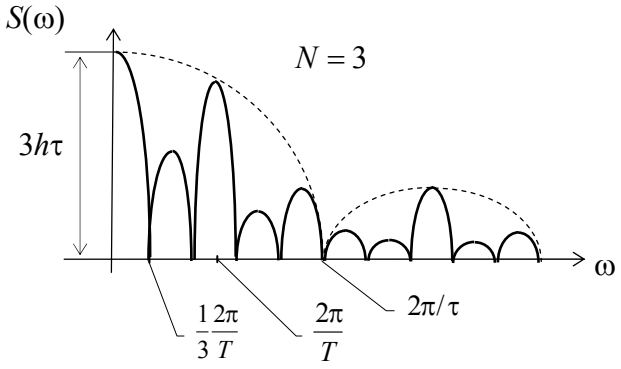


Рис. 2.11. Пачка прямоугольных импульсов

2.6. ЭНЕРГЕТИЧЕСКОЕ ТОЛКОВАНИЕ СПЕКТРА СИГНАЛА

Пусть колебание $x(t)$ представляет собой сложную периодическую функцию времени с периодом T. Энергия такого колебания, длящегося от $-\infty$ до $+\infty$, бесконечно велика. Основной интерес представляют средняя мощность периодического колебания и распределение этой мощности между отдельными гармониками. Очевидно, что средняя мощность колебания, рассматриваемого на всей оси времени, совпадает с мощностью, средней за один период T.

Таким образом, средняя мощность

$$\overline{x^2(t)} = \frac{1}{t_2 - t_1} \sum_{k=-\infty}^{\infty} |C_k|^2 \|\varphi_k\|^2 = \sum_{k=-\infty}^{\infty} |C_k|^2 = C_0^2 + 2\sum_{k=1}^{\infty} |C_k|^2 .$$

Используя тригонометрическую форму ряда Фурье и учитывая, что $C_0 = \frac{A_0}{2}$, $|C_k| = \frac{|A_k|}{2}$, получаем

$$\overline{x^2(t)} = \left(\frac{A_0}{2}\right)^2 + 2\sum_{k=1}^{\infty} \left(\frac{|A_k|}{2}\right)^2 = \left(\frac{A_0}{2}\right)^2 + \frac{1}{2}\sum_{k=1}^{\infty} |A_k|^2 . \qquad (2.6)$$

Если $x(t)$ представляет собой ток $i(t)$, то при прохождении его через сопротивление r выделяется средняя мощность

$$r \cdot \overline{i^2(t)} = r\left(I_0^2 + \frac{I_1^2}{2} + \frac{I_2^2}{2} + \ldots \right),$$

где I_0 – постоянная составляющая, а I_k – амплитуда k-й гармоники тока $i(t)$.

Итак, ***средняя мощность*** равна сумме средних мощностей, выделяемых отдельно постоянной составляющей I_0 и гармониками с амплитудами $I_1, I_2, \ldots$. Средняя мощность не зависит от фаз отдельных гармоник.

Рассмотрим теперь распределение ***энергии*** в спектре непериодического сигнала. Если функция $x(t)$ описывает непериодический ток конечной длительности, то $\mathcal{Э} = \int_{-\infty}^{\infty} [x(t)]^2 dt$ определяет энергию, выделяемую током в резисторе в 1 Ом.

Для установления распределения энергии по спектру непериодического сигнала выразим $\mathcal{Э}$ через модуль спектральной плотности сигнала $S(\omega)$:

$$\mathcal{Э} = \int_{-\infty}^{\infty} [x(t)]^2 dt = \int_{-\infty}^{\infty} x(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega \right] dt =$$

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

Поскольку произведение комплексно-сопряженных величин можно представить как квадрат модуля спектральной плотности $S(j\omega)S(-j\omega) = [S(\omega)]^2$, то энергию сигнала запишем в виде

$$\mathcal{Э} = \frac{1}{2\pi} \int_{-\infty}^{\infty} [S(\omega)]^2 d\omega = \frac{1}{\pi} \int_{0}^{\infty} [S(\omega)]^2 d\omega . \qquad (2.7)$$

Выражение (2.7), получившее название ***равенства Парсеваля***, показывает, что энергия сигнала может быть представлена в виде

суммы бесконечно малых слагаемых $\frac{1}{\pi}[S(\omega)]^2 d\omega$, соответствующих бесконечно малым участкам частотного спектра. Более того, выражение $\frac{1}{\pi}[S(\omega)]^2 d\omega$ представляет собой энергию, содержащуюся в спектральных составляющих сигнала, расположенных в полосе частот $d\omega$ в окрестности частоты ω (рис. 2.12).

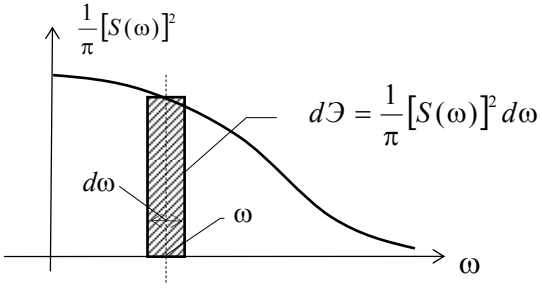


Рис. 2.12. Распределение энергии по частотному спектру

Между выражениями (2.6) и (2.7) имеется существенное различие. Сначала речь шла о средней мощности периодического колебания. Операция усреднения осуществлялась делением энергии отрезка колебания за один период на величину T. В случае же непериодического колебания конечной длительности усреднение энергии за бесконечно большой период дает нуль, и, следовательно, средняя мощность такого колебания равна нулю.

2.7. ПРАКТИЧЕСКАЯ ШИРИНА СПЕКТРА СИГНАЛА

Каждый реальный сигнал имеет конечную длительность и, следовательно, обладает бесконечным частотным спектром. Практически все каналы связи имеют ограниченную полосу пропускания. В результате при передаче сигнала через реальный канал связи может быть передана лишь часть его частотного спектра. Поэтому приходится беспокоиться о том, чтобы обеспечить пропускание через канал связи наиболее существенной части спектра. В связи с этим введено понятие практической ширины спектра сигнала.

За практическую ширину спектра сигнала принимают диапазон частот, в пределах которого находится наиболее существенная

часть спектра сигнала. Выбор количественной меры практической ширины спектра сигнала, как правило, осуществляется с помощью энергетического критерия.

С энергетической точки зрения, практическая ширина спектра определяется как область частот, в пределах которой сосредоточена подавляющая часть всей энергии сигнала.

В качестве примера рассмотрим одиночный прямоугольный импульс длительностью τ и величиной h. Энергия сигнала, сосредоточенная в полосе частот от 0 до ω_1, выражается функцией

$$Э(\omega_1) = \frac{1}{\pi}\int_0^{\omega_1}\left[S(\omega)\right]^2 d\omega = \frac{\tau^2 h^2}{\pi}\int_0^{\omega_1}\left|\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}\right|^2 d\omega .$$

Относительная величина энергии одиночного импульса, сосредоточенная в полосе частот от 0 до ω_1, равна

$$\lambda(\omega_1) = \frac{Э(\omega_1)}{Э} = \frac{\tau}{\pi}\int_0^{\omega_1}\left|\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}\right|^2 d\omega ,$$

где $Э = \int_{-\infty}^{\infty}\left|x(t)\right|^2 dt = \tau h^2$ – полная энергия одиночного прямоугольного импульса.

График функции $\lambda(\omega_1)$, называемой ***интегральной кривой распределения энергии*** сигнала в спектре, приведен на рис. 2.13. Как видно из рисунка, в полосе частот от 0 до $\omega_1 = 2\pi/\tau$ сосредоточено более 90 % всей энергии сигнала, и данный диапазон частот может быть взят в качестве практической ширины спектра одиночного прямоугольного импульса. При этом дальнейшее увеличение практической ширины спектра ведет к незначительному увеличению энергии в данном диапазоне частот, так как при $\omega_1 > 2\pi/\tau$ кривая $\lambda(\omega_1)$ довольно пологая.

Аналогично определяется практическая ширина спектра непериодических сигналов любой другой формы.

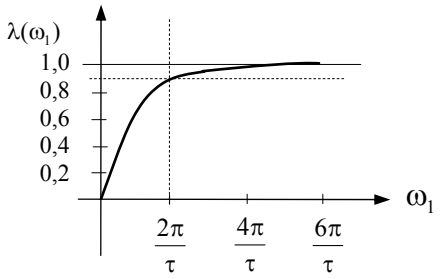


Рис. 2.13. Интегральная кривая распределения энергии

Из рис. 2.14 видно, что спектры сигналов обладают бесконечной протяженностью и имеют тенденцию к затуханию с увеличением частоты. Форма и характер его затухания зависят от формы импульсов и его длительности. Следовательно, форма и ширина импульса влияют на действительную ширину спектра.

Наиболее экономичным с этой точки зрения является колокольный импульс, который требует наименьшей полосы частот при заданной длительности.

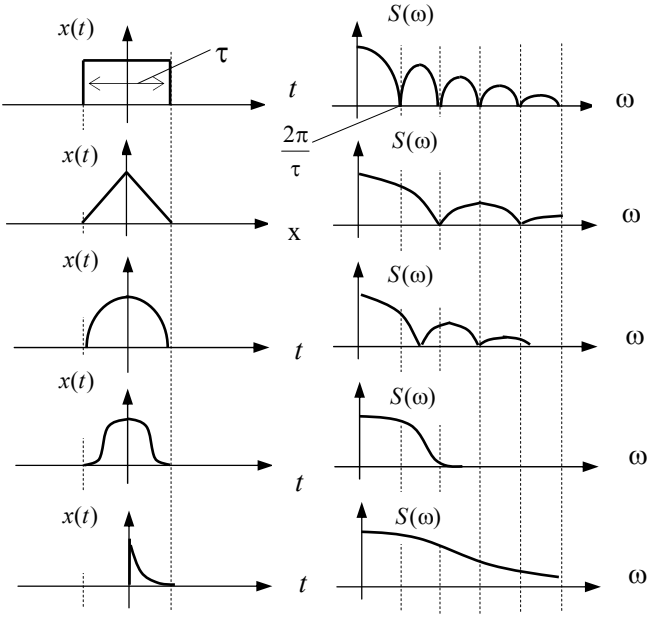


Рис. 2.14. Спектры одиночных импульсов

2.8. СПЕКТР ДЕЛЬТА-ФУНКЦИИ

Под ***дельта-функцией*** понимают сигнал, представленный на рис. 2.15. Дельта-функцию можно трактовать как предел прямоугольного импульса длительности τ и амплитуды $1/\tau$, получаемый при $\tau \to 0$ (рис. 2.16).

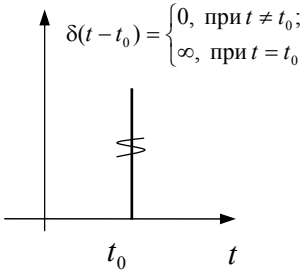


Рис. 2.15. Дельта-функция

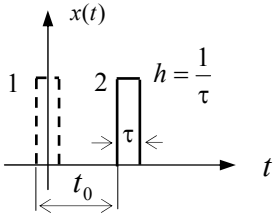


Рис. 2.16. Иллюстрация предельного перехода

Сдвиг импульса в сторону запаздывания можно учесть следующим образом:

$$S_2(j\omega) = S_1(j\omega)e^{-j\omega t_0}, \text{ где } S_1(j\omega) = h\tau \operatorname{sinc}\frac{\omega\tau}{2}.$$

Тогда, приняв амплитуду импульса равной $h = 1/\tau$, получим

$$S_\delta(j\omega) = \lim_{\tau\to 0}\left[\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}e^{-j\omega t_0}\right] = e^{-j\omega t_0}.$$

На рис. 2.17 представлены модуль и фаза спектральной плотности дельта-функции.

Прямое и обратное преобразование Фурье для дельта-функции можно представить в виде

$$S_\delta(j\omega) = \int_{-\infty}^{\infty}\delta(t-t_0)e^{-j\omega t}dt = e^{-j\omega t_0}, \quad \delta(t-t_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{j\omega(t-t_0)}d\omega.$$

Данные выражения полностью проясняют смысл известного свойства дельта-функции $\int_{-\infty}^{\infty}\delta(t-t_0)f(t)dt = f(t_0)$.

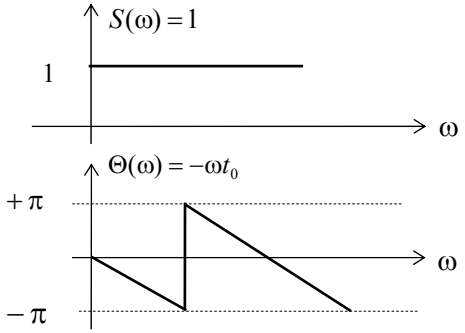


Рис. 2.17. Спектр дельта-функции

ЗАДАЧИ

2.1. Найти спектральную плотность одиночного импульса высокочастотных колебаний (рис. 2.18)

$$x(t) = \begin{cases} h\cos\omega_0 t, & -\dfrac{\tau}{2} \le t \le \dfrac{\tau}{2}; \\ 0, & |t| > \dfrac{\tau}{2}. \end{cases}$$

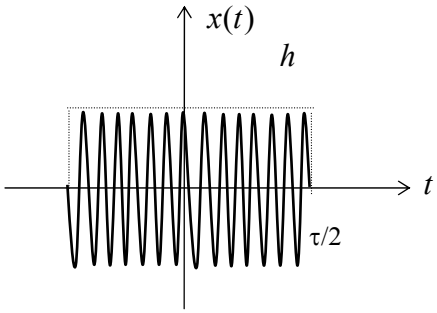


Рис. 2.18. Одиночный импульс высокочастотных колебаний

2.2. Найти спектральную плотность одиночного экспоненциального импульса (рис. 2.19)

$$x(t) = \begin{cases} he^{-\beta(t-t_0)}, & t \geq t_0; \\ 0, & t < t_0. \end{cases}$$

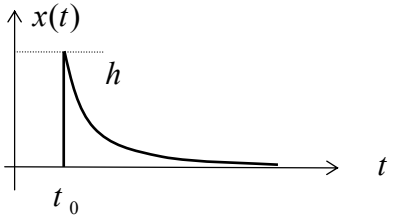


Рис. 2.19. Экспоненциальный импульс

2.3. Вывести формулы модуля $S(\omega)$ и аргумента $\Theta(\omega)$ спектральной плотности сигнала включения (рис. 2.20)

$$x(t) = h \cdot 1(t) = \begin{cases} h, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

Рис. 2.20. Сигнал включения

2.4. Найти реакцию фильтра нижних частот с граничной частотой ω_c на дельта-импульс амплитудой h (рис. 2.21) при условии, что $\omega_c < \dfrac{2\pi}{\tau}$.

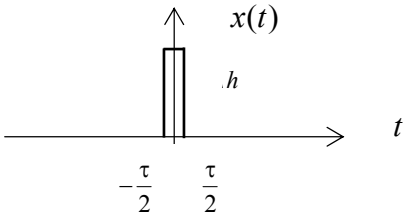


Рис. 2.21. Дельта-импульс

2.5. Для дельта-функции (рис. 2.22) доказать равенство

$$\delta(t - t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega(t-t_0)} d\omega \ .$$

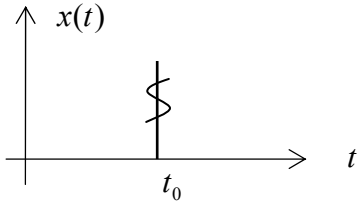


Рис. 2.22. Дельта-функция

2.6. На вход цепи (рис. 2.23) поступает сигнал (рис. 2.24)

$$x(t) = \begin{cases} \dfrac{1}{RC} e^{\frac{(t-T_0)}{RC}}, & t \le T_0; \\ 0, & t > T_0. \end{cases}$$

Определить выходной сигнал $x^*(t)$.

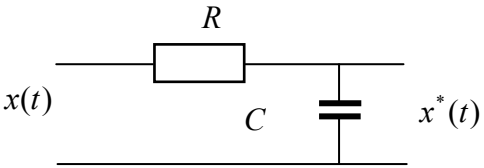


Рис. 2.23. Интегрирующая цепочка

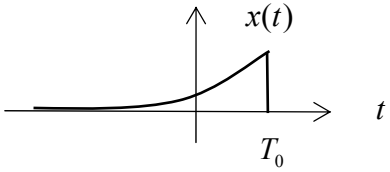


Рис. 2.24. Входной сигнал

2.7. Найти спектральную плотность импульса (рис. 2.25)

$$x(t) = h\omega_m \frac{\sin(\omega_m t / 2)}{\omega_m t / 2} \quad .$$

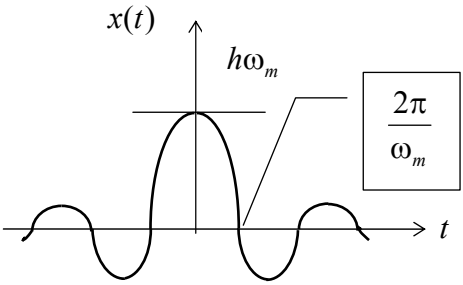


Рис. 2.25. Одиночный импульс

3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СЛУЧАЙНЫХ СИГНАЛОВ

К случайным сигналам относят сигналы, значения которых заранее неизвестны и могут быть предсказаны с некоторой вероятностью. По существу, любой сигнал, несущий в себе информацию, должен рассматриваться как случайный.

До приема сообщения сигнал следует рассматривать как случайный процесс, представляющий собой совокупность функций времени, подчиняющихся некоторой общей для них статистической закономерности. Одна из этих функций, ставшая полностью известной после приема сообщения, называется реализацией случайного процесса. Эта реализация является уже не случайной, а детерминированной функцией времени.

На рис. 3.1 приведены несколько характерных случайных процессов.

В реальных условиях все сигналы имеют случайный характер. Вследствие этого получателю заранее неизвестно, каким будет сигнал.

Однако нельзя также утверждать, что приемная сторона не располагает абсолютно никакими предварительными (априорными) данными о сигналах. Во-первых, обычно предварительно известно все множество, весь ансамбль возможных сигналов. Во-вторых, как правило, имеются сведения об ожидаемой вероятности тех или иных сигналов из общего ансамбля сигналов.

Таким образом, предварительные сведения о сигналах, которыми мы располагаем, носят статистический характер. Поэтому для исследования прохождения сигналов через информационные системы следует применять статистические методы. Целесообразность применения статистических методов обусловлена еще и тем, что на сигнал воздействуют помехи, представляющие, как правило, неизвестную функцию времени.

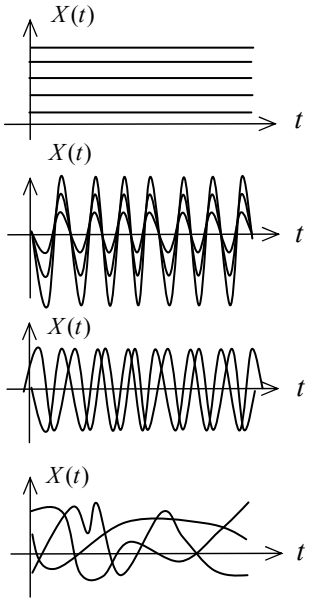
Основным содержанием задачи приема сигналов на фоне помех является наиболее полное извлечение информации из сигнала. Успешного решения этой задачи можно достичь только на основе использования статистических методов приема.

Для характеристики и анализа случайных сигналов применяется

статистический подход. В качестве основных характеристик случайных сигналов принимают:

– закон распределения вероятностей;

– спектральное распределение средней мощности сигнала.



– сигнал в виде постоянного напряжения случайного уровня;

– гармоническое колебание со случайной амплитудой;

– гармоническое колебание со случайной фазой;

– возможные реализации произвольного случайного процесса

Рис. 3.1. Примеры совокупностей функций времени, подчиняющихся некоторой общей для них статистической закономерности

3.1. СЛУЧАЙНЫЕ СИГНАЛЫ И ИХ ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ

Конкретный вид, принимаемый случайным процессом $X(t)$ в результате опыта, называется ***реализацией процесса*** $x_k(t)$. Отдельные наблюдения над случайным процессом, протекающим в однотипных системах при одинаковых условиях опыта, дадут различные реализации случайного процесса $x_1(t), x_2(t), ..., x_k(t), ...$ Вид функции $x_k(t)$ случайным образом меняется от одного опыта к другому (рис. 3.2). Совокупность реализаций случайного процесса, полученных в результате опытов, называется ***ансамблем реали-***

заций случайного процесса $X(t)$.

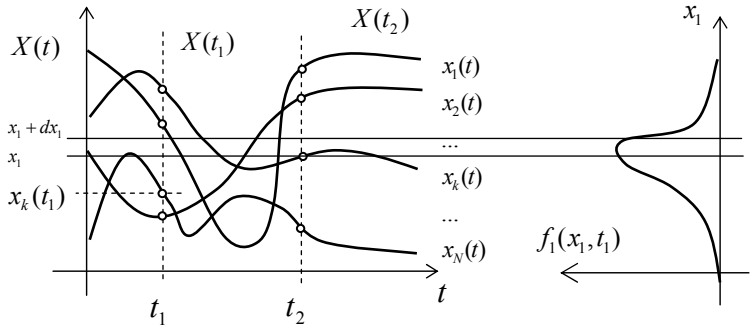


Рис. 3.2. Реализации случайного процесса *X*(*t*)

Величина k -й реализации случайного процесса в определенный момент времени (например, t_1) называется ***выборкой случайного процесса*** $x_k(t_1)$. Совокупность значений выборок в определенный момент времени t_1 образует ***случайную величину*** $X(t_1)$.

Вероятность того, что в момент времени t_1 величина $X(t_1)$ находится в интервале между x_1 и $x_1 + dx_1$, имеет вид

$$P\left[x_1 \le X(t_1) < x_1 + dx_1\right] = f_1(x_1, t_1)dx_1,$$

где $f_1(x_1, t_1)$ – ***одномерная плотность распределения вероятностей*** случайной величины $X(t_1)$.

Плотность вероятности $f_1(x_1, t_1)$ есть функция времени, так как зависит от t_1 и является частной производной от функции распределения случайной величины $X(t_1)$:

$$f_1(x_1, t_1) = \frac{\partial F_1(x_1, t_1)}{\partial x_1}; \quad F_1(x_1, t_1) = P\left[X(t_1) < x_1\right] = \int_{-\infty}^{x_1} f_1(\xi, t_1)d\xi.$$

Одномерный закон плотности распределения вероятностей является простейшей статистической характеристикой случайного процесса. Он дает представление о процессе лишь в отдельные, фиксированные моменты времени. Другими словами, характеризует процесс статически и не дает представления о динамике его

развития.

Для более полной характеристики случайного процесса необходимо знать связь между вероятными значениями случайной функции при двух произвольных моментах времени t_1 и t_2. Эта связь выражается через двумерную плотность распределения вероятностей и формулируется следующим образом:

«Вероятность нахождения любой из функций $X(t_i)$, входящих в случайный процесс $X(t)$, в интервале $(x_1, x_1 + dx_1)$ в момент времени t_1 и в интервале $(x_2, x_2 + dx_2)$ в момент времени t_2, имеет вид

$P[x_1 \le X(t_1) < x_1 + dx_1; x_2 \le X(t_2) < x_2 + dx_2] = f_2(x_1, t_1; x_2, t_2)dx_1 dx_2$,

где $f_2(x_1, t_1; x_2, t_2)$ – ***двумерная плотность распределения вероятностей*** случайного процесса $X(t)$».

Рассуждая аналогичным образом, можно ввести понятие о n-мерной $f_n(x_1, t_1; x_2, t_2; ...; x_n, t_n)$ плотности распределения вероятностей случайного процесса. Тогда вероятность сложного события, заключающегося в том, что в момент времени t_i функция $X(t_i)$ находится в интервале $(x_i, x_i + dx_i)$, $i = \overline{1, n}$, равна

$$P\left[x_i \le X(t_i) < x_i + dx_i\big|_{i=\overline{1,n}}\right] = f_n(x_1, t_1; ...; x_n, t_n)dx_1 ... dx_n .$$

Чем больше число n, тем точнее n-мерная функция плотности характеризует статистические свойства случайного процесса. Однако n-мерные функции могут быть получены с помощью довольно сложной и трудоемкой обработки множества реализаций случайного процесса. При пользовании n-мерными функциями встречаются существенные математические трудности. Поэтому на практике чаще всего оперируют конечным числом числовых характеристик, которые дают, безусловно, менее полную картину процесса, но достаточны для решения ряда важных задач.

3.2. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНОГО ПРОЦЕССА

Простейшими моментными функциями, в основном используемыми для характеристики случайных процессов, являются момен-

ты первых двух порядков: математическое ожидание, дисперсия и корреляционная функция случайного процесса.

Математическое ожидание или первый начальный момент одномерного закона распределения выражается формулой

$$m[X(t_1)] = \int_{-\infty}^{\infty} x_1 f_1(x_1, t_1) dx_1 .$$

Физически математическое ожидание выражает среднее значение совокупности выборок случайного процесса (случайной величины $X(t_1)$) в момент времени t_1.

Дисперсия или второй центральный момент одномерного закона распределения – это математическое ожидание квадрата отклонения величины $X(t_1)$ от математического ожидания в момент времени t_1:

$$D[X(t_1)] = m\{[X(t_1) - m[X(t_1)]]^2\} = \int_{-\infty}^{\infty} \{x_1 - m[X(t_1)]\}^2 f_1(x_1, t_1) dx_1 .$$

Дисперсия выражает меру разброса значений случайной величины $X(t_1)$ около математического ожидания. С учетом того, что $\int_{-\infty}^{\infty} f_1(x_1, t_1) dx_1 = 1$, можно получить более простое выражение: $D[X(t_1)] = m[X^2(t_1)] - m^2[X(t_1)]$.

Корень квадратный из дисперсии принято называть ***среднеквадратическим отклонением*** случайной величины

$$\sigma[X(t_1)] = \sqrt{D[X(t_1)]} .$$

Аналогично можно найти среднее значение квадрата случайной величины $X(t_1)$:

$$m[X^2(t_1)] = \int_{-\infty}^{\infty} x_1^2 f_1(x_1, t_1) dx_1 .$$

При этом $\sqrt{m[X^2(t_1)]}$ называется ***среднеквадратическим значением*** $X(t_1)$.

Приведенные характеристики приближенно характеризуют поведение случайного процесса в отдельные моменты времени, но совершенно не затрагивают связь между значениями случайного

процесса в различные моменты времени. Эту связь выражает ***автокорреляционная функция*** $K_{XX}(t_1,t_2)$, определяемая как среднее значение произведения значений случайных величин в моменты времени t_1 и t_2:

$$K_{XX}(t_1,t_2)=m[X(t_1)X(t_2)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2f_2(x_1,t_1;x_2,t_2)dx_1dx_2\,.$$

Большое число задач, связанных с описанием случайных сигналов, удается решать на основе двумерной плотности распределения вероятностей $f_2(x_1,t_1;x_2,t_2)$.

Иногда рассматривают нормированную автокорреляционную функцию (***коэффициент автокорреляции***)

$$r_{XX}(t_1,t_2)=\frac{K_{XX}(t_1,t_2)}{\sqrt{m[X^2(t_1)]m[X^2(t_2)]}}\,.$$

Усредненные характеристики могут быть также получены путем обработки одной из реализаций $x_k(t)$ случайного процесса на достаточно большом интервале времени.

Среднее по времени значение случайного процесса определяется выражением

$$\overline{x}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x_k(t)dt\,,$$

где $x_k(t)$ – реализация случайного процесса $X(t)$, T – время наблюдения процесса.

По аналогии пользуются понятиями среднего по времени значения от функции $x_k^2(t)$, от квадрата разности $[x_k(t)-\overline{x}]^2$ и от произведения $x_k(t)x_k(t+\tau)$:

$$\overline{x^2}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x_k^2(t)dt\,;$$

$$\overline{[x-\overline{x}]^2}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}[x_k(t)-\overline{x}]^2dt=\overline{x^2}-(\overline{x})^2\,;$$

$$\overline{x(t)x(t+\tau)}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x_k(t)x_k(t+\tau)dt\,.$$

Если $X(t)$ представляет собой электрический сигнал (ток), то $\bar{x}$ – постоянная составляющая случайного сигнала, $\overline{x^2}$ – средняя мощность, рассеиваемая на сопротивлении в 1 Ом, $\overline{[x-\bar{x}]^2}$ – средняя мощность флуктуации сигнала.

3.3. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

ОПРЕДЕЛЕНИЕ 3.1. Под ***стационарными процессами*** понимаются случайные процессы, для которых функция плотности распределения вероятностей $f_n(x_1,t_1;x_2,t_2;...;x_n,t_n)$ произвольного порядка n не меняется при любом сдвиге всей группы точек $t_1,t_2,...,t_n$ вдоль оси времени, т.е. для любых n и τ

$$f_n(x_1,t_1;x_2,t_2;...;x_n,t_n)=f_n(x_1,t_1+\tau;x_2,t_2+\tau;...;x_n,t_n+\tau).$$

Стационарные процессы имеют вид непрерывных случайных колебаний около некоторого среднего значения, причем ни среднее значение, ни характер этих колебаний не претерпевают существенных изменений во времени. Строго говоря, стационарные процессы бесконечны во времени, т.е. не имеют ни начала, ни конца. Таких процессов практически нет. Однако многие случайные процессы на определенных отрезках времени с определенным приближением можно считать стационарными.

Для стационарных процессов:

1. Одномерная функция плотности распределения вероятностей не зависит от времени, т.е.

$$f_1(x_1,t_1)=f_1(x_1,t_1+\tau)=f_1(x).$$

2. Двумерная функция плотности зависит только от разности времени $t_2-t_1=\tau$, т.е.

$$f_2(x_1,t_1;x_2,t_2)=f_2(x_1,x_2;t_2-t_1)=f_2(x_1,x_2;\tau).$$

Поскольку математическое ожидание и дисперсия выражаются через одномерную функцию плотности, то можно утверждать, что для стационарного процесса математическое ожидание и дисперсия не зависят от времени. Вследствие зависимости двумерной функции плотности только от разности времен $\tau=t_2-t_1$, автокорреляционная функция стационарного процесса также зависит только от разности τ.

Следовательно,

$$m[X(t_1)] = m[X],$$
$$D[X(t_1)] = D[X],$$
$$m[X^2(t_1)] = m[X^2],$$
$$K_{XX}[t_1, t_2] = K_{XX}(\tau).$$

Более того, существует класс случайных стационарных процессов, обладающих важным для практических приложений свойством эргодичности.

Стационарный процесс называется ***эргодическим***, если усреднение по множеству с вероятностью, сколь угодно близкой к единице, равно усреднению по времени:

$$m[X] = \int_{-\infty}^{\infty} x f_1(x)dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_k(t)dt = \bar{x},$$

$$D[X] = \int_{-\infty}^{\infty} (x - m[X])^2 f_1(x)dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [x_k(t) - \bar{x}]^2 dt = \overline{[x - \bar{x}]^2},$$

$$m[X^2] = \int_{-\infty}^{\infty} x^2 f_1(x)dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_k^2(t)dt = \overline{x^2},$$

$$K_{XX}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_2(x_1, x_2, \tau)dx_1 dx_2 = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_k(t)x_k(t+\tau)dt = \overline{x(t)x(t+\tau)}.$$

Стационарные случайные процессы по своей природе проще, чем нестационарные и описываются более простыми характеристиками. Ввиду того, что стационарные процессы встречаются на практике очень часто, получила широкое применение специальная теория стационарных случайных процессов.

Так как свойства стационарного процесса во многом определяются свойствами автокорреляционной функции, то для изучения такого процесса нужно, в первую очередь, определить свойства автокорреляционной функции.

3.4. СВОЙСТВА АВТОКОРРЕЛЯЦИОННОЙ ФУНКЦИИ СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

1. Определим, как ведет себя автокорреляционная функция $K_{XX}(\tau)$ при неограниченном увеличении разностного временного

интервала $\tau = t_2 - t_1$.

По мере увеличения τ зависимость между величинами $X(t)$ и $X(t+\tau)$ ослабевает. В пределе при $\tau \to \infty$ эти величины становятся независимыми. Тогда с учетом того, что математическое ожидание произведения случайных независимых величин равно произведению математических ожиданий сомножителей и что для стационарного процесса математическое ожидание не зависит от времени, получим

$$K_{XX}(\infty) = \lim_{\tau\to\infty} K_{XX}(\tau) = \lim_{\tau\to\infty} m[X(t)X(t+\tau)] = m[X(t)]m[X(t+\infty)] = m^2[X].$$

Таким образом, при $\tau \to \infty$ автокорреляционная функция стремится к квадрату математического ожидания случайного процесса, который можно трактовать как мощность постоянной составляющей реализаций случайного стационарного процесса.

2. При уменьшении интервала τ зависимость между величинами $X(t)$ и $X(t+\tau)$ усиливается, и в пределе при $\tau \to 0$ получим

$$K_{XX}(0) = \lim_{\tau\to 0} K_{XX}(\tau) = \lim_{\tau\to 0} m[X(t)X(t+\tau)] = m[X^2(t)] = m[X^2].$$

Таким образом, при $\tau = 0$ автокорреляционная функция равна начальному моменту второго порядка функции $X(t)$, который можно трактовать как среднюю мощность случайного стационарного процесса.

3. Дисперсия случайного стационарного процесса удовлетворяет равенству

$$D[X] = m[X^2] - m^2[X] = K_{XX}(0) - K_{XX}(\infty) \ .$$

4. В силу независимости $f_2(x_1, x_2; \tau)$ от начала отсчета

$$K_{XX}(\tau) = K_{XX}(-\tau) \ .$$

5. Автокорреляционная функция по абсолютному значению максимальна при $\tau = 0$:

$$\left|K_{XX}(\tau)\right| \le K_{XX}(0) \ .$$

Типичные кривые автокорреляционной функции $K_{XX}(\tau)$ имеют вид, представленный на рис. 3.3. Как видно из рисунка, асимптотическое приближение функции $K_{XX}(\tau)$ к установившемуся значению $m^2[X]$ может происходить как монотонно, так и немонотонно.

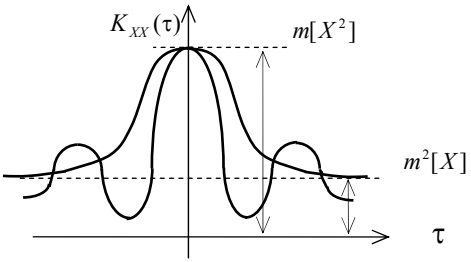


Рис. 3.3. Вид автокорреляционной функции процесса $X(t)$

На практике часто вместо случайного процесса $X(t)$ рассматривают его отклонение от математического ожидания $\overset{\circ}{X}(t) = X(t) - m[X]$, называемое ***пульсациями процесса***.

Автокорреляционная функция пульсаций стационарного случайного процесса равна

$$\overset{\circ}{K}_{XX}(\tau) = m[\overset{\circ}{X}(t)\,\overset{\circ}{X}(t+\tau)] = m\{[X(t) - m[X]][X(t+\tau) - m[X]]\} =$$
$$= m[X(t)X(t+\tau)] - 2m^2[X] + m^2[X] = K_{XX}(\tau) - m^2[X].$$

Математическое ожидание пульсаций равно нулю $m[\overset{\circ}{X}] = 0$, а дисперсия

$$D[\overset{\circ}{X}] = \overset{\circ}{K}_{XX}(0) - \overset{\circ}{K}_{XX}(\infty) = K_{XX}(0) - m^2[X] = m[X^2] - m^2[X] = D[X].$$

Нормированная автокорреляционная функция пульсаций случайного процесса приведена на рис. 3.4.

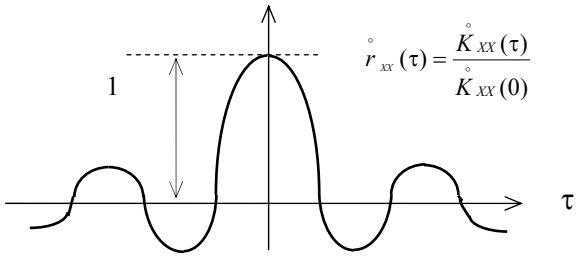


Рис. 3.4. Вид коэффициента автокорреляции пульсаций $\overset{\circ}{X}(t)$

3.5. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ

Наряду со спектральным подходом к описанию детерминированных сигналов часто на практике оказывается необходимой характеристика, которая давала бы представление о некоторых свойствах сигнала (скорость изменения во времени) без разложения его на гармонические составляющие.

В качестве такой временной характеристики широко используется автокорреляционная функция.

Для детерминированного сигнала $x(t)$ конечной длительности автокорреляционная функция определяется следующим выражением:

$$K_{XX}^{\text{д}}(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)dt\,.$$

Из этого выражения видно, что $K_{XX}^{\text{д}}(\tau)$ характеризует степень связи сигнала $x(t)$ со своей копией, сдвинутой на величину τ по оси времени. Ясно, что эта функция достигает максимума при $\tau = 0$, так как любой сигнал полностью коррелирован с самим собой. При этом $K_{XX}^{\text{д}}(0) = \int_{-\infty}^{\infty} x^2(t)dt = Э$, т.е. максимальное значение автокорреляционной функции равно энергии сигнала.

С увеличением τ функция $K_{XX}^{\text{д}}(\tau)$ убывает и при относительном сдвиге сигналов $x(t)$ и $x(t+\tau)$ на величину, превышающую длительность сигнала, обращается в нуль.

На рис. 3.5 показано построение автокорреляционной функции для простейшего сигнала в виде прямоугольного импульса

$$K_{XX}^{\text{д}}(\tau) = \int_{t_1}^{t_2-\tau} h^2 dt = h^2[(t_2-\tau)-t_1] = h^2(\tau_{\text{и}}-\tau)\,.$$

Для периодического сигнала, энергия которого бесконечна, определение автокорреляционной функции аналогично неприемлемо. В этом случае исходят из следующего определения:

$$K_{XX}^{\text{дп}}(\tau) = \lim_{T' \to \infty} \frac{1}{T'} \int_{-T'/2}^{T'/2} x(t)x(t+\tau)dt \, .$$

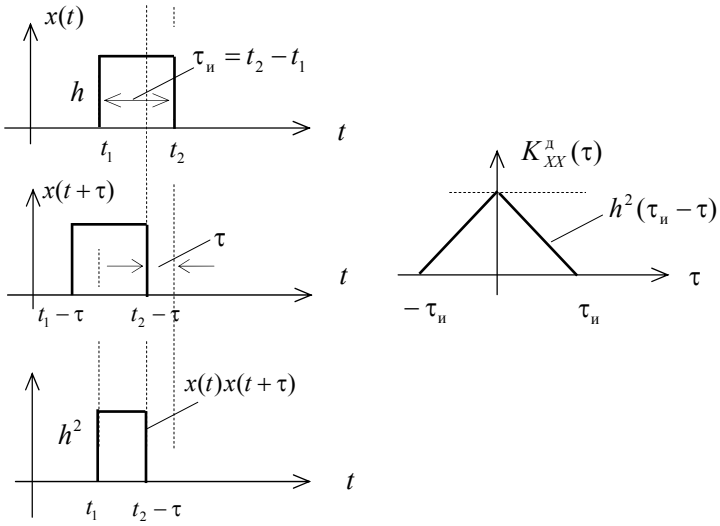


Рис. 3.5. Иллюстрация $K_{XX}^{\text{д}}(\tau)$ для прямоугольного импульса

При таком определении $K_{XX}^{\text{дп}}(\tau)$ приобретает размерность мощности, причем $K_{XX}^{\text{дп}}(0)$ равна средней мощности периодического сигнала. Ввиду периодичности усреднение по бесконечно большому отрезку T' должно совпадать с усреднением по периоду T. Поэтому $K_{XX}^{\text{дп}}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t+\tau)dt$.

Представив периодическую функцию разложением в ряд Фурье, можно выяснить природу автокорреляционной функции

$$K_{XX}^{\text{дп}}(\tau) = \frac{1}{4T} \int_{-T/2}^{T/2} \sum_{l=-\infty}^{\infty} A_l e^{jl\omega_0 t} \sum_{k=-\infty}^{\infty} A_k e^{jk\omega_0(t+\tau)} dt =$$

$$= \frac{1}{4T} \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} A_l A_k e^{jk\omega_0\tau} \int_{-T/2}^{T/2} e^{j(l+k)\omega_0 t} dt =$$

$$= \frac{1}{4} \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} A_l A_k \frac{\sin(l+k)\omega_0 \frac{T}{2}}{(l+k)\omega_0 \frac{T}{2}} e^{jk\omega_0 \tau}.$$

Так как $\omega_0 \frac{T}{2} = \pi$, то дробь

$$\frac{\sin(l+k)\omega_0 \frac{T}{2}}{(l+k)\omega_0 \frac{T}{2}} = \begin{cases} 0, \text{ при } l+k \neq 0; \\ 1, \text{ при } l+k = 0. \end{cases}$$

Следовательно,

$$K_{XX}^{\text{дп}}(\tau) = \frac{1}{4} \sum_{k=-\infty}^{\infty} A_{-k} A_k e^{jk\omega_0 \tau} = \frac{1}{4} \sum_{k=-\infty}^{\infty} |A_k|^2 e^{jk\omega_0 \tau} =$$

$$= \left(\frac{A_0}{2} \right)^2 + \frac{1}{2} \sum_{k=1}^{\infty} |A_k|^2 \cos k\omega_0 \tau \ .$$

Как видно из полученного выражения, автокорреляционная функция $K_{XX}^{\text{дп}}(\tau)$ периодического сигнала с периодом T представляет собой периодическую функцию аргумента τ с тем же периодом T. Амплитуда гармоники корреляционной функции $K_{XX}^{\text{дп}}(\tau)$ равна половине квадрата амплитуды соответствующей гармоники $x(t)$.

Более того, при $\tau = 0$ в чистом виде получаем распределение средней мощности по частотному спектру

$$K_{XX}^{\text{дп}}(0) = \left(\frac{A_0}{2} \right)^2 + \frac{1}{2} \sum_{k=1}^{\infty} |A_k|^2 ,$$

что согласуется с результатом, полученным в подразделе 2.6 настоящего пособия.

3.6. СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ МОЩНОСТИ СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

При изучении детерминированных сигналов весьма удобным оказался гармонический анализ. В связи с этим желательно ис-

пользовать аппарат преобразований Фурье к случайным процессам.

Однако необходимо иметь в виду, что реализациям, обладающим различной формой, соответствуют различные спектральные характеристики. Усреднение комплексной спектральной плотности по всем реализациям приводит к нулевому спектру процесса из-за случайности и независимости фаз спектральных составляющих в различных реализациях.

Тем не менее, можно ввести понятие спектральной плотности среднего квадрата случайной функции, поскольку величина среднего квадрата не зависит от соотношения фаз суммируемых гармоник.

Выделим из ансамбля какую-либо реализацию $x_k(t)$ случайного процесса $X(t)$ и ограничим ее длительность конечным интервалом T (рис. 3.6). Для такой функции справедливо преобразование Фурье

$$S_k^{'}(j\omega) = \int_{-T/2}^{T/2} x_k^{'}(t) e^{-j\omega t} dt .$$

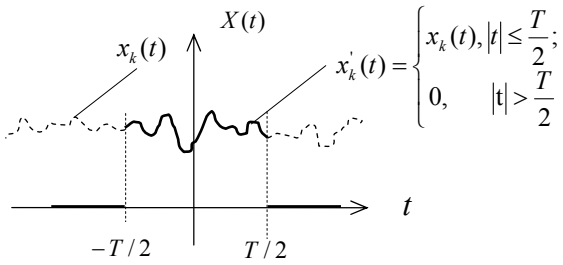


Рис. 3.6. Реализация $x_k(t)$ случайного процесса $X(t)$

Тогда энергию рассматриваемого отрезка реализации можно вычислить с помощью формулы Парсеваля

$$Э_k^{'} = \int_{-T/2}^{T/2} \left[x_k^{'}(t)\right]^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|S_k^{'}(j\omega)\right|^2 d\omega .$$

Разделив эту энергию на T, получим среднюю мощность k - й реализации на отрезке T

$$\frac{Э_k^{'}}{T}=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{\left|S_k^{'}(j\omega)\right|^2}{T}d\omega=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}G_k^{'}(\omega)d\omega\,,$$

где $G_k^{'}(\omega)=\dfrac{\left|S_k^{'}(j\omega)\right|^2}{T}$ представляет собой спектральную плотность средней мощности рассматриваемой k-й реализации на отрезке T.

В общем случае величина $G_k^{'}(\omega)$ должна быть усреднена по множеству реализаций:

$$m[G_k^{'}(\omega)]=\frac{1}{T}m\left[\left|S_k^{'}(j\omega)\right|^2\right]=\frac{1}{T}m\left[S_k^{'}(j\omega)S_k^{'}(-j\omega)\right]=$$

$$=\frac{1}{T}m\left[\int\limits_{-T/2}^{T/2}x_k^{'}(t_1)e^{-j\omega t_1}dt_1\int\limits_{-T/2}^{T/2}x_k^{'}(t_2)e^{j\omega t_2}dt_2\right]=$$

$$=\frac{1}{T}\int\limits_{-T/2}^{T/2}\int\limits_{-T/2}^{T/2}m\left[x_k^{'}(t_1)x_k^{'}(t_2)\right]e^{-j\omega(t_1-t_2)}dt_1dt_2\ .$$

Так как $m\left[x_k^{'}(t_1)x_k^{'}(t_2)\right]=K_{XX}(t_1,t_2)=K_{XX}(\tau)$, то

$$m[G_k^{'}(\omega)]=\frac{1}{T}\int\limits_{-T/2}^{T/2}K_{XX}(\tau)e^{-j\omega\tau}d\tau\int\limits_{-T/2}^{T/2}dt_2=\int\limits_{-T/2}^{T/2}K_{XX}(\tau)e^{-j\omega\tau}d\tau\,.$$

Наконец, переходя к пределу при $T\to\infty$, окончательно получим

$$G_{XX}(\omega)=\lim_{T\to\infty}m[G_k^{'}(\omega)]=\int\limits_{-\infty}^{\infty}K_{XX}(\tau)e^{-j\omega\tau}d\tau=2\int\limits_{0}^{\infty}K_{XX}(\tau)\cos\,\omega\tau\,d\tau.\quad(3.1)$$

Таким образом, ***спектральная плотность мощности*** $G_{XX}(\omega)$, являющаяся усредненной характеристикой совокупности реализаций случайного процесса, представляет собой прямое преобразование Фурье для корреляционной функции. Тогда обратное преобразование Фурье будет иметь вид

$$K_{XX}(\tau)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}G_{XX}(\omega)e^{j\omega\tau}d\omega=\frac{1}{\pi}\int\limits_{0}^{\infty}G_{XX}(\omega)\cos\,\omega\tau\,d\omega.\qquad(3.2)$$

Преобразования (3.1) и (3.2), связывающие функции $G_{XX}(\omega)$ и $K_{XX}(\tau)$, носят название ***преобразований Хинчина–Винера***.

Для выяснения физического смысла функции $G_{XX}(\omega)$ примем в

(3.2) $\tau = 0$. Так как $K_{XX}(0) = \frac{1}{\pi}\int_0^\infty G_{XX}(\omega)d\omega$ выражает среднюю мощность сигнала, то $G_{XX}(\omega)$ дает усредненную энергетическую картину распределения мощности сигнала по частотному спектру, а элементарная составляющая $\frac{1}{\pi}G_{XX}(\omega)d\omega$ представляет собой долю средней мощности, приходящуюся на диапазон частот $d\omega$.

3.7. БЕЛЫЙ ШУМ

Случайный процесс, имеющий равномерный на всех частотах спектр, называют ***белым шумом*** (рис. 3.7,а).

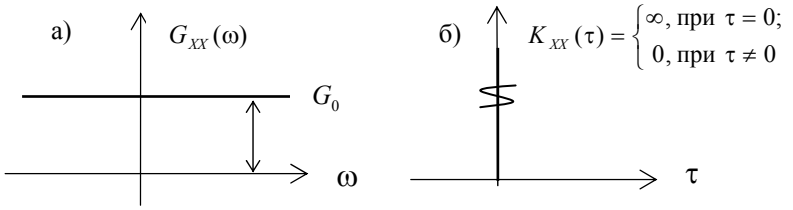


Рис. 3.7. Спектральные характеристики белого шума:
а – спектральная плотность мощности; б – автокорреляционная функция

Автокорреляционная функция белого шума может быть определена следующим образом (рис. 3.7,б)

$$K_{XX}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_0 e^{j\omega\tau} d\omega = G_0\delta(\tau),$$

так как интеграл $\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega\tau} d\omega$ является обратным преобразованием Фурье для дельта-функции $\delta(\tau)$.

Очевидно, что белый шум $X(t)$ характеризуется тем, что значения в любые два (сколь угодно близкие) момента времени некоррелированы. Такой процесс называют абсолютно случайным.

Необходимо подчеркнуть, что понятие белый шум основано на спектральном свойстве случайного процесса и совершенно не связано с законами распределения вероятностей. В частности, если

белый шум имеет нормальный закон распределения, то его называют нормальным белым шумом или Гауссовым шумом.

Белый шум в точном смысле является идеализацией, никогда не встречающейся в реальных условиях, хотя бы потому, что достаточно близкие значения случайной функции практически всегда зависимы, а также и потому, что реальные процессы имеют конечную мощность, а для белого шума полная мощность процесса бесконечна. Однако подобная идеализация во многих важных практических случаях значительно упрощает математический анализ и не вносит сколько-нибудь существенных погрешностей.

Спектры реальных процессов практически ограничены полосой частот вследствие ограниченности полосы пропускания реальных систем.

Если белый шум пропустить через идеальный фильтр низких частот с граничными частотами $(0, \omega_c)$, то на выходе получим шум с ограниченным спектром (рис. 3.8,а). Автокорреляционная функция такого сигнала принимает вид (рис. 3.8,б)

$$K_{XX}(\tau) = \frac{1}{\pi}\int_0^{\infty} G_0 \cos \omega\tau \, d\omega = \frac{1}{\pi\tau} G_0 \sin \omega\tau \Big|_0^{\omega_c} = \frac{\omega_c G_0}{\pi} \frac{\sin \omega_c \tau}{\omega_c \tau}.$$

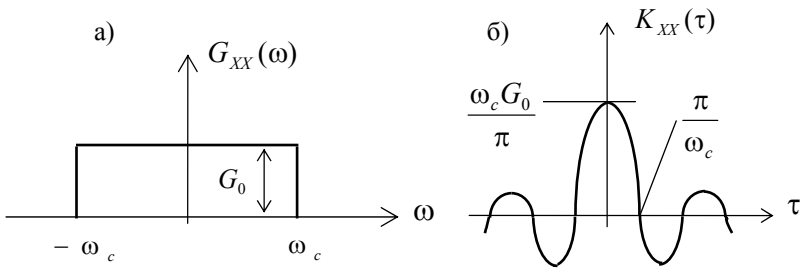


Рис. 3.8. Спектральные характеристики практического белого шума:
а – спектральная плотность мощности; б – автокорреляционная функция

3.8. СВОЙСТВА СПЕКТРАЛЬНОЙ ПЛОТНОСТИ МОЩНОСТИ СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Отметим некоторые свойства функции $G_{XX}(\omega)$:

1. Из определения $G_{XX}(\omega) = \lim\limits_{T\to\infty} \frac{1}{T} m\left[\left|S_k'(j\omega)\right|^2\right]$ следует, что

спектральная плотность мощности является неотрицательной функцией.

2. Дисперсия стационарного случайного процесса $X(t)$ с нулевым математическим ожиданием равна интегралу от его спектральной плотности мощности во всем диапазоне частот $D[X(t)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_{XX}(\omega)d\omega$. Это утверждение следует из известного свойства $D[\overset{\circ}{X}(t)] = \overset{\circ}{K}_{XX}(0)$.

3. Спектральная плотность мощности $G_{XX}(\omega)$ является четной функцией $G_{XX}(\omega) = G_{XX}(-\omega)$. Доказательством этого свойства является соотношение $G_{XX}(\omega) = 2\int_{0}^{\infty} K_{XX}(\tau)\cos\omega\tau d\tau$.

Таким образом, функция $G_{XX}(\omega)$ описывает распределение средней мощности по частотам, в связи с чем она носит название энергетического спектра или спектральной плотности мощности случайного сигнала.

Для случайных сигналов, реализации которых изображены на рис. 3.9, показаны характер корреляционной функции и энергетического спектра. Функции $x_2(t)$ и $x_3(t)$ при одинаковых математических ожиданиях и дисперсиях характеризуются различной степенью статистической связи в соседних сечениях. При сильной статистической связи $x_3(t)$ корреляционная функция затухает медленнее, а соответствующая ей кривая энергетического спектра спадает быстрее.

Функции $x_1(t)$ и $x_4(t)$ представляют собой предельные случаи:

– полностью коррелированного сигнала $x_1(t)$, т.е. детерминированного сигнала;

– совершенно не коррелированного сигнала $x_4(t)$, называемого белым шумом.

В первом случае корреляционная функция постоянна, а спектральная плотность мощности вырождается в дельта-функцию на нулевой частоте.

Во втором случае корреляционная функция отлична от нуля

только при $\tau = 0$, а спектральная плотность мощности имеет равномерный характер и неограниченную полосу.

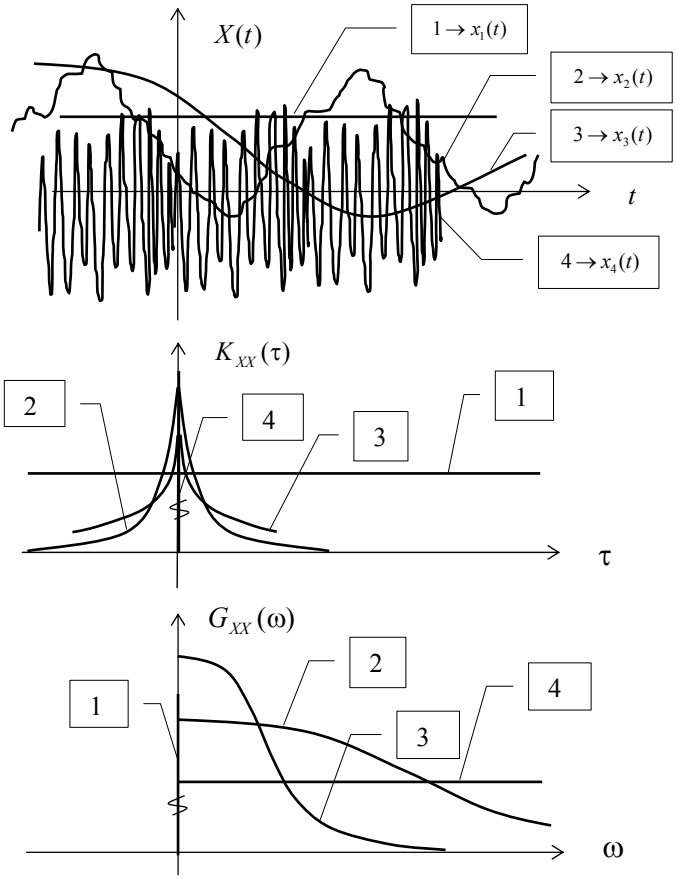


Рис. 3.9. Спектральные характеристики случайных процессов:
1 – постоянный уровень; 2 – слабо коррелированный стационарный процесс;
3 – сильно коррелированный стационарный процесс; 4 – белый шум

3.9. ИНТЕРВАЛ КОРРЕЛЯЦИИ И ЭФФЕКТИВНАЯ ШИРИНА СПЕКТРА СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Нормированная автокорреляционная функция $\overset{\circ}{r}_{XX}(\tau)$ стационарного случайного процесса показывает, как изменяется статистическая зависимость между двумя сечениями процесса при

изменении расстояния τ между ними. Из рис. 3.10 видно, что корреляционная связь между сечениями процесса, расположенными очень близко друг к другу, практически равна единице. По мере увеличения расстояния между сечениями, $\overset{\circ}{r}_{xx}(\tau)$ проявляет тенденцию к убыванию. Различные процессы имеют различные скорости убывания.

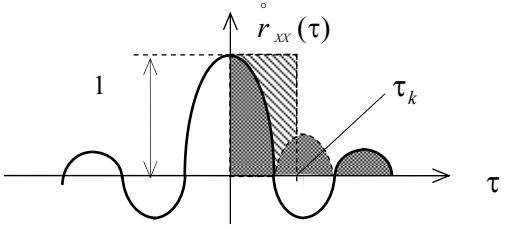


Рис. 3.10. Способ определения интервала корреляции τ_k

Для стационарного случайного процесса можно найти такое минимальное значение аргумента τ_k функции $\overset{\circ}{r}_{xx}(\tau)$, начиная с которого статистической зависимостью между сечениями можно пренебречь, т.е. считать $\overset{\circ}{r}_{xx}(\tau) = 0$ при всех $|\tau| > \tau_k$. Такое значение аргумента называют ***интервалом корреляции***. Чем меньше интервал корреляции случайного процесса, тем быстрее убывает корреляционная связь между сечениями.

Способ определения τ_k основан на построении прямоугольника единичной высоты с площадью, равной половине площади под графиком модуля $\overset{\circ}{r}_{XX}(\tau)$. Интервал корреляции τ_k определяется как основание такого прямоугольника:

$$\tau_k = \frac{1}{2}\int_{-\infty}^{\infty}\left|\overset{\circ}{r}_{XX}(\tau)\right|d\tau = \int_{0}^{\infty}\left|\overset{\circ}{r}_{XX}(\tau)\right|d\tau .$$

Кроме τ_k – числовой характеристики $\overset{\circ}{r}_{xx}(\tau)$, пользуются числовыми характеристиками спектральной плотности мощности стационарного случайного процесса.

Одной из таких характеристик является эффективная ширина спектра. ***Эффективной шириной спектра*** называют отношение площади под спектральной плотностью мощности стационарного случайного процесса к удвоенному максимуму спектральной плотности мощности

$$\Delta\omega = \frac{\int_{-\infty}^{\infty} G_{XX}(\omega)d\omega}{2G_{XX}(\omega_m)} = \frac{\int_{0}^{\infty} G_{XX}(\omega)d\omega}{G_{XX}(\omega_m)}.$$

Способ определения $\Delta\omega$ основан на построении прямоугольника (рис. 3.11) площадью $\int_0^{\infty} G_{XX}(\omega)d\omega$, высота которого равна максимуму спектральной плотности мощности $G_{XX}(\omega_m)$ на частоте ω_m.

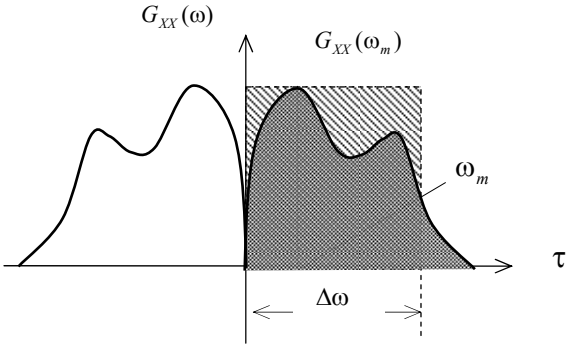


Рис. 3.11. Способ определения эффективной ширины спектра $\Delta\omega$

3.10. ЗАДАЧИ

3.1. Найти корреляционную функцию детерминированного периодического сигнала, представленного на рис. 3.12.

3.2. Найти корреляционную функцию периодического колебания произвольной формы, представленного на рис. 3.13.

3.3. Найти спектральную плотность средней мощности стационарного случайного процесса с корреляционной функцией, представленной на рис. 3.14.

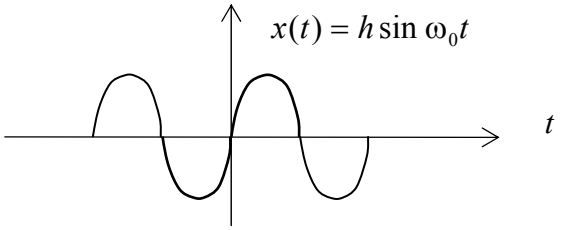


Рис. 3.12. Периодический сигнал

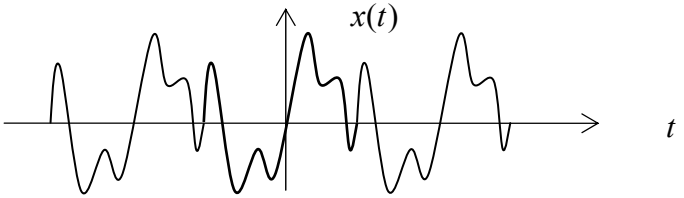


Рис. 3.13. Периодическое колебание

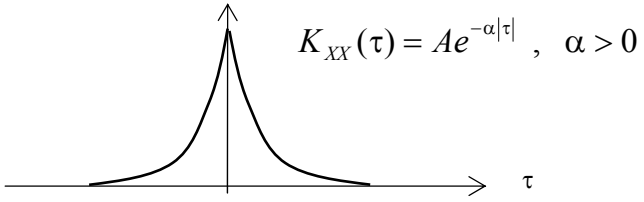


Рис. 3.14. Корреляционная функция

3.4. Найти спектральную плотность средней мощности стационарного случайного процесса с корреляционной функцией, представленной на рис. 3.15.

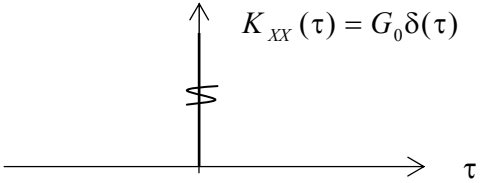


Рис. 3.15. Функция $K_{XX}(\tau)$

3.5. Найти корреляционную функцию стационарного случайного процесса со спектральной плотностью мощности, представленной на рис. 3.16.

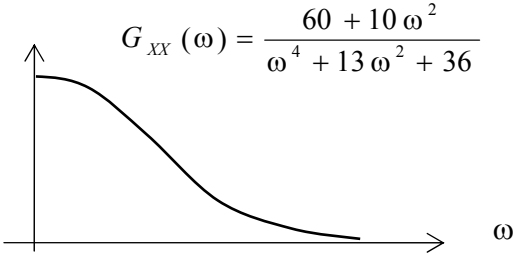


Рис. 3.16. Функция $G_{XX}(\omega)$

3.6. Необходимо найти дисперсию стационарного случайного процесса по заданной на рис. 3.16 спектральной плотности мощности.

3.7. Найти среднеквадратическое значение для случайного процесса со спектральной плотностью средней мощности в виде δ - функции (рис. 3.17).

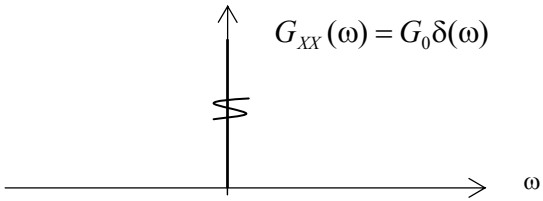


Рис. 3.17. Спектральная плотность средней мощности

3.8. Для стационарного случайного процесса со спектральной плотностью средней мощности (рис. 3.18) определить математическое ожидание $m[X]$ и дисперсию $D[X]$.

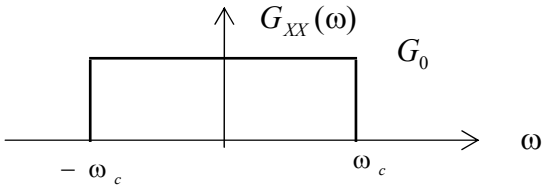


Рис. 3.18. Спектральная плотность $G_{XX}(\omega)$

3.9. Белый шум подается на вход электрической цепи, представленной на рис. 3.19. Найти спектральные характеристики выходного сигнала.

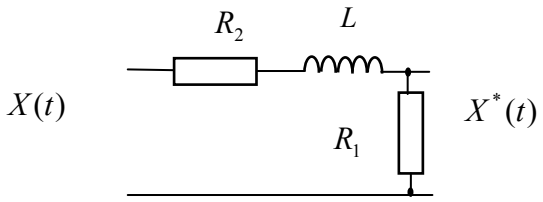


Рис. 3.19. Электрическая цепь

4. ЭЛЕМЕНТЫ ТЕОРИИ ДИСКРЕТИЗАЦИИ НЕПРЕРЫВНЫХ ФУНКЦИЙ

При передаче непрерывных сигналов в информационных системах весьма широкое применение получила кодоимпульсная модуляция сигналов (КИМ). КИМ складывается из трех операций:

– дискретизация по времени;

– квантование по уровню;

– кодирование.

Дискретизация по времени заключается в замене непрерывного во времени сигнала $x(t)$ дискретным сигналом $x_{д}(t)$, значения которого для дискретных моментов времени $t_0, t_1, ..., t_n$ совпадают с мгновенными значениями непрерывного сигнала (рис. 4.1).

Квантование по уровню заключается в замене непрерывного множества значений сигнала $x_{д}(t)$ множеством дискретных значений. При этом шкала возможных значений сигнала разбивается на определенное количество интервалов, и непрерывное значение сигнала заменяется ближайшим дискретным.

Полученные дискретные значения затем кодируются (обычно двоичным кодом). КИМ обеспечивает существенное повышение помехоустойчивости передачи данных. Кроме того, дискретизация позволяет использовать одни и те же устройства для большого числа различных сигналов.

При КИМ весьма важным является правильный выбор способа дискретизации сигналов по времени и квантования по уровню. Рассмотрим некоторые вопросы теории дискретизации непрерывных функций по времени.

4.1. ЧАСТОТНЫЙ КРИТЕРИЙ ДИСКРЕТИЗАЦИИ В.А. КОТЕЛЬНИКОВА

Итак, при дискретизации по времени непрерывная по аргументу функция $x(t)$ преобразуется в функцию $x_{д}(t)$ дискретного аргумента. Такое преобразование может быть выполнено путем взятия отсчетов функции $x(t)$ в определенные дискретные моменты вре-

мени. В результате функция $x(t)$ заменяется совокупностью мгновенных значений $x(t_i),\ i=\overline{0,n-1}$.

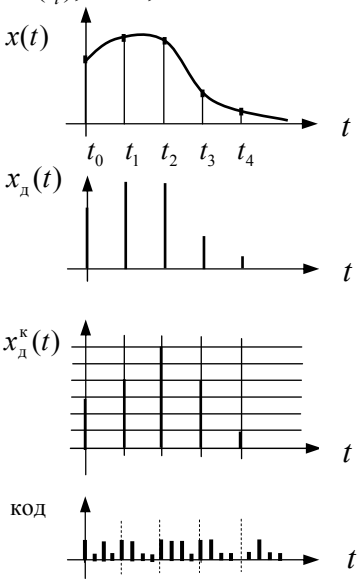


Рис. 4.1. Кодоимпульсная модуляция

Временной интервал $\Delta t = t_i - t_{i-1}, i=\overline{1,n-1}$ между двумя соседними фиксированными моментами времени, в которых задается дискретная функция, называется ***интервалом дискретизации***. Величина $f_д = \dfrac{1}{\Delta t}$ называется ***частотой дискретизации***. Устройства, с помощью которых проводится дискретизация сигналов, носят название дискретизаторов (рис. 4.2).

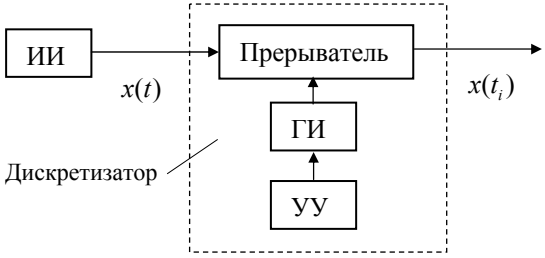


Рис. 4.2. Структура дискретизатора

Частота дискретизации должна выбираться таким образом, чтобы по отсчетным значениям $x(t_i)$ можно было бы с заданной точностью получить исходную функцию.

Известно несколько критериев выбора частоты $f_д$ дискретизации по времени. Рассмотрим ***частотный критерий*** В.А. Котельникова. Данный критерий, который получил название теоремы Котельникова, основывается на следующей модели сигналов:

– сигнал представляет собой стационарный случайный процесс;

– спектр реализации сигнала сплошной и ограничен некоторой частотой, за пределами которой он тождественно равен нулю.

ТЕОРЕМА 4.1 (***теорема Котельникова***). Если непрерывная функция $x(t)$ удовлетворяет условиям Дирихле (ограничена, кусочно-непрерывна и имеет конечное число экстремумов), и ее спектр ограничен некоторой частотой $f_c = \dfrac{\omega_c}{2\pi}$, то она полностью определяется последовательностью своих значений в точках, отстающих на расстоянии $\Delta t = \dfrac{1}{2f_c} = \dfrac{\pi}{\omega_c}$ друг от друга.

ДОКАЗАТЕЛЬСТВО. Пусть сигнал, описываемый непрерывной функцией времени $x(t)$, имеет ограниченный спектр ω_c (рис. 4.3), т.е. преобразование Фурье $S(j\omega) = \int\limits_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$ удовлетворяет условию $S(j\omega) = 0$ при $|\omega| > \omega_c$.

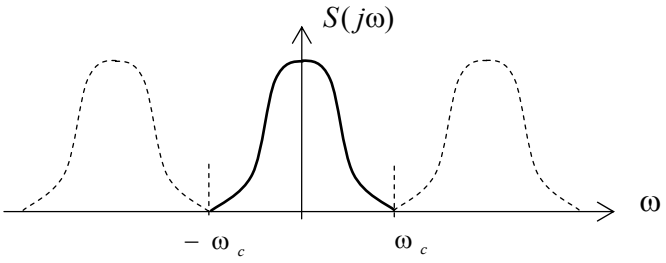


Рис. 4.3. Спектральная плотность

В представлении сигнала $x(t)$ интегралом Фурье пределы интегрирования можно ограничить интервалом $[-\omega_c, \omega_c]$

$$x(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} S(j\omega) e^{j\omega t} d\omega . \qquad (4.1)$$

Дополним функцию $S(j\omega)$ до периодической с периодом $2\omega_c$ и разложим ее в ряд Фурье на интервале $[-\omega_c, \omega_c]$ ($\omega_0 = 2\pi/T \to \pi/\omega_c$):

$$S(j\omega) = \sum_{k=-\infty}^{\infty} C_k e^{jk\frac{\pi}{\omega_c}\omega} , \qquad (4.2)$$

где коэффициент Фурье имеет вид

$$C_k = \frac{1}{2\omega_c} \int_{-\omega_c}^{\omega_c} S(j\omega) e^{-jk\frac{\pi}{\omega_c}\omega} d\omega . \qquad (4.3)$$

Сравнивая (4.1) и (4.3), замечаем, что они совпадают с точностью до постоянного множителя $\Delta t = \frac{\pi}{\omega_c}$, если принять $t = -k\frac{\pi}{\omega_c}$.

Следовательно, $C_k = \frac{\pi}{\omega_c} x(-k\Delta t)$. Подставив найденное выражение для C_k в (4.2), получаем

$$S(j\omega) = \sum_{k=-\infty}^{\infty} \frac{\pi}{\omega_c} x(-k\Delta t) e^{jk\frac{\pi}{\omega_c}\omega} . \qquad (4.4)$$

После подстановки (4.4) в (4.1), замены знака при k (так как суммирование производится по всем положительным и отрицательным значениям k) и перестановки операций суммирования и интегрирования получим

$$x(t) = \frac{1}{2\omega_c} \sum_{k=-\infty}^{\infty} x(k\Delta t) \int_{-\omega_c}^{\omega_c} e^{j\omega(t-k\Delta t)} d\omega . \qquad (4.5)$$

Вычислим интеграл

$$\int_{-\omega_c}^{\omega_c} e^{j\omega(t-k\Delta t)} d\omega = \int_{-\omega_c}^{\omega_c} \cos[\omega(t-k\Delta t)] d\omega + j \int_{-\omega_c}^{\omega_c} \sin[\omega(t-k\Delta t)] d\omega = \\ = \frac{2\sin[\omega_c(t-k\Delta t)]}{t-k\Delta t} \tag{4.6}$$

с учетом равенства нулю второго интеграла разложения.

После подстановки (4.6) в (4.5) окончательно имеем

$$x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t) \frac{\sin[\omega_c(t-k\Delta t)]}{\omega_c(t-k\Delta t)}. \tag{4.7}$$

Этим, собственно, и доказывается теорема отсчетов.

Таким образом, выражение (4.7) показывает, что реализация $x(t)$ полностью определяется совокупностью отсчетов, взятых в моменты времени $k\Delta t = \frac{k}{2f_c}$ и отстоящих друг от друга на величину $\Delta t = \frac{1}{2f_c}$.

Ряд (4.7) в технической литературе получил название ***ряда Котельникова***.

4.2. ПРЕДСТАВЛЕНИЕ СИГНАЛОВ С ОГРАНИЧЕННОЙ ЧАСТОТНОЙ ПОЛОСОЙ В ВИДЕ РЯДА КОТЕЛЬНИКОВА

Остается решить вопрос, каким образом восстановить функцию $x(t)$ по ее дискретным отсчетам $x(k\Delta t)$, если в этом возникает необходимость.

Из (4.7) видно, что непрерывная функция, обладающая ограниченным спектром, может быть представлена разложением в ряд, каждый член которого выражается одинаковой функцией вида $\mathrm{sinc}\,[\omega_c(t-k\Delta t)]$, но с различными коэффициентами $x(k\Delta t)$.

Другими словами, ряд Котельникова представляет собой разложение реализации случайного процесса ***координатными детерминированными функциями времени*** $A(t_k) = \mathrm{sinc}\,[\omega_c(t-k\Delta t)]$ с весовыми коэффициентами $x(k\Delta t)$ равными мгновенным значениям сигнала в точках $k\Delta t$.

Как известно, функция $\operatorname{sinc} x(t)$ представляет собой реакцию фильтра нижних частот (ФНЧ) с граничной частотой ω_c на дельта-импульс.

Следовательно, если в приемном устройстве поместить такой фильтр и пропустить через него квантованный сигнал $x_{\text{д}}^{\text{к}}(t)$, представляющий собой последовательность с частотой $2f_c = \dfrac{\omega_c}{\pi}$ кратковременных импульсов, амплитуды которых пропорциональны отсчетам исходной непрерывной функции, то, суммируя выходные сигналы ФНЧ, можно воспроизвести с высокой степенью точности исходный непрерывный сигнал (рис. 4.4).

Практически реализовать это достаточно трудно. Кроме того, на практике приходится иметь дело с сигналами, ограниченными во времени и, следовательно, обладающими бесконечно широким спектром, что противоречит основному условию теоремы Котельникова.

Однако на практике не требуется идеально точного воспроизведения передаваемого сигнала. Поэтому с целью использования теоремы для дискретизации сигналов реальный спектр сигнала, простирающийся от 0 до ∞, условно ограничивают некоторым диапазоном частот от 0 до ω_c, в котором сосредоточена основная часть энергии спектра. Энергия отсекаемой части спектра сигнала определяет погрешность

$$\varepsilon_{\omega_c} = \frac{\int\limits_{\omega_c}^{\infty} [S(\omega)]^2 d\omega}{\int\limits_{0}^{\infty} [S(\omega)]^2 d\omega} = 1 - \lambda(\omega_c),$$

возникающую за счет ограничений спектра сигнала.

Кроме того, известна математическая модель сигнала Н.А. Железнова, в которой принимается, что спектр сигнала конечной длительности отличен от нуля на всей оси частот и известна автокорреляционная функция $K_{XX}(\tau)$. В этом случае шаг дискретизации выбирается равным интервалу корреляции τ_k.

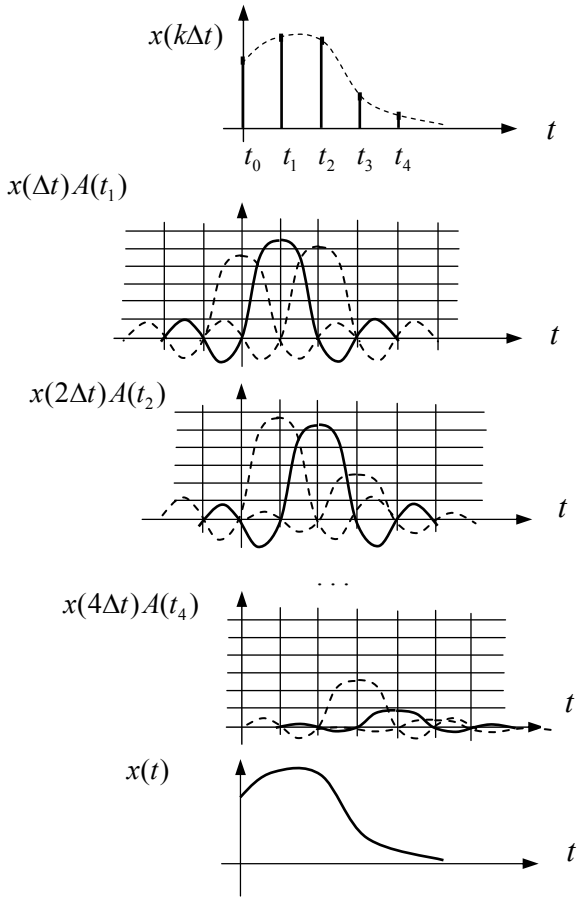


Рис. 4.4. Восстановление сигнала по дискретным отсчетам

Корреляционный критерий выбора отсчетов является разновидностью частотного критерия. А.А. Харкевич показал, что $\Delta t = \tau_k \approx \dfrac{\pi}{\omega_c}$, то есть отсчеты по теореме Котельникова представляют собой ближайшие некоррелированные значения сигнала.

Таким образом, несмотря на то, что теорема Котельникова базируется на модели сигналов с ограниченным спектром, она имеет большую теоретическую и практическую ценность в технике преобразований сигналов, их передачи и обработки.

4.3. ДИСКРЕТНЫЕ СИГНАЛЫ И ИХ СПЕКТРЫ

При вычислении спектра на ЭВМ анализируемая функция $x(t)$ представляется в виде своих отсчетов в дискретные моменты времени $x(nT)$, где T – интервал дискретизации по времени (ранее обозначался Δt), а $n = 0,\pm1,\pm2,...$.

Спектральный анализ, основанный на цифровой обработке сигналов, предполагает преобразование последовательности $x(nT)$ в последовательность чисел $A_k = A(jk\Omega)$, где Ω – интервал дискретизации частоты, а $k = 0,\pm1,\pm2,...$ [18].

Определим спектр дискретизированной периодической функции, представленной на рис. 4.5. Положим, что ее период T_0 содержит ровно N отсчетов:

$$x(nT):\ x(0), x(T), x(2T),..., x[(N-1)T].$$

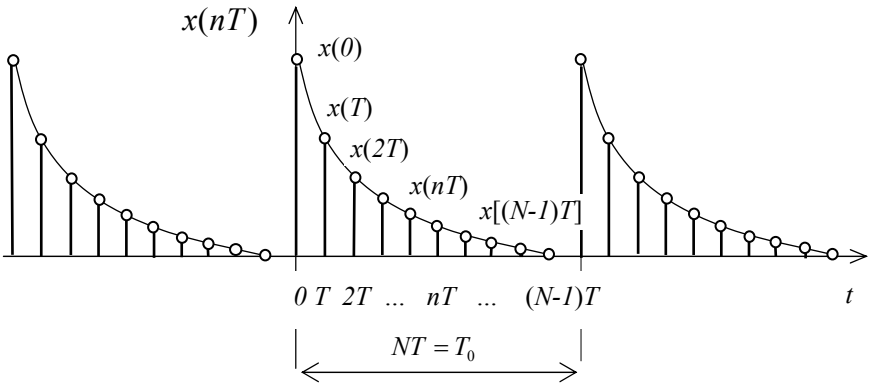


Рис. 4.5. Дискретный периодический сигнал

Тогда $x(t)$ можно разложить в ряд

$$x(t) = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{jk\Omega t},$$

где $\Omega = 2\pi / NT$.

В моменты отсчетов nT имеем

$$x(nT) = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{jk\frac{2\pi}{NT}nT} = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{jk\frac{2\pi}{N}n}. \qquad (4.8)$$

Правая часть не зависит от T. Поэтому вместо $x(nT)$ можно писать $x(n)$.

Но $e^{jk\frac{2\pi}{N}n} = e^{j(k\pm sN)\frac{2\pi}{N}n}$, где s – любое целое положительное число.

Следовательно, можно записать:

$$x(nT) = \frac{1}{2}\sum_{k=-\infty}^{\infty} A_k e^{j(k\pm sN)\frac{2\pi}{N}n}. \qquad (4.9)$$

Сравнивая (4.8) и (4.9), приходим к выводу, что комплексные амплитуды при дискретизированных гармониках, кратных N, не отличаются друг от друга, и имеется только N различных составляющих дискретного спектра $A_0, A_1, A_2, ..., A_{N-1}$. Таким образом, получено важное свойство периодических дискретизированных во времени сигналов: их спектр также является периодическим и дискретным с периодом, равным N отсчетам частоты ($N\Omega$).

Поэтому бесконечные пределы в (4.8) можно заменить одним периодом:

$$x(nT) = \frac{1}{2}\sum_{k=0}^{N-1} A_k e^{jk\frac{2\pi}{N}n}. \qquad (4.10)$$

Для определения коэффициентов A_k умножим обе части этого равенства на $e^{-jm\frac{2\pi}{N}n}$, где m – целое число в пределах от 0 до $N-1$, и рассмотрим сумму по переменной n, содержащую N таких произведений:

$$\sum_{n=0}^{N-1} x(nT) e^{-jm\frac{2\pi}{N}n} = \frac{1}{2}\sum_{k=0}^{N-1} A_k \sum_{n=0}^{N-1} e^{j(k-m)\frac{2\pi}{N}n}.$$

Если $k = m$, то $e^{j(k-m)\frac{2\pi}{N}n} = 1$ независимо от величины n,

при этом

$$\sum_{n=0}^{N-1} e^{j(k-m)\frac{2\pi}{N}n} = N .$$

При $k \neq m$ комплексная экспонента описывает единичный вектор, фаза которого при изменении величины n изменяется скачками на равноотстоящие углы $\frac{2\pi}{N}(k-m)$. Сумма этих векторов за N отсчетов равна нулю.

Учитывая это, получаем

$$\frac{1}{2}\sum_{k=0}^{N-1} A_k \sum_{n=0}^{N-1} e^{j(k-m)\frac{2\pi}{N}n} = \frac{1}{2} NA_m .$$

Следовательно,

$$A_m = \frac{2}{N}\sum_{n=0}^{N-1} x(nT)e^{-jm\frac{2\pi}{N}n} . \qquad (4.11)$$

Объединяя (4.10) и (4.11) и заменяя в A_m индекс m на индекс k, окончательно получаем:

$$\begin{cases} A_k = \frac{2}{N}\sum_{n=0}^{N-1} x(nT)e^{-jk\frac{2\pi}{N}n}; \\ x(nT) = \frac{1}{2}\sum_{k=0}^{N-1} A_k e^{jk\frac{2\pi}{N}n}. \end{cases} \qquad (4.12)$$

Эти соотношения называются ***прямым и обратным дискретными преобразованиями Фурье*** (ДПФ). ДПФ A_k, как и сама последовательность $x(nT)$, является периодической функцией по аргументу k с периодом N.

Характер зависимости модуля функции A_k от частоты приведен на рис. 4.6.

Период спектра в масштабе частот равен $N\Omega = \frac{2\pi}{T}$. Заметим, что в пределах периода составляющие спектра симметричны относительно середины этого интервала, причем $|A_k| = |A_{N-k}|$ (это

следует, например, из того факта, что $|A_k| = |A_{-k}|$ и $|A_{-k}| = |A_{N-k}|, k = 1,2,3,...$) и $\arg A_k = -\arg A_{N-k}$. Поэтому вся полезная информация имеется уже в интервале частот, равном $\frac{\pi}{T}$.

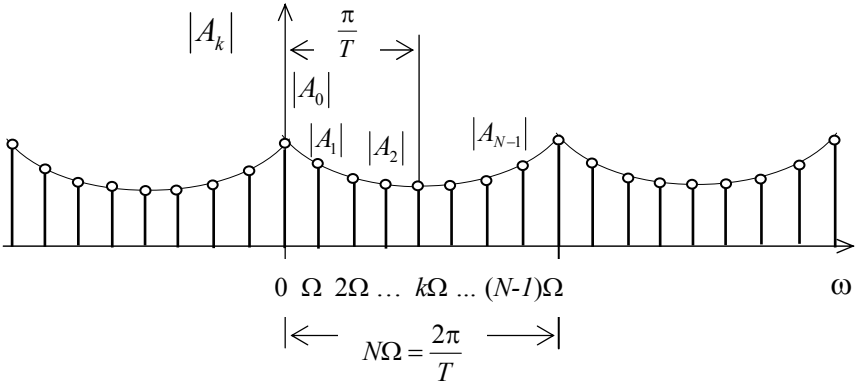


Рис. 4.6. Спектр амплитуд дискретного периодического сигнала

Рассмотрим непериодический аналоговый сигнал $x(t)$, заданный на интервале T_0. Спектр сигнала $x(t)$ получается в результате прямого преобразования Фурье

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

Сам сигнал по его спектру вычисляется с помощью обратного преобразования Фурье

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

Пусть дискретный сигнал $x(nT)$ конечной длины N, определенный при $n = 0,1,2,..., N-1$ и равный нулю вне интервала $[0, N-1]$, получен дискретизацией непрерывного сигнала $x(t)$ с интервалом T (рис. 4.7).

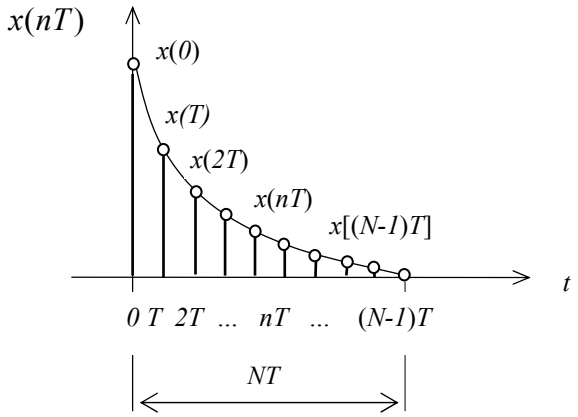


Рис. 4.7. Дискретный непериодический сигнал

Пара преобразований Фурье для дискретного сигнала имеет вид

$$S(e^{j\omega T}) = \sum_{n=0}^{N-1} x(nT)e^{-j\omega nT}, \qquad (4.13)$$

$$x(nT) = \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} S(e^{j\omega T})e^{j\omega nT}\,d\omega, \qquad (4.14)$$

где $S(e^{j\omega T})$ – спектр дискретного сигнала. Спектры $S(e^{j\omega T})$ и $S(j\omega)$ связаны соотношением

$$S(e^{j\omega T}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} S\left[j(\omega + k\omega_D)\right],$$

т.е. спектр последовательности $x(nT)$ равен с точностью до множителя $\frac{1}{T}$ сумме спектров соответствующего сигнала $x(t)$, смещенных по оси частот на все возможные значения частоты, кратные частоте дискретизации $\omega_D = 2\pi/T$.

$S(e^{j\omega T})$ является периодической функцией по частоте ω с периодом, равным частоте дискретизации $\omega_D = 2\pi/T$. Действительно,

$$S(e^{j\omega T}) = S(e^{j(\omega + k\cdot 2\pi/T)T}), \quad k = 1,2,\ldots .$$

Именно поэтому при вычислении $x(nT)$ интегрирование проводится в пределах $[-\pi/T, \pi/T]$. Очевидно, что свойство периодичности спектра распространяется также на его модуль $\left|S(e^{j\omega T})\right|$ и аргумент $\arg S(e^{j\omega T})$.

Предположим, что спектр $S(j\omega)$ непрерывного сигнала ограничен частотой ω_c. Тогда, для того чтобы спектр дискретизированного сигнала содержал наибольшее количество информации об исходном сигнале $x(t)$, необходимо выполнить соотношение

$$\omega_c \le \frac{\pi}{T},$$

т.е. условие теоремы Котельникова. В этом случае спектр $\left|S(e^{j\omega T})\right|$ повторяет по форме спектр исходного сигнала (рис. 4.8), только в отличие от последнего он имеет периодическую структуру.

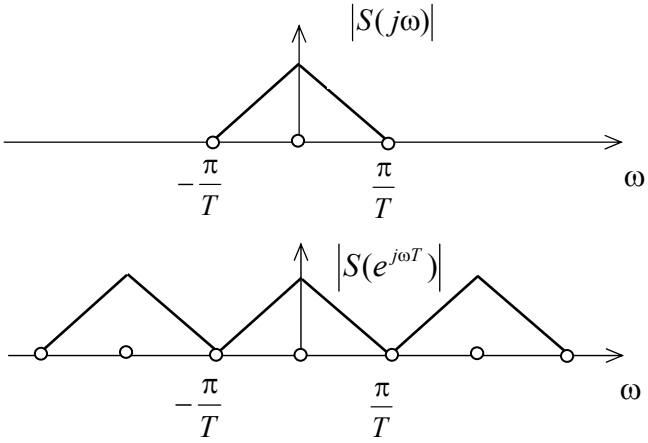


Рис. 4.8. Спектр дискретного непериодического сигнала

Если же $\omega_c > \frac{\pi}{T}$, происходит искажение $x(t)$ вследствие вза-

имного наложения спектральных составляющих соседних участков (рис. 4.9).

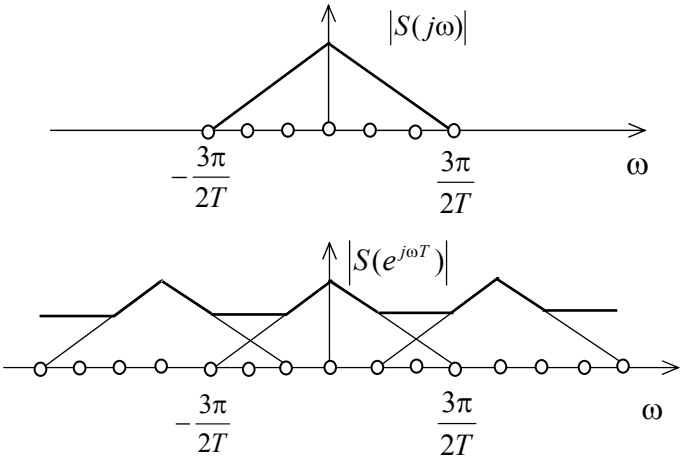


Рис. 4.9. Взаимное наложение спектральных составляющих

Таким образом, приступая к анализу, следует предварительно выбрать интервал дискретизации T в соответствии с критерием Котельникова.

Если сравнить спектр конечного дискретного сигнала, выражаемый формулой $S(e^{j\omega T})$, с ДПФ этого сигнала, то увидим, что ДПФ представляет собой N отсчетов спектра, взятых на периоде с интервалом дискретизации по частоте, равным $\Omega = \frac{2\pi}{NT}$. Из этого следует, что свойства ДПФ аналогичны свойствам спектров.

Следовательно, дискретное преобразование Фурье может быть использовано и для представления последовательности $x(nT)$ конечной длины N, определенной при $n = 0,1,2,..., N-1$ и равной нулю вне интервала $[0, N-1]$. Действительно, такую последовательность можно рассматривать как один период соответствующей периодической последовательности и использовать преобразования ДПФ; следует только считать, что вне интервала $[0, N-1]$ A_k и $x(nT)$ равны нулю.

Преобразования (4.13) и (4.14) в большей степени носят исключительно теоретический интерес.

4.4. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Для ДПФ (4.12), в принципе, справедливы все свойства интегральных преобразований Фурье, однако при этом следует учитывать периодичность дискретных функций и спектров.

Из выражений ДПФ можно видеть, что для вычисления каждой гармоники нужно N операций комплексного умножения и сложения и соответственно N^2 операций на полное выполнение ДПФ. При больших объемах массивов данных это может приводить к существенным временным затратам. Ускорение вычислений достигается при использовании ***быстрого преобразования Фурье*** (БПФ).

БПФ базируется на том, что при вычислениях среди множителей (синусов и косинусов) есть много периодически повторяющихся значений (в силу периодичности функций). Алгоритм БПФ группирует слагаемые с одинаковыми множителями в пирамидальный алгоритм, значительно сокращая число умножений за счет исключения повторных вычислений. В результате быстродействие БПФ в зависимости от N может в сотни раз превосходить быстродействие стандартного алгоритма. При этом следует подчеркнуть, что алгоритм БПФ даже точнее стандартного, так как сокращая число операций, он приводит к меньшим ошибкам округления.

Алгоритмы БПФ основываются на свойствах комплексной экспоненты $e^{-j(2\pi/N)kn} = W_N^{kn}$: ее симметрии $W_N^{kn} = W_N^{-(N-k)n} = W_N^{-(N-n)k}$ и периодичности $W_N^{kn} = W_N^{(k+lN)(n+mN)}$ с периодом, равным длине обрабатываемой реализации сигнала N (числу точек БПФ). С учетом последнего свойства, экспоненте $W_N^{pkn} = W_{N/p}^{kn}$ соответствует период N/p, где p – целые числа, на которые делится N. Использование данных свойств в алгоритмах БПФ исключает большое число повторяющихся при вычислении ДПФ операций.

Алгоритм БПФ заключается в разбиении ДПФ исходной последовательности на ДПФ подпоследовательностей меньшей длины,

вплоть до минимально возможной (равной основанию БПФ), через которые и вычисляется ДПФ исходной последовательности.

Разбиение означает прореживание последовательностей во временной или в частотной области. В связи с этим различают БПФ с прореживанием по времени и БПФ с прореживанием по частоте.

В отличие от ДПФ, БПФ может вычисляться только по определенному числу точек $N = m^L$, соответствующему целой степени L его основания m. Целая степень L определяет число этапов прореживания. К наиболее используемым относятся БПФ по основаниям $m = 2,4,8$.

В качестве примера рассмотрим алгоритм БПФ с прореживанием по времени по основанию 2.

Пусть задана последовательность $x(nT) = x(n)_N$ конечной длины $n = 0,1,...N-1$. Нужно найти ее ДПФ: $X(jk) = \sum_{n=0}^{N-1} x(n) W_N^{kn}$ для $k = 0,1,...N-1$ (номера бинов ДПФ) с минимальным объемом вычислений.

Решение этой задачи в данном алгоритме БПФ находится следующим образом.

Исходную последовательность $x(n)$ длиной N разобьем на две подпоследовательности длиной $N/2$ (рис. 4.10) – четную (включающую отсчеты $x(n)$ с четными индексами $n : x_1(n) = x(2n)$) и нечетную: $x_2(n) = x(2n+1)$, $n = 0,1,...N-1$. Это соответствует первому прореживанию сигнала по времени.

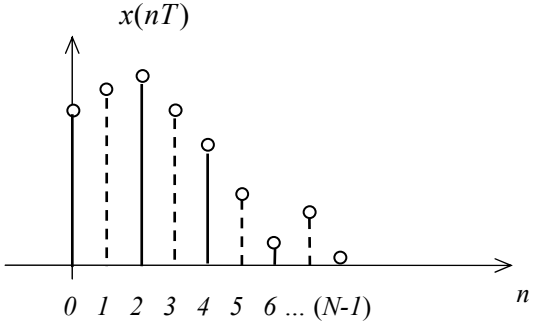


Рис. 4.10. Иллюстрация прореживания сигнала по времени

Обозначим их ДПФ, как $X_1(jk)$ и $X_2(jk)$. Выразим ДПФ исходной последовательности $x(nT) = x(n)_N$ через ДПФ подпоследовательностей $x_1(n)_{N/2}$, $x_2(n)_{N/2}$:

$$X(jk) = \sum_{n=0}^{N/2-1} x_1(n) e^{-j\frac{2\pi}{N/2}kn} + \sum_{n=0}^{N/2-1} x_2(n) e^{-j\frac{2\pi}{N/2}kn} e^{-j\frac{2\pi}{N}k} = \\ = X_1(jk) + W_N^k X_2(jk), \quad k = 0,1,..., \frac{N}{2} - 1 \ . \tag{4.15}$$

Это первые $N/2$ частотных выборок ДПФ.

Вторую половину частотных выборок $X(jk)$ для $k = N/2,..., N-1$ найдем с учетом свойства периодичности:

$$X\left[j(k + \frac{N}{2}) \right] = X_1(jk) + W_N^{(k+N/2)} X_2(jk) = \\ = X_1(jk) - W_N^k X_2(jk), \quad k = 0,1,..., \frac{N}{2} - 1 \ . \tag{4.16}$$

Выражения (4.15) и (4.16) определяют базовую операцию БПФ (операцию объединения):

$$X(jk) = X_1(jk) + W_N^k X_2(jk), \\ X(j(k + N/2)) = X_1(jk) - W_N^k X_2(jk), \ k = 0,1,..., \frac{N}{2} - 1. \tag{4.17}$$

Входящий в (4.17) множитель W_N^k, равный по модулю единице, называют поворачивающим. Вычисления в соответствии с (4.17) включают одно комплексное умножение и пару сложения – вычитания. Базовую операцию представляют графически с помощью сигнального графа (бабочки БПФ) (рис. 4.11).

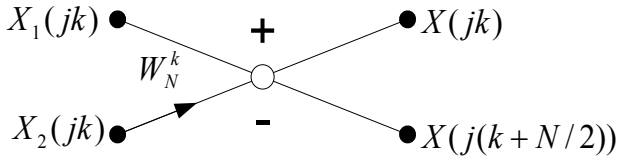


Рис. 4.11. Сигнальный граф базовой операции БПФ

На нем узел означает операцию сложения (верхний выход) и вычитания (нижний выход), а стрелка $\longrightarrow$ соответствует умножению на поворачивающий множитель W_N^k.

Дальше каждую из последовательностей $x_1(n)$ и $x_2(n)$ можно разбить еще на две подпоследовательности вдвое меньшей длины: $x_{11}(n), x_{12}(n)$ и $x_{21}(n), x_{22}(n)$ (четную и нечетную) и повторить вышеприведенные операции объединения их ДПФ с помощью базовых операций.

Такое прореживание выполняем L раз до получения $N/2$ двухточечных последовательностей, ДПФ которых вычисляется в соответствии с базовой операцией (4.17).

На рис. 4.12 в качестве примера приведен полный граф БПФ для $N = 8$. В соответствии с графом на каждом из L этапов вычисления – объединения ДПФ выполняются $N/2$ базовых операций, а общий объем вычислений для комплексных операций умножения и сложения – вычитания составляет

$$K_{умн} = (NL)/2 = (N \log_2 N)/2,$$

$$K_{слож} = NL = N \log_2 N.$$

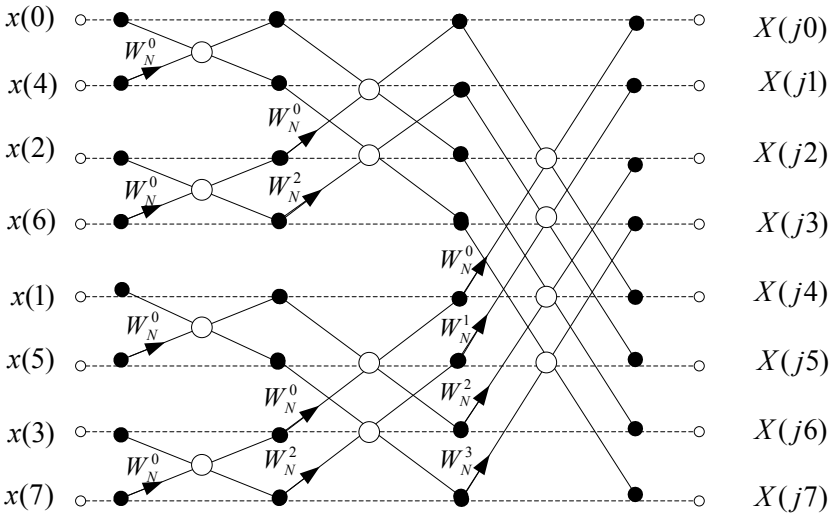


Рис. 4.12. Полный граф БПФ для $N = 8$

Особенностью алгоритма БПФ с прореживанием по времени является требуемый им неестественный порядок отсчетов входного сигнала, обусловленный его многократными разбиениями на четные и нечетные подпоследовательности (n = 0, 4, 2, 6, 1, 5, 3, 7 для $N = 8$). Такой порядок следования называют двоично-инверсным. Это приводит к необходимости предварительной перестановки отсчетов исходной последовательности до начала вычислений. Для этого естественные номера отсчетов последовательности $x(n)$ представляются в L - разрядном двоичном коде, коды эти прочитываются в обратном порядке, т.е. справа налево и преобразуются затем снова в десятичную форму, соответствующую номеру отсчета переставленной последовательности.

ЗАДАЧИ

4.1. Для сигнала, представленного на рис. 2.19, найти и обосновать практическую ширину спектра.

4.2. Для сигнала, представленного на рис. 2.19, вычислить величину интервала дискретизации в соответствии с теоремой Котельникова.

4.3. Для экспоненциального сигнала отобразить координатные детерминированные функции времени, входящие в ряд Котельникова.

4.4. Определить интервал частот, в пределах которого сосредоточена вся полезная информация спектра амплитуд дискретного периодического сигнала.

5. ЭЛЕМЕНТЫ ТЕОРИИ ОПТИМАЛЬНОГО ПРИЕМА И СТАТИСТИЧЕСКИХ РЕШЕНИЙ

Следует отметить, что общая схема передачи информации, представленная во введении, соответствует случаю, когда информация вводится в начале канала связи, т.е. непосредственно в передатчике. Несколько иначе обстоит дело, например, в радиолокационном канале, где информация о цели вводится в результате отражения радиоволны от цели в свободном пространстве.

Проблема оптимального приема состоит в рациональном использовании избыточности, а также имеющихся сведений о свойствах полезного сигнала, помехи и канала для увеличения вероятности правильного приема.

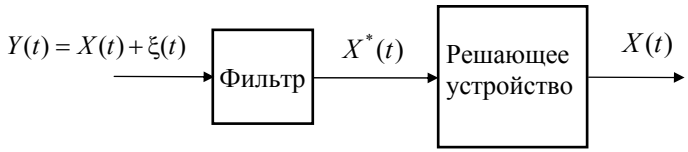
При этом рассматриваются следующие задачи:

1. Обнаружение сигнала, когда требуется только дать ответ, имеется ли в принятом колебании полезный сигнал или оно обусловлено одним шумом.
2. Оценка параметров, когда требуется с наибольшей точностью определить значение одного или нескольких параметров полезного сигнала.
3. Различение сигналов, когда на входе возможно присутствие нескольких сигналов, и нужно указать, какие именно сигналы присутствуют.
4. Воспроизведение первоначальной формы сигнала, искаженной действием шумов.
5. Предсказание сигнала, когда следует, располагая историей сигнала, предсказать его наиболее вероятные значения в будущем.

Результатом воздействия помех является частичная или полная потеря информации, переносимой полезным сигналом. Приемное устройство, осуществляя обработку входного сигнала, должно обеспечить извлечение из принятого сигнала возможно большего количества необходимой информации.

Вследствие того, что на вход приемника поступает сумма полезного сигнала и помехи, вероятность правильного приема будет определяться отношением полезного сигнала к помехе. Для повышения вероятности правильного приема должна быть произведена

инятого сигнала по схеме, представленной на рис. 5.1. Фильтр обеспечивает улучшение отношения мощностей сигнал/помеха $\rho = \left(\frac{p_x}{p_\xi} \right)_{\text{вых}}$, а решающее устройство выполняет главные функции приема (обнаружение, различение или восстановление сигналов).



Используемые характеристики : $G_{XX}(\omega)$, $G_{\xi\xi}(\omega)$, $P(a_i)$, $f(X/a_i)$, $K_{XX}(\tau)$, $K_{\xi\xi}(\tau)$, r_{12}, r_{21}.

Рис. 5.1. Блок-схема предварительной обработки принятого сигнала

Рассмотрим основные задачи оптимального приема и статистических решений.

5.1. МЕТОДЫ ФИЛЬТРАЦИИ

Известны следующие методы фильтрации, обеспечивающие улучшение отношения сигнал/помеха:

– частотная фильтрация;
– метод накопления;
– корреляционный метод;
– согласованная фильтрация.

Все эти методы основаны на использовании различных свойств полезного сигнала и помехи.

5.1.1. Частотная фильтрация

Идея частотной фильтрации основана на отличии спектров полезного сигнала и помехи. При этом используются линейные частотные фильтры, позволяющие подавлять помеху и улучшать тем самым отношение сигнал/помеха.

Параметры фильтра определяются спектральными характеристиками сигнала и помехи. На практике наиболее часто встречаются следующие случаи (рис. 5.2):

а) на вход приемного устройства поступает узкополосный сигнал и широкополосная помеха. В этом случае в тракт приемного устройства включается узкополосный фильтр с полосой пропускания $\Delta\omega_X$;

б) на вход приемника поступает широкополосный сигнал и узкополосная помеха. В таких случаях в тракт приемника включается фильтр, обеспечивающий подавление помехи в полосе $\Delta\omega_\xi$.

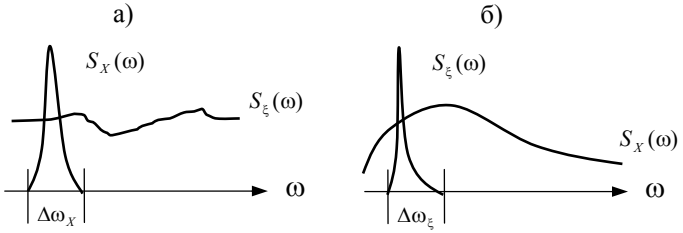


Рис. 5.2. Иллюстрация частотной фильтрации:
а – узкополосный сигнал и широкополосная помеха; б – широкополосный сигнал и узкополосная помеха

5.1.2. Метод накопления

Метод применим в том случае, если полезный сигнал в течение времени приема постоянен.

Пусть в течение времени передачи T_X сигнал постоянен и равен a. На сигнал воздействует аддитивная помеха с $m[\xi(t)] = 0$. Предположим, что в течение T_X (т.е. на интервале, равном длительности передаваемого импульса) осуществляется многократное его измерение (рис. 5.3). Покажем, что суммирование этих отсчетов позволяет увеличить отношение сигнал/помеха $\rho = \left(\frac{p_x}{p_\xi}\right)_{\text{вых}}$.

Рассмотрим случай, когда указанные отсчеты не коррелированны, т.е. берутся через интервалы времени $\Delta t \geq \tau_k$, где τ_k – интервал корреляции.

На выходе накопителя будет сигнал

$$y = \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} (a + \xi_i) = na + \sum_{i=1}^{n} \xi_i ,$$

где n – количество отсчетов.

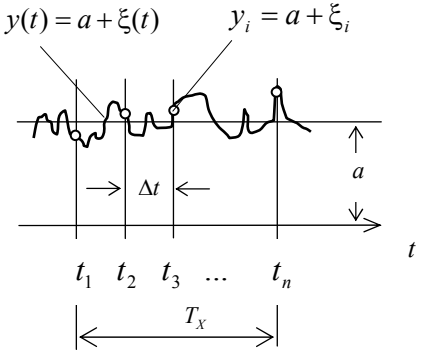


Рис. 5.3. Метод накопления

Отношение мощностей сигнала и помехи на выходе накопителя

$$\rho = \left(\frac{p_x}{p_\xi} \right)_{\text{вых}} = \frac{(na)^2}{D[\sum_{i=1}^{n} \xi_i]},$$

где $D[\sum_{i=1}^{n} \xi_i]$ – дисперсия помехи на выходе накопителя.

В предположении, что значения ξ_i некоррелированы и помеха представляет собой стационарный случайный процесс, дисперсия суммы отсчетов ξ_i равна сумме дисперсий отсчетов

$$D[\sum_{i=1}^{n} \xi_i] = \sum_{i=1}^{n} D[\xi_i] = nD[\xi] .$$

Следовательно,

$$\rho = \left(\frac{p_x}{p_\xi} \right)_{\text{вых}} = \frac{(na)^2}{nD[\xi]} = n \frac{a^2}{D[\xi]} = n \left(\frac{p_x}{p_\xi} \right)_{\text{вх}} .$$

Таким образом, в результате n - кратного отсчета, отношение мощностей сигнала и помехи увеличивается в n раз.

5.1.3. Корреляционный метод

Сущность метода заключается в использовании различия между корреляционными функциями сигнала и помехи. Метод эффективен лишь в случае приема периодических сигналов.

Пусть полезный сигнал является периодическим, а помеха – типа гауссова шума.

В приемном устройстве определяется корреляционная функция поступающей на вход суммы $y(t) = x(t) + \xi(t)$ полезного сигнала и помехи

$$K_{YY}(\tau) = \frac{1}{T}\int_{-T/2}^{T/2}[x(t)+\xi(t)][x(t+\tau)+\xi(t+\tau)]dt = \frac{1}{T}\int_{-T/2}^{T/2}[x(t)x(t+\tau)]dt +$$

$$+\frac{1}{T}\int_{-T/2}^{T/2}[x(t)\xi(t+\tau)]dt + \frac{1}{T}\int_{-T/2}^{T/2}[\xi(t)x(t+\tau)]dt + \frac{1}{T}\int_{-T/2}^{T/2}[\xi(t)\xi(t+\tau)]dt =$$

$$= K_{XX}(\tau) + K_{X\xi}(\tau) + K_{\xi X}(\tau) + K_{\xi\xi}(\tau).$$

В полученном выражении $K_{X\xi}(\tau)$ и $K_{\xi X}(\tau)$ есть взаимные корреляционные функции сигнала и помехи, а $K_{XX}(\tau)$ и $K_{\xi\xi}(\tau)$ – автокорреляционные функции сигнала и помехи соответственно.

Поскольку передаваемый сигнал и помеха статистически независимы, то $K_{X\xi}(\tau) = K_{\xi X}(\tau) = 0$. Следовательно,

$$K_{YY}(\tau) = K_{XX}(\tau) + K_{\xi\xi}(\tau).$$

Как известно, корреляционная функция $K_{XX}(\tau)$ периодического сигнала является периодической функцией аргумента τ с тем же периодом.

Функция $K_{\xi\xi}(\tau)$ с увеличением τ стремится к нулю и при $\tau \geq \tau_k$ практически равно нулю (рис. 5.4).

Следовательно, выбирая такое время τ, при котором значением $K_{\xi\xi}(\tau)$ можно пренебречь, мы обеспечим тем самым получение функции $K_{YY}(\tau) = K_{XX}(\tau)$, отображающей полезный сигнал, т.е. выделение полезного сигнала из смеси полезного сигнала с помехой.

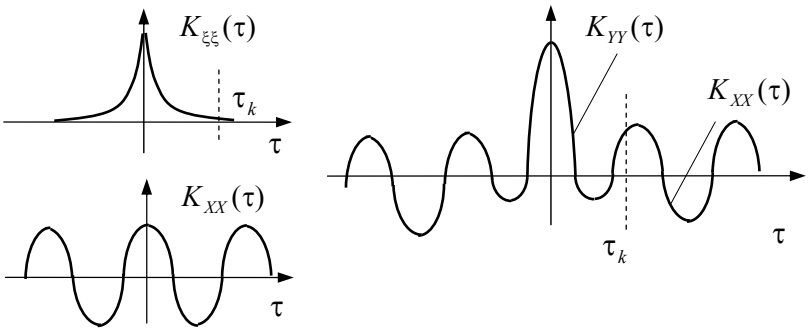


Рис. 5.4. Иллюстрация корреляционного метода

Схема корреляционного приема приведена на рис. 5.5. Схема содержит блок задержки (БЗ), устройство умножения (УУ) и интегратор (И).

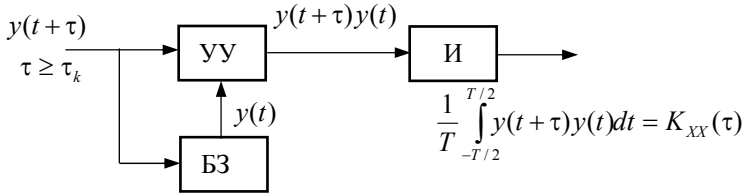


Рис. 5.5. Схема корреляционного приема

5.1.4. Согласованная фильтрация

Согласованные фильтры предназначены для выделения сигналов известной формы на фоне шумов.

Рассмотрим импульсный сигнал $x(t)$ произвольной формы, определенный на интервале $(0, T_X)$ и равный нулю вне этого интервала. Предположим, что принятый сигнал $y(t) = x(t) + \xi(t)$ пропускается через линейный фильтр, импульсная переходная функция которого имеет вид $\Phi(t) = Bx(T_X - t)$ (рис. 5.6). Фильтр с такой характеристикой носит название согласованного.

На выходе фильтра имеет место реакция, определяемая интегралом свертки $y(t) * \Phi(t)$:

$$x_{\Phi}(t) = \int_{-\infty}^{\infty} y(\tau)\Phi(t-\tau)d\tau = B\int_{-\infty}^{\infty}[x(\tau)+\xi(\tau)]x(T_X - t + \tau)d\tau =$$

$$= B\int_{-\infty}^{\infty} x(\tau)x(T_X - t + \tau)d\tau + B\int_{-\infty}^{\infty}\xi(\tau)x(T_X - t + \tau)d\tau .$$

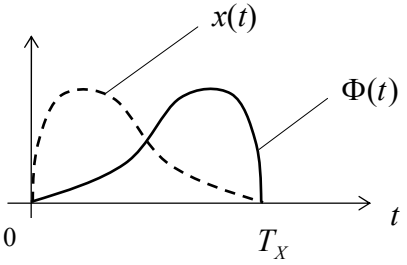


Рис. 5.6. Импульсная переходная функция

Полученные два интеграла представляют собой с точностью до постоянного множителя B корреляционные функции

$$x_{\Phi}(t) = B[K_{XX}(T_X - t) + K_{\xi X}(T_X - t)].$$

Таким образом, выходная реакция фильтра совпадает с выходным сигналом корреляционного фильтра, но при согласованной фильтрации нет необходимости использования устройства перемножения (рис. 5.7).

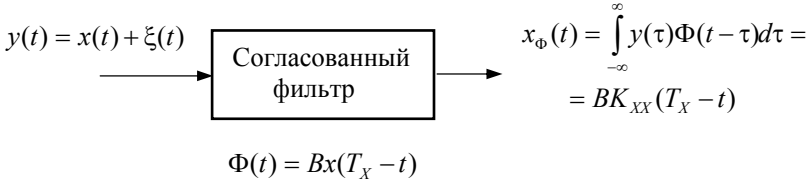


Рис. 5.7. Схема согласованной фильтрации

5.2. СУЩНОСТЬ ОСНОВНОЙ ЗАДАЧИ ПРИЕМА СИГНАЛОВ

Основная задача приема сигналов фактически представляет собой процесс распознавания. Процесс распознавания состоит в классификации явлений по имеющейся информации путем отнесе-

ния воспринимаемой совокупности признаков $X = (x_1, x_2, ..., x_m)$ к области, характеризующей одно из состояний a_i источника информации [2].

С этой целью m - пространство признаков разбивается по какому-либо критерию на n областей $X_1, X_2, ..., X_n$, соответствующих точкам n - пространства состояний $A = (a_1, a_2, ..., a_n)$ источника информации (рис. 5.8).

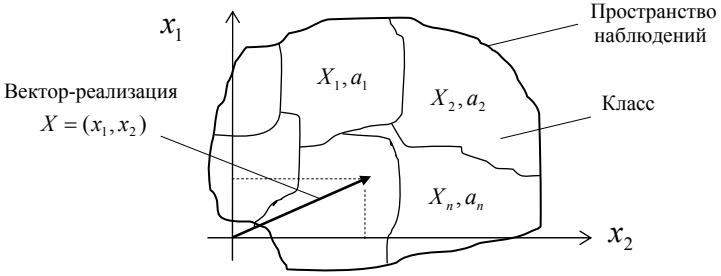


Рис. 5.8. Двумерное m - пространство признаков

Если разбиение производится на две области, распознавание становится двухальтернативным и представляет собой обнаружение некоторого явления. Векторы X в m -пространстве признаков носят название векторов-реализаций, а области разбиения образуют множество классов.

В общем случае зависимость реализаций X от состояний a_i источника носит вероятностный характер. Можно выделить три характерных случая этих зависимостей (для простоты на рис. 5.9 изображены одномерные распределения):

1. В первом случае возможно безошибочное распознавание. Для этого области $X_1, X_2, ..., X_n$ в m - пространстве признаков должны быть выбраны так, чтобы каждая из них включала все возможные значения только одного соответствующего ей признака. При этом вероятность $p(a_i / X_i)$ правильного распознавания i -го состояния при условии $X \in X_i$, которую можно подсчитать по формуле Байеса

$$p(a_i / X_i) = \frac{p(a_i)p(X_i / a_i)}{\sum_j p(a_j)p(X_i / a_j)} = \frac{p(a_i)\int_{X_i} f(X / a_i)dX}{\sum_j p(a_j)\int_{X_i} f(X / a_j)dX}, \quad (5.1)$$

равна единице, так как $f(X / a_j) = 0$ в области $X_i (i \neq j)$.

2. Во втором случае распознавание по вектору-реализации X невозможно $p(a_i / X_i) = p(a_i)$. Отнесение вектора X к какой-либо области не увеличивает вероятности распознавания.

3. В третьем случае распознавание состояний возможно с вероятностями $p(a_i) < p(a_i / X_i) < 1$.

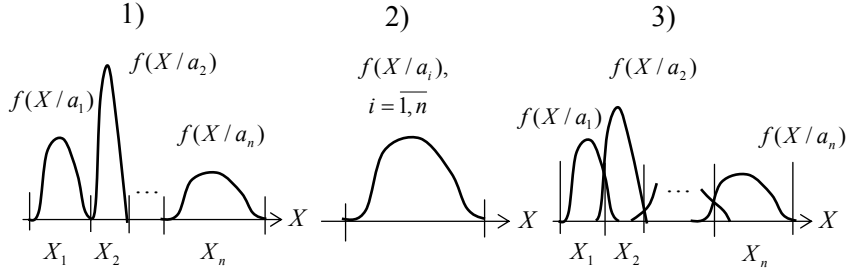


Рис. 5.9. Условные плотности распределений:
1 – случай непересекающихся распределений; 2 – случай тождественных распределений; 3 – случай пересекающихся распределений

Основными характеристиками качества распознавания являются средние ошибки распознавания и средние потери.

Предположим, что при $X \in X_i$ принимается гипотеза о наличии состояния a_i. Вероятность правильного решения составляет при этом $p(a_i / X_i)$, а вероятность ошибки $p_{i\text{ош}} = 1 - p(a_i / X_i)$. Средняя вероятность ошибки распознавания для всех возможных состояний источника равна

$$p_{\text{ош}} = \sum_i p(X_i) p_{i\text{ош}} = 1 - \sum_i p(X_i) p(a_i / X_i) = \\ = 1 - \sum_i P(a_i) \int_{X_i} f(X / a_i) dX \ , \quad (5.2)$$

где $p(X_i)$ означает $p\{X \in X_i\}$.

Если ошибки распознавания отдельных состояний неравноценны, то для характеристики качества могут быть приняты средние потери.

Обозначим через r_{ij} положительное число, определяющее коэффициент потерь от ошибки в результате заключения о состоянии a_i, в то время как источник информации находится в некотором другом состоянии a_j. При попадании вектора X в область X_i условные потери составят

$$r_i = \sum_j p(a_j / X_i) r_{ij} ,$$

а средние потери

$$r = \sum_i p(X_i) r_i = \sum_j p(a_j) \sum_i r_{ij} \int_{X_i} f(X/a_j) dX . \qquad (5.3)$$

Величина r_i носит название условного, а r – среднего риска распознавания. В дальнейшем коэффициенты потерь r_{ij} $(i = j)$, связанные с правильными решениями, предполагаются равными нулю.

5.3. ОБНАРУЖЕНИЕ СИГНАЛА

Для обеспечения возможности обнаружения m-мерное пространство признаков должно быть разбито на две области: X_1 и X_2. Граница этих областей называется решающей поверхностью. В процессе обнаружения решающее устройство определяет, какой области принадлежит вектор-реализация X, и делает заключение о состоянии a_i источника.

Например, в случае трехмерного пространства (рис. 5.10) пространство принятых сигналов условно разбивается на две части:

– область X_2, соответствующая состоянию a_2 источника, когда в принятом сигнале содержится полезный сигнал;

– область X_1, соответствующая состоянию a_1 источника, когда в принятом сигнале отсутствует полезный сигнал.

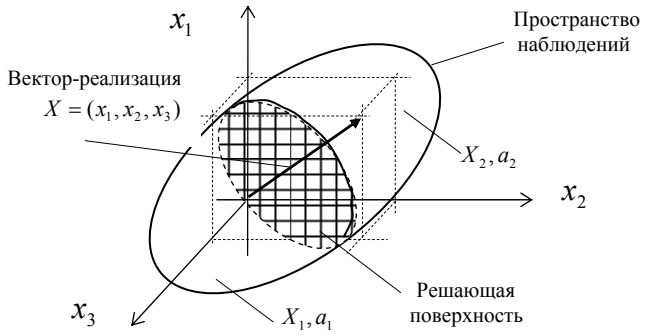


Рис. 5.10. Трехмерное m - пространство признаков

Поскольку ***решающая поверхность*** многомерна, ее задание и хранение могут встретить значительные трудности. Поэтому многомерный случай приводится к одномерному путем перехода к новой переменной, функционально связанной с вектором *X*. Эта переменная носит название отношения ***функций правдоподобия***

$$\lambda = \frac{L(a_2)}{L(a_1)} = \frac{f(X/a_2)}{f(X/a_1)},$$

где $f(X/a_i) = f(x_1, x_2, \ldots, x_m / a_i)$ – многомерная условная плотность распределения вероятностей.

Вместо уравнения решающей поверхности в этом случае достаточно запомнить одно число λ_0, с которым сравнивается текущее значение коэффициента правдоподобия λ. Неравенству $\lambda \le \lambda_0$ соответствует $X \in X_1$ и наличие состояния a_1, когда в принятом сигнале отсутствует полезный сигнал. Неравенству $\lambda > \lambda_0$ соответствует $X \in X_2$ и наличие состояния a_2, когда в принятом сигнале присутствует полезный сигнал.

Уравнение границы (решающей поверхности) в этом случае определяется соотношением

$$F(X) = f(X/a_2) - \lambda_0 f(X/a_1) = 0.$$

Для того чтобы выбрать то или иное правило принятия решения, необходимо руководствоваться определенными критериями.

5.3.1. Критерий максимума правдоподобия (критерий Фишера)

Этот критерий формулируется следующим образом: наиболее правдоподобно то значение параметра a_i, для которого функция правдоподобия $L(a_i)$ максимальна. Другими словами, решающее правило имеет вид

$$\begin{cases} \lambda = \dfrac{L(a_2)}{L(a_1)} > \lambda_0 = 1 \rightarrow a_2; \\ \lambda = \dfrac{L(a_2)}{L(a_1)} \le \lambda_0 = 1 \rightarrow a_1. \end{cases}$$

Практическое достоинство данного критерия заключается в том, что при его применении не требуется знания априорных вероятностей $p(a_1)$ и $p(a_2)$ источника. При этом предполагается, что $\lambda_0 = 1$ соответствует случаю $p(a_1) = p(a_2)$.

На рис. 5.11 приведена графическая иллюстрация данного критерия в предположении, что вектор X одномерен, полезный сигнал представляет собой постоянный уровень величиной A, а помеха типа гауссова шума с $m[\xi] = 0$. В этом случае решающая поверхность вырождается в точку $X_{\text{П}}$.

5.3.2. Критерий идеального наблюдателя (критерий Зигерта–Котельникова)

Согласно данному критерию принимается та гипотеза, при которой обеспечивается минимум средней ошибки принятия решения.

При решении задачи обнаружения сигнала могут иметь место ошибки двух типов:

1. При отсутствии полезного сигнала вектор-реализация X оказывается в области X_2 и принимается в соответствии с этим гипотеза a_2. Это ошибка первого рода или ошибка типа ложной тревоги.

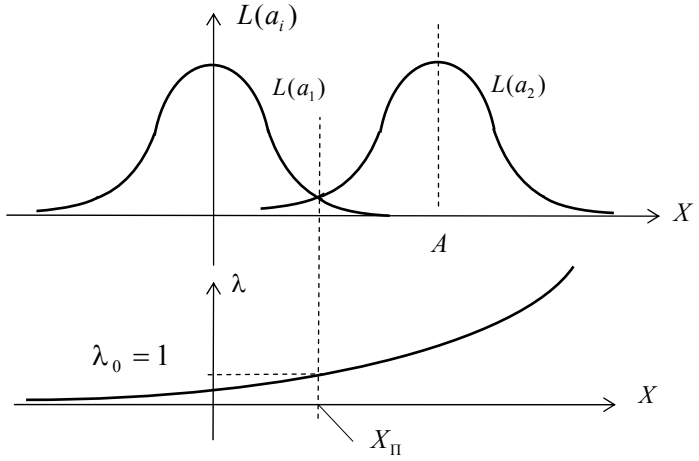


Рис. 5.11. Графики функций правдоподобия $L(a_i)$ и отношения функций правдоподобия λ

2. При наличии полезного сигнала вектор-реализация X оказывается в области X_1 и принимается в соответствии с этим гипотеза a_1. Это ошибка второго рода или ошибка типа пропуска сигнала.

Количественно ошибки первого и второго рода оцениваются условными вероятностями ошибочных решений о наличии полезного сигнала, когда в действительности он отсутствует

$$\alpha = p(X \in X_2 / a_1) = \int_{X_2} f(X / a_1) dX ,$$

и об отсутствии сигнала, когда в действительности он имеется

$$\beta = p(X \in X_1 / a_2) = \int_{X_1} f(X / a_2) dX .$$

Среднюю вероятность ошибки распознавания (5.2) приведем к более удобному виду

$$p_{\text{ош}} = 1 - \sum_i P(a_i) \int_{X_i} f(X / a_i) dX =$$

$$= p(a_1) + p(a_2) - p(a_1) \int_{X_1} f(X / a_1) dX - p(a_2) \int_{X_2} f(X / a_2) dX =$$

$$= p(a_1)[1 - \int_{X_1} f(X/a_1)dX] + p(a_2)[1 - \int_{X_2} f(X/a_2)dX] =$$

$$= p(a_1)\alpha + p(a_2)\beta .$$

Условие оптимального решения по критерию идеального наблюдателя состоит в минимизации

$$p_{\text{ош}} = p(a_1)\int_{X_2} f(X/a_1)dX + p(a_2)\int_{X_1} f(X/a_2)dX .$$

Ошибку второго рода можно представить в виде

$$\beta = p(X \in X_1/a_2) = 1 - p(X \in X_2/a_2) = 1 - \int_{X_2} f(X/a_2)dX .$$

Следовательно,

$$p_{\text{ош}} = p(a_2) - \int_{X_2} [p(a_2)f(X/a_2) - p(a_1)f(X/a_1)]dX .$$

Минимум $p_{\text{ош}}$ будет обеспечен, если интеграл будет максимален, а для этого необходимо, чтобы выполнялось неравенство

$$p(a_2)f(X/a_2) - p(a_1)f(X/a_1) > 0 .$$

Это условие определяет принадлежность вектора X к области X_2, т.е. окончательно решающее правило можно трактовать следующим образом

$$\begin{cases} \lambda = \dfrac{L(a_2)}{L(a_1)} > \lambda_0 = \dfrac{p(a_1)}{p(a_2)} \rightarrow a_2; \\ \lambda = \dfrac{L(a_2)}{L(a_1)} \le \lambda_0 = \dfrac{p(a_1)}{p(a_2)} \rightarrow a_1. \end{cases}$$

Таким образом, при наличии априорных данных $p(a_1)$ и $p(a_2)$ целесообразно применять данный критерий, так как при этом имеется возможность пользоваться дополнительной информацией, позволяющей точнее решить задачу обнаружения сигнала.

5.3.3. Критерий минимального риска (критерий Байеса)

Этот критерий экономически наиболее целесообразен, но требует максимальной априорной информации. Для его использова-

ния должны быть известны, кроме условных распределений $f(X/a_1)$ и $f(X/a_2)$, априорные вероятности состояний источника $p(a_1)$, $p(a_2)$ и коэффициенты потерь r_{12}, r_{21}.

Разбиение m-пространства признаков производится таким образом, чтобы минимизировался средний риск. Это означает, что при достаточно большом числе актов распознавания экономические потери от ошибок будут минимальными.

Средний риск (5.3) для случая обнаружения имеет вид

$$r = r_{21}p(a_1)\alpha + r_{12}p(a_2)\beta =$$

$$= r_{21}p(a_1)\int_{X_2} f(X/a_1)dX + r_{12}p(a_2)\int_{X_1} f(X/a_2)dX =$$

$$= r_{12}p(a_2) - \int_{X_2}[r_{12}p(a_2)f(X/a_2) - r_{21}p(a_1)f(X/a_1)]dX.$$

Минимум r будет при условии, если подынтегральная функция положительная:

$$r_{12}p(a_2)f(X/a_2) - r_{21}p(a_1)f(X/a_1) > 0.$$

Отсюда получаем следующее правило принятия решения:

$$\begin{cases} \lambda = \dfrac{L(a_2)}{L(a_1)} > \lambda_0 = \dfrac{r_{21}p(a_1)}{r_{12}p(a_2)} \rightarrow a_2; \\ \lambda = \dfrac{L(a_2)}{L(a_1)} \le \lambda_0 = \dfrac{r_{21}p(a_1)}{r_{12}p(a_2)} \rightarrow a_1. \end{cases}$$

5.3.4. Критерий Неймана–Пирсона

Часто ошибки первого и второго рода могут привести к существенно различным последствиям, которые невозможно оценить в виде потерь. Это имеет место, например, если одна из этих ошибок приводит к непредсказуемым последствиям. Такую ошибку необходимо ограничить некоторой очень малой величиной. Вторую из ошибок всегда желательно сделать возможно меньшей. Критерий, отражающий эти требования, формулируется как экстремальная задача с ограничением:

$$\begin{cases} \min[\beta = \int\limits_{X_1} f(X / a_2) dX]; \\ \alpha = \int\limits_{X_2} f(X / a_1) dX \le \varepsilon_0 . \end{cases}$$

Величина λ_0 находится в этом случае по соотношению

$$\alpha = \int\limits_{\lambda_0}^{\infty} f(\lambda / a_1) d\lambda = \varepsilon_0 ,$$

так как большие значения λ_0 приводят к большим условным ошибкам β, а меньшие – к недопустимым ошибкам $\alpha > \varepsilon_0$. Правило принятия решения в этом случае имеет традиционный вид

$$\begin{cases} \lambda = \dfrac{L(a_2)}{L(a_1)} > \lambda_0 \ \rightarrow \ a_2; \\ \lambda = \dfrac{L(a_2)}{L(a_1)} \le \lambda_0 \ \rightarrow \ a_1. \end{cases}$$

5.4. РАЗЛИЧЕНИЕ СИГНАЛОВ

При различении сигналов имеет место многоальтернативная ситуация, когда полезный сигнал может иметь много значений и приемное устройство должно определить, какое именно значение из этого множества имеет место в действительности. Методы многоальтернативных решений являются обобщением соответствующих методов двухальтернативных решений.

Пусть источник может иметь n возможных состояний $a_1, a_2, ..., a_n$ с априорными вероятностями $p(a_1), p(a_2), ..., p(a_n)$. При этом m - пространство признаков разбивается на n областей $X_1, X_2, ..., X_n$, соответствующих принятию гипотез $a_1, a_2, ..., a_n$ соответственно. Правила принятия решений и разбивка пространства признаков на области могут производиться в соответствии с любым из критериев, рассмотренных для двухальтернативной ситуации и обобщенных на случай многоальтернативной ситуации.

Процедура работы решающего устройства приемника при раз-

личении сигналов следующая. По данным вектор-реализации X определяются функции правдоподобия

$$L(a_1) = f(X/a_1),\ L(a_2) = f(X/a_2),...,\ L(a_n) = f(X/a_n)$$

и вычисляются отношения правдоподобия

$$\lambda_{ij} = \frac{f(X/a_i)}{f(X/a_j)};\ \ i, j = \overline{1,n};\ i \neq j$$

для всех возможных сочетаний пар a_i и a_j. Полученные значения отношений правдоподобия сравниваются с пороговым значением λ_0 и выбирается такое значение сигнала a_i, для которого все $n-1$ неравенства $\lambda_{ij} > \lambda_0,\ \ j = \overline{1,n};\ i \neq j$ выполняются. λ_0 в зависимости от критерия принимает значения:

$$1;\ \ \frac{p(a_j)}{p(a_i)};\ \ \frac{r_{ij}\,p(a_j)}{r_{ji}\,p(a_i)}.$$

На рис. 5.12 приводится графическая иллюстрация процесса различения сигналов в предположении, что $m = 1$, $n = 3$ и $\lambda_0 = 1$.

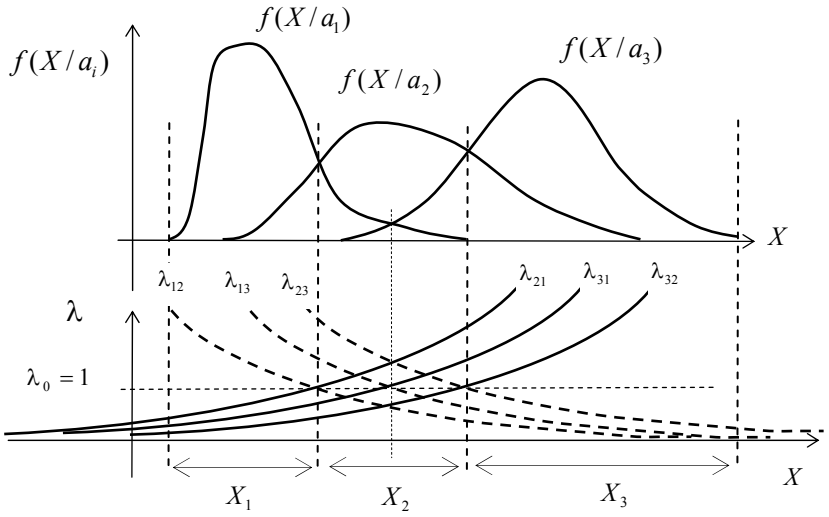


Рис. 5.12. Процесс различения сигналов ($m = 1$ и $n = 3$)

5.5. СИНТЕЗ СТРУКТУРЫ РЕШАЮЩЕГО УСТРОЙСТВА

Рассмотрим общий случай многоальтернативной ситуации, когда полезный сигнал $x(t)$ может принимать n значений $x_1,\ldots,x_n$. Будем полагать помеху $\xi(t)$ нормальной с нулевым математическим ожиданием и аддитивной (рис. 5.13). Следовательно, принимаемый сигнал имеет вид $y(t)=x(t)+\xi(t)$.

Представим сигнал $y(t)$ в виде совокупности из m отсчетов, полагая, что отсчеты осуществляются через интервал времени $\Delta t=\dfrac{1}{2f_c}$, где f_c – граничная полоса пропускания канала связи.

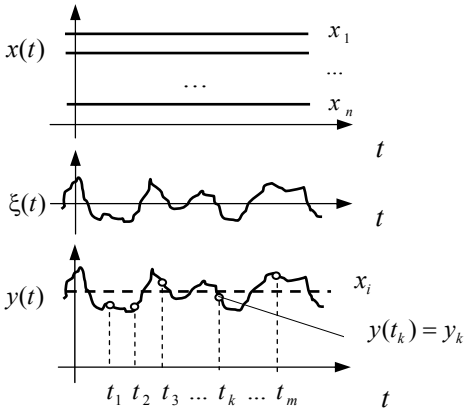


Рис. 5.13. Графическая интерпретация исходных данных

Тогда для любого отсчетного значения принятого сигнала y_k можно записать $y_k=x_i+\xi_k$, где x_i – текущее состояние полезного сигнала; ξ_k – отсчетное значение помехи, распределенное по нормальному закону

$$f(\xi_k)=\frac{1}{\sigma_\xi\sqrt{2\pi}}e^{-\frac{\xi_k^2}{2\sigma_\xi^2}}.$$

При взаимной независимости полезного сигнала и помехи функция $f(y_k/x_i)$ определяется законом распределения помехи.

Так как элемент случайности вносит только помеха, то при любом фиксированном x_i вероятность того, что случайная величина y_k примет значение между y_k и $y_k + dy_k$, равна вероятности того, что помеха будет находиться в пределах между ξ_k и $\xi_k + d\xi_k$, т.е. $f(y_k / x_i)dy_k = f(\xi_k)d\xi_k$. Но $\frac{dy_k}{d\xi_k} = \frac{dx_i}{d\xi_k} + \frac{d\xi_k}{d\xi_k} = 1$, так что

$$f(y_k / x_i) = f(\xi_k) = \frac{1}{\sigma_\xi \sqrt{2\pi}} e^{-\frac{(y_k - x_i)^2}{2\sigma_\xi^2}},$$ где $\xi_k = y_k - x_i$.

Вектор отсчетов помехи определяется многомерной плотностью распределения вероятностей $f(\xi) = f(\xi_1 \xi_2 ... \xi_k ... \xi_m)$, где m – объем выборки. Полагая помеху стационарной и отсчеты некоррелированными, можно многомерный закон распределения помехи представить в виде

$$f(\xi) = f(\xi_1) f(\xi_2) ... f(\xi_m) = \left(\frac{1}{\sigma_\xi \sqrt{2\pi}} \right)^m e^{-\frac{\sum_{k=1}^{m} \xi_k^2}{2\sigma_\xi^2}},$$

а многомерного вектора-реализации Y

$$f(Y / x_i) = f(y_1 / x_i) ... f(y_m / x_i) = \left(\frac{1}{\sigma_\xi \sqrt{2\pi}} \right)^m e^{-\frac{\sum_{k=1}^{m} (y_k - x_i)^2}{2\sigma_\xi^2}}.$$

Оптимальное решающее устройство должно строиться таким образом, чтобы оно могло вычислить отношения функций правдоподобия λ_{ij} и сопоставить их с пороговым значением λ_0. В качестве истинного выбирается такой сигнал x_i, для которого выполняются $n-1$ неравенств

$$\lambda_{ij} > \lambda_0 ; \quad i, j = \overline{1, n}; i \neq j .$$

В случае цифровой обработки отношения функций правдоподобия вычисляются микропроцессором по формуле

$$\lambda_{ij} = \frac{f(Y/x_i)}{f(Y/x_j)} = e^{\frac{\sum_{k=1}^{m}(y_k - x_j)^2 - \sum_{k=1}^{m}(y_k - x_i)^2}{2\sigma_\xi^2}}.$$

В случае аналоговой обработки, когда сигнал принимается непрерывно в течение определенного времени T, проделаем некоторые преобразования. Умножим числитель и знаменатель показателя степени экспоненты на $\Delta t = \frac{T}{m} = \frac{1}{2f_c}$ и перейдем в числителе показателя степени от суммирования отсчетов к интегрированию в пределах от 0 до T:

$$\lambda_{ij} = e^{\frac{\int_0^T [y(t)-x_j]^2\,dt - \int_0^T [y(t)-x_i]^2\,dt}{\sigma_\xi^2 / f_c}} = e^{\frac{2\int_0^T y(t)(x_i - x_j)\,dt - E_i + E_j}{P_0}},$$

где E_i и E_j – энергии сигналов x_i и x_j, $P_0 = \frac{\sigma_\xi^2}{f_c}$ – мощность помехи, приходящаяся на единицу полосы спектра.

Пологая энергии сигналов x_i и x_j одинаковыми, получаем

$$\lambda_{ij} = e^{\frac{2\int_0^T y(t)(x_i - x_j)\,dt}{P_0}}.$$

Наконец, решающее правило можно представить в виде

$$\int_0^T y(t)x_i dt - \int_0^T y(t)x_j dt > \frac{1}{2}P_0 \ln\lambda_0 = \lambda_0';\quad i, j = \overline{1,n};\ i \neq j.$$

Структура решающего устройства представлена на рис. 5.14. Устройство содержит наборы генераторов сигналов $x_1, \ldots, x_n$ (Г), устройств умножения (УУ), интеграторов (И) и схему сравнения (СС).

На выходе УУ получаются произведения $y(t)x_i$, которые затем интегрируются $\int_0^T y(t)x_i dt$. Схема сравнения СС определяет раз-

ность между различными сочетаниями выходных сигналов интеграторов, сравнивает полученные результаты с порогом $\lambda_0^{'}$ и выносит решение в пользу того x_i, для которого выполняется решающее правило. Величина $\lambda_0^{'}$ определяется критерием, положенным в основу синтеза решающего устройства. Например,

$$\lambda_0^{'} = \frac{1}{2} P_0 \ln \frac{p(a_j)}{p(a_i)}.$$

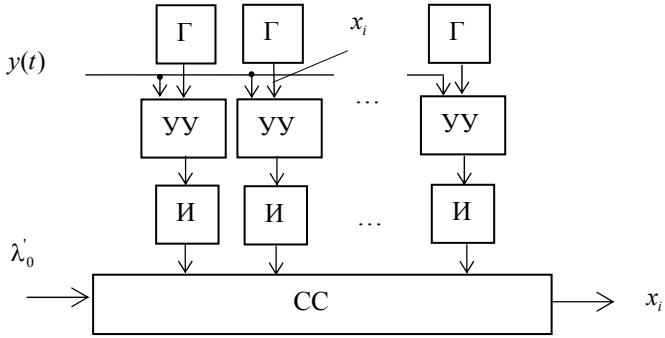


Рис. 5.14. Структура решающего устройства

5.6. ВОССТАНОВЛЕНИЕ СИГНАЛА

Восстановление сигналов сводится к оценке некоторого числа неизвестных параметров полезного сигнала. Ограничимся рассмотрением случая оценки одного из параметров сигнала, например амплитуды A, при заданной форме сигнала. При этом помеху будем полагать аддитивной типа гауссова шума.

Представим входной сигнал в виде

$$y(t) = x(t) + \xi(t) = Af(t) + \xi(t),$$

где $f(t)$ – известная функция времени; A – параметр сигнала.

Задача состоит в том, чтобы по принятой выборке входного сигнала $Y = [y(t_1), y(t_2), ..., y(t_m)]$ определить, каково значение параметра A в полезном сигнале $x(t)$.

Произведем оценку параметра A методом максимума правдо-

подобия. Если отсчет принятого сигнала производится в дискретные моменты времени, то функция правдоподобия для параметра A будет равна

$$L(A) = f(Y/A) = \left(\frac{1}{\sigma_\xi \sqrt{2\pi}}\right)^m e^{-\frac{\sum_{k=1}^{m}[y(t_k) - Af(t_k)]^2}{2\sigma_\xi^2}}.$$

Необходимо найти такое значение параметра A, для которого функция правдоподобия $L(A)$ максимальна. Максимуму функции правдоподобия соответствует минимальное значение $\Phi(A)$ – части показателя степени в выражении $L(A)$:

$$\Phi(A) = \sum_{k=1}^{m}[y(t_k) - Af(t_k)]^2 =$$

$$= \sum_{k=1}^{m}[y(t_k)]^2 - 2A\sum_{k=1}^{m} y(t_k) f(t_k) + A^2 \sum_{k=1}^{m}[f(t_k)]^2.$$

Из условия минимума имеем

$$\frac{d\Phi(A)}{dA} = -2\sum_{k=1}^{m} y(t_k) f(t_k) + 2A\sum_{k=1}^{m}[f(t_k)]^2 = 0,$$

откуда получаем оценочное значение параметра полезного сигнала в случае цифровой обработки сигнала

$$A^* = \frac{\sum_{k=1}^{m} y(t_k) f(t_k)}{\sum_{k=1}^{m}[f(t_k)]^2}.$$

Осуществив переход к непрерывному примеру, т.е. аналоговой обработке, получим

$$A^* = \frac{\int_0^T y(t) f(t) dt}{\int_0^T [f(t_k)]^2 dt} = \frac{\int_0^T y(t) f(t) dt}{E_0},$$

где E_0 – энергия полезного сигнала при амплитуде $A = 1$; T – время приема входного сигнала.

Решающее устройство, осуществляющее операцию оценки параметра сигнала, представлено на рис. 5.15. Устройство содержит генератор сигнала (Г), устройство умножения (УУ), осуществляющее умножение $y(t)$ на $f(t)$, и интегратор (И), производящий интегрирование произведения.

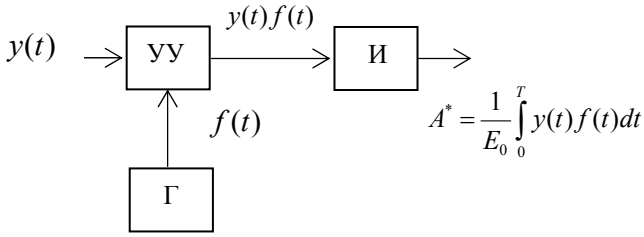


Рис. 5.15. Решающее устройство

Для оценки точности восстановления сигнала воспользуемся известным соотношением $y(t) = Af(t) + \xi(t)$. Тогда

$$A^* = \frac{A}{E_0}\int_0^T [f(t)]^2 dt + \frac{1}{E_0}\int_0^T \xi(t) f(t) dt = A + \frac{1}{E_0}\int_0^T \xi(t) f(t) dt,$$

$$A^* - A = \frac{1}{E_0}\int_0^T \xi(t) f(t) dt.$$

Наконец, относительную погрешность можно найти в соответствии с известным подходом

$$\varepsilon = \frac{|A^* - A|}{A^*} 100\,\% = \frac{\left|\frac{1}{E_0}\int_0^T \xi(t) f(t) dt\right|}{A^*} 100\,\%.$$

ЗАДАЧИ

5.1. По каналу связи, в котором действует аддитивная стационарная помеха, передается периодическая последовательность прямоугольных импульсов. Параметры полезного сигнала: $a = 2$ В, период следования $T = 100$ мс. Помеха имеет нормальное распределение. Среднеквадратическое отклонение помехи

$\sigma_\xi = 5$ В , математическое ожидание $m_\xi = 0$. Обработка сигнала на приемной стороне осуществляется методом синхронного накопления (рис. 5.16). Строб поступает синхронно с полезным сигналом и обеспечивает отпирание приемника на время измерения полезного сигнала (рис. 5.17). Определить время обработки сигналов, необходимое для обеспечения превышения сигнала над помехой в 4 раза $\left(\frac{a_x}{a_\xi} = 4\right)$.

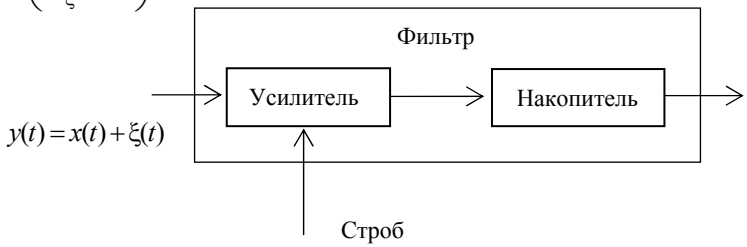


Рис. 5.16. Синхронный накопитель

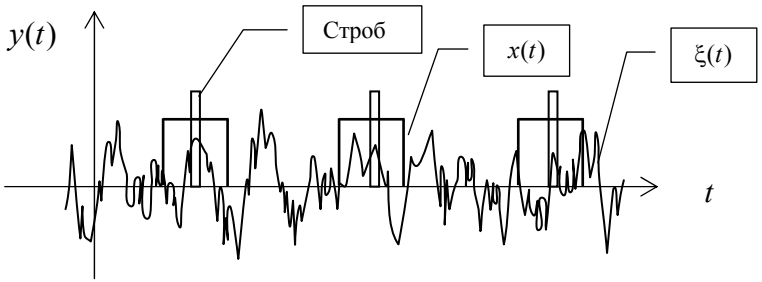


Рис. 5.17. Схема поступления стробов

5.2. Необходимо обнаружить постоянный сигнал величиной $a = 2$ В на фоне аддитивной помехи (рис. 5.18) с нормальным распределением и средним значением, равным нулю. Метод приема – однократный отсчет. Произвести синтез приемного устройства,

работающего на основе критерия максимума правдоподобия, и определить пороговый уровень измерения сигнала $X_{\text{П}}^{1}$.

5.3. Необходимо обнаружить постоянный сигнал величиной $a = 2$ В на фоне аддитивной помехи (см. рис. 5.18) с нормальным распределением $(m_{\xi} = 0{,}1\ ,\ \sigma_{\xi} = 0{,}5)$. Метод приема – однократный отсчет. Произвести синтез приемного устройства, работающего на основе критерия идеального наблюдателя, и определить пороговый уровень измерения сигнала $X_{\text{П}}^{2}$, если отношение априорных вероятностей равно $\dfrac{p(a_1)}{p(a_2)} = 10$.

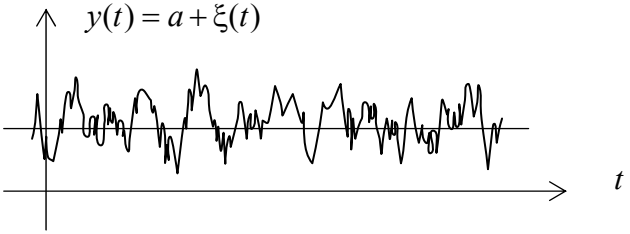


Рис. 5.18. Смесь полезного сигнала и помехи

5.4. Необходимо обнаружить постоянный сигнал величиной $a = 2$ В на фоне аддитивной помехи (см. рис. 5.18) с нормальным распределением $(m_{\xi} = 0{,}1\ ,\ \sigma_{\xi} = 0{,}5)$. Метод приема – однократный отсчет. Произвести синтез приемного устройства, работающего на основе критерия минимального риска, и определить пороговый уровень измерения сигнала $X_{\text{П}}^{3}$, если известны отношения априорных вероятностей $\dfrac{p(a_1)}{p(a_2)} = 10$ и коэффициентов потерь $\dfrac{r_{21}}{r_{12}} = 2$.

5.5. Необходимо обнаружить постоянный сигнал величиной $a = 2$ В на фоне аддитивной помехи (см. рис. 5.18) с нормальным

распределением $(m_\xi = 0\ ,\ \sigma_\xi = 1{,}25)$. Метод приема – однократный отсчет. Произвести синтез приемного устройства, работающего на основе критерия Неймана – Пирсона, и определить пороговый уровень измерения сигнала X_Π, если вероятность ошибки первого рода не должна превышать величины $\varepsilon_0 = 0{,}05$.

5.6. Необходимо обнаружить постоянный сигнал величиной $a = 2$ В на фоне аддитивной помехи с нормальным распределением $(m_\xi = 0{,}1\ ,\ \sigma_\xi = 0{,}5)$. Метод приема – двукратный отсчет. Произвести синтез приемного устройства, работающего на основе критерия минимального риска, если

$$p(a_1) = 0{,}9\ ,\ \ p(a_2) = 0{,}1\ ,\ r_{12} = 5\ ,\ \ r_{21} = 2.$$

5.7. На вход приемника поступает смесь постоянного сигнала с амплитудой a и аддитивной помехи (рис. 5.19), распределенной по нормальному закону. Произведены два замера входного сигнала $y_1(t)$, $y_2(t)$ $(t_2 - t_1 > \tau_k)$. Априорные вероятности гипотез равны. Построить условные плотности распределения вероятностей выборки $Y = \{y_1(t), y_2(t)\}$ при разных гипотезах в предположении, что отсчеты помех $\xi_1(t)$ и $\xi_2(t)$ статистически независимы.

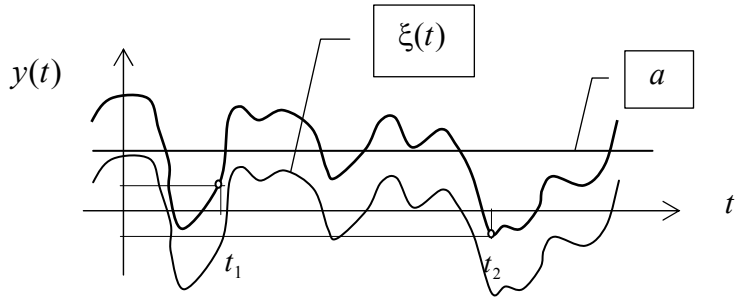


Рис. 5.19. Схема отсчетов входного сигнала

5.8. На вход приемника поступает смесь постоянного сигнала с амплитудой $a = 0{,}4$ В и аддитивной помехи (рис. 5.19), распре-

деленной по нормальному закону $(m_\xi = 0\ ,\ \sigma_\xi = 0{,}25)$. Произведено два замера входного сигнала

$$y_1(t) = 0{,}25 \text{ В}, \quad y_2(t) = -0{,}3 \text{ В} \quad (t_2 - t_1 > \tau_k).$$

Априорные вероятности гипотез равны. Найти апостериорные вероятности гипотез после указанных выше замеров.

6. ИЗМЕРЕНИЕ ИНФОРМАЦИИ

Изложение информационных моделей сигналов также начнем с краткого анализа общей модели системы связи (рис. 6.1). Однако теперь от рассмотрения непрерывных источников перейдем к рассмотрению дискретных источников и связанных с ним соответствующих технических средств.

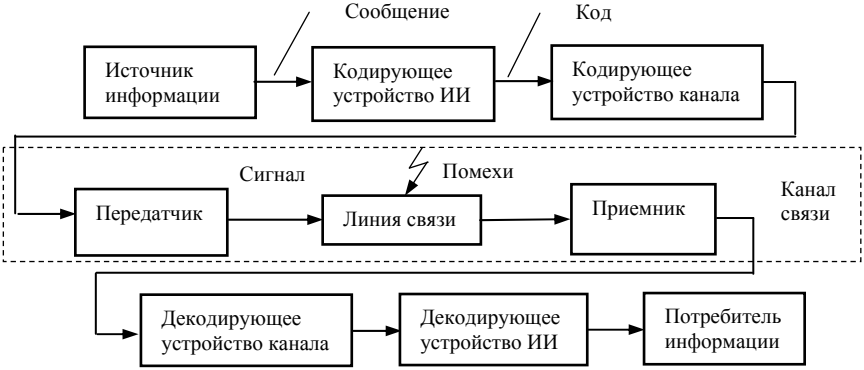


Рис. 6.1. Общая модель системы связи

Источник вырабатывает информацию в виде дискретных сообщений, т.е. ***сообщение*** – это форма представления информации. Целью кодирования, как правило, является согласование источника информации с каналом связи, обеспечивающее либо максимально возможную скорость передачи информации, либо заданную помехоустойчивость. Кодирующие устройства преобразуют сообщения в сигнал, пригодный для передачи по каналу связи, т.е. сигнал – материальный переносчик сообщения.

Введение двух разных кодирующих и декодирующих устройств вызвано желанием подчеркнуть различие между операциями кодирования и декодирования, зависящими от характеристик пары «источник – адресат», и операциями, которые зависят от характеристик канала.

Задача системы связи состоит в том, чтобы воспроизвести сообщение источника адресату в месте удобном для него. При этом мы не рассчитываем, естественно, на абсолютно точное воспроиз-

ведение, а полагаемся на воспроизведение сообщения, удовлетворяющее некоторым специфическим целям.

Для целей инженерного проектирования необходимо ввести количественную меру информации. При этом постулируется, что сообщение может быть описано должным образом только в терминах теории вероятностей как элемент соответствующим образом определенных множеств [11,12].

В статистической теории связи существенным является лишь то, что посылаемое сообщение принадлежит некоторому множеству сообщений. Если это множество конечно, то число сообщений или любую монотонную функцию от числа сообщений можно рассматривать как меру информации. При этом информация как бы создается случайным выбором одного сообщения из множества возможных.

Семантические аспекты сообщения, т.е. его значение, физическая или умозрительная сущность, не имеют отношения к технической стороне вопроса передачи информации.

6.1. ВЗАИМНАЯ ИНФОРМАЦИЯ

Рассмотрим два дискретных множества $X = \{x_k\}$ и $Y = \{y_i\}$. ***Декартовое произведение множеств*** XY определяется, как известно, как совокупность всех упорядоченных пар (x_k, y_i). При этом пары $(x_k, y_i) \in XY$ и $(y_i, x_k) \in YX$ считаются различными, даже если $X = Y$. Элементы x_k, y_i – компоненты пары (x_k, y_i).

Пусть на множестве X задано распределение вероятностей $p(x_k)$, на множестве Y – $p(y_i)$, а на множестве XY – $p(x_k y_i)$. Естественно, что распределения носят дискретный характер, с учетом дискретности рассматриваемых множеств.

Для всех множеств очевидны соотношения

$$\sum_{x_k \in X} p(x_k) = 1; \quad \sum_{y_i \in Y} p(y_i) = 1; \quad \sum_{(x_k, y_i) \in XY} p(x_k y_i) = 1 .$$

Распределения безусловных $p(x_k)$, $p(y_i)$, $p(x_k y_i)$ и условных $p(x_k / y_i)$, $p(y_i / x_k)$ вероятностей, в свою очередь, связаны формулами

$$\sum_{y_i \in Y} p(x_k y_i) = p(x_k); \quad \sum_{x_k \in X} p(x_k y_i) = p(y_i);$$

$$p(x_k / y_i) = \frac{p(x_k y_i)}{p(y_i)}; \quad p(y_i / x_k) = \frac{p(x_k y_i)}{p(x_k)} .$$

Все приведенные соотношения удобно интерпретировать графически (рис. 6.2).

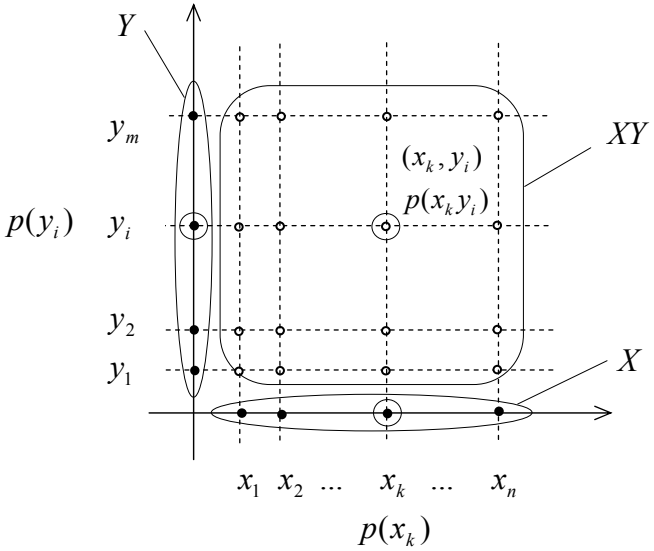


Рис. 6.2. Графическая иллюстрация декартового произведения

Произведение множеств XY образовано совокупностью точек (x_k, y_i). Каждый столбец этой совокупности соответствует точке $x_k \in X$, а каждая строка – точке $y_i \in Y$. Другими словами, точки множеств X и Y могут рассматриваться как координаты точек $(x_k, y_i) \in XY$.

Вероятность $p(x_k)$ есть сумма вероятностей $p(x_k y_i)$ по всем точкам столбца с x_k. Аналогично, $p(y_i)$ есть сумма вероятностей $p(x_k y_i)$ по всем элементам строки с y_i. Условная вероятность $p(x_k / y_i)$ равна вероятности, предписанной точке (x_k, y_i) –

$p(x_k y_i)$, деленной на сумму по строке y_i, а условная вероятность $p(y_i / x_k)$ равна $p(x_k y_i)$, деленной на сумму по столбцу x_k.

Подобным образом можно определить произведение большего числа множеств и ввести соответствующие распределения.

Введем способ измерения информации, содержащейся в y_i относительно x_k. Будем трактовать x_k и y_i как случайные события или сообщения на входе и выходе некоторой системы. В этом нет ничего принципиально нового, с таким подходом мы встречаемся в кибернетике достаточно часто (рис. 6.3).

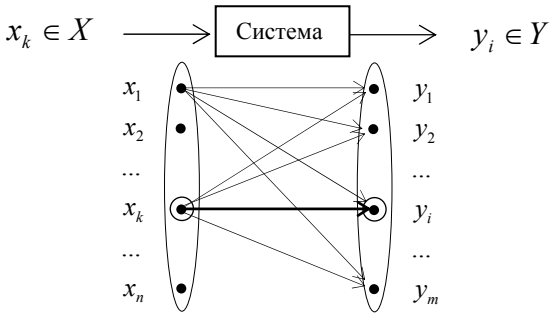


Рис. 6.3. Передача информации через систему

Встанем на позиции внешнего наблюдателя, который следит за системой со стороны и которому известны события и на входе, и на выходе. С позиции этого наблюдателя определим, в какой степени y_i определяет x_k. Другими словами, определим меру количества информации, переданной через систему, или какую информацию в количественном смысле содержит y_i относительно x_k.

Предполагается, что вход системы в некотором смысле определяет ее выход. Это требование вполне естественно, иначе о какой-либо передаче сообщений не могло быть и речи.

Пусть некоторый источник информации может порождать семь различных равновероятных сообщений, которые кодируются двоичным кодом и посимвольно передаются адресату (табл. 6.1).

Предположим, что передается сообщение с номером 2. Априорная вероятность получить это сообщение равна 1/7. После того как

принят первый символ "0", из рассмотрения исключаются коды с номерами 4, 5, 6, 7, начинающиеся с символа "1", и вероятность получения кода 2 возрастает. Расчет этой вероятности прост и сводится к простой нормировке $p(u_2 / y_1^0) = \dfrac{1/7}{1/7+1/7+1/7} = 1/3$. После принятия второго символа "1" вероятность получения кода 2 равна 1/2 , а после принятия последнего символа "0" неопределенность полностью исчезает. В этом случае вероятность приема сообщения с номером 2 при последовательной посимвольной передаче изменяется от 1/7 до 1.

Таблица 6.1

Сообщение u_k	Вероятность сообщения $p(u_k)$	Двоичный код $y_1y_2y_3$	Вероятность появления на выходе k - го сообщения при условии появления одного, двух и трех символов		
			$p(u_k / y_1^0)$	$p(u_k / y_1^0 y_2^1)$	$p(u_k / y_1^0 y_2^1 y_3^0)$
u_1	1/7	001	1/3	0	0
u_2	1/7	010	1/3	1/2	1
u_3	1/7	011	1/3	1/2	0
u_4	1/7	100	0	0	0
u_5	1/7	101	0	0	0
u_6	1/7	110	0	0	0
u_7	1/7	111	0	0	0

Разобранный пример показывает, что информация, содержащаяся в $y_i(010)$ относительно $x_k(u_2)$, сводится к изменению вероятности от априорного значения $p(x_k) = p(u_2) = 1/7$ к апостериорному $p(x_k / y_i) = p(u_2 / y_1^0 y_2^1 y_3^0) = 1$.

Удобной мерой этого изменения оказался логарифм отношения апостериорной вероятности $p(x_k / y_i)$ к априорной $p(x_k)$. Количество информации, содержащееся в событии y_i относительно события x_k, определяется как $J(x_k; y_i) = \log_a \dfrac{p(x_k / y_i)}{p(x_k)}$ и носит название ***взаимной информации***.

Основание логарифма определяет единицу измерения количества информации. Если берется двоичный логарифм, то информация измеряется в битах, если натуральный логарифм, то в натах, если десятичный логарифм, то в дитах (хартли). Количественно эти меры соотносятся следующим образом:

1 нат = 1,44 бит;

1 дит = 2,30 нат = 3,32 бит.

На практике чаще применяют двоичный логарифм, так как он непосредственно дает количество информации в битах, что хорошо согласуется с двоичной логикой вычислительных машин. В теоретических расчетах удобнее пользоваться натуральным логарифмом из-за простоты операций дифференцирования и интегрирования.

Наконец, выбранная логарифмическая мера количества информации удобна по следующим причинам:

1. Многие технические параметры зависят линейно от логарифма числа возможностей. Например, добавление одного разряда к регистру удваивает число возможных состояний регистра. Тем самым прибавляется единица к логарифму этого числа при основании два $n+1 \to \log_2 2^n + 1$.

2. Логарифмическая мера ближе к нашему интуитивному представлению о подходящей мере. Мы измеряем количества с помощью линейного сравнения с принятыми эталонами. Естественно для нас считать, что два одинаковых запоминающих устройства должны обладать вдвое большей памятью для хранения информации, чем одно; а два идентичных канала должны иметь удвоенную пропускную способность.

3. С математической точки зрения логарифмическая мера более проста, так как многие операции, в частности операции предельного перехода, осуществляются наиболее просто при использовании логарифмов.

6.2. ОСНОВНЫЕ СВОЙСТВА ВЗАИМНОЙ ИНФОРМАЦИИ

1. Взаимная информация обладает свойством симметричности:

$$J(x_k; y_i) = \log_2 \frac{p(x_k / y_i)p(y_i)}{p(x_k)p(y_i)} = \log_2 \frac{p(x_k y_i)}{p(x_k)p(y_i)} =$$

$$= \log_2 \frac{p(x_k)p(y_i / x_k)}{p(x_k)p(y_i)} = \log_2 \frac{p(y_i / x_k)}{p(y_i)} = J(y_i; x_k).$$

Другими словами, информация, содержащаяся в y_i относительно x_k, равна информации, содержащейся в x_k относительно y_i. Поэтому $J(x_k; y_i)$ называют взаимной информацией между событиями x_k и y_i. Взаимная информация является мерой статистической связи двух событий, сообщений. Действительно, если события независимы, то $p(x_k y_i) = p(x_k)p(y_i)$ и $J(x_k; y_i) = 0$.

2. Взаимная информация положительна $J(x_k; y_i) > 0$, если условная вероятность появления одного события при условии, что второе событие уже произошло, больше безусловной вероятности его появления – $p(x_k / y_i) > p(x_k)$. Взаимная информация отрицательна $J(x_k; y_i) < 0$, если условная вероятность осуществления события меньше безусловной вероятности – $p(x_k / y_i) < p(x_k)$. Характер зависимости взаимной информации от апостериорной и априорной вероятностей приведен на рис. 6.4.

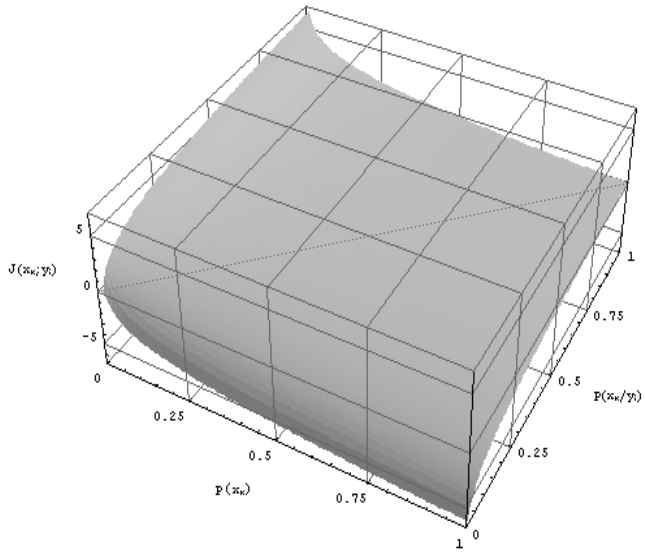


Рис. 6.4. Взаимная информация

Для определения других свойств взаимной информации положим, что (x_k, y_i, z_j) – упорядоченная тройка множества XYZ, и введем в рассмотрение модель, в которой будем считать, что x_k – событие на входе системы, а y_i, z_j – символы, порождаемые на выходе этой системы.

3. Условная взаимная информация определяется следующим образом: $J(x_k; z_j / y_i) = \log_2 \dfrac{p(x_k / y_i z_j)}{p(x_k / y_i)}$.

4. ***Свойство аддитивности***. Пусть $J(x_k; y_i z_j)$ – взаимная информация события x_k и пары (y_i, z_j). Выразим взаимную информацию $J(x_k; y_i z_j)$ через информацию, которую несет символ y_i относительно события x_k, и информацию, которую несет символ z_j относительно того же события x_k:

$$J(x_k; y_i z_j) = \log_2 \frac{p(x_k / y_i z_j) p(x_k / y_i)}{p(x_k) p(x_k / y_i)} = \log_2 \frac{p(x_k / y_i)}{p(x_k)} +$$

$$+ \log_2 \frac{p(x_k / y_i z_j)}{p(x_k / y_i)} = J(x_k; y_i) + J(x_k; z_j / y_i). \qquad (6.1)$$

Информация, содержащаяся в символах y_i и z_j, равна сумме информации, содержащейся в символе y_i относительно x_k, и информации, содержащейся в символе z_j, при условии, что y_i известно.

Порядок символов в (6.1) безразличен, поэтому

$$J(x_k; y_i z_j) = J(x_k; y_i) + J(x_k; z_j / y_i) = J(x_k; z_j) + J(x_k; y_i / z_j).$$

Суммируя оба выражения, получаем

$$J(x_k; y_i z_j) = \frac{1}{2}[J(x_k; y_i) + J(x_k; z_j) + J(x_k; y_i / z_j) + J(x_k; z_j / y_i)].$$

В силу свойства симметричности можно рассматривать $J(y_i z_j; x_k)$ как количество информации, которое содержится в x_k относительно пары (y_i, z_j): $J(y_i z_j; x_k) = J(y_i; x_k) + J(z_j; x_k / y_i)$.

Рассмотрим пример. Пусть на вход кодера может поступать семь сообщений, которые он кодирует в трехразрядный двоичный код (табл. 6.2). На выходе кодера порождается последовательность в виде трех символов. Используя свойство аддитивности информации, вычислим, какое количество информации несет кодовая последовательность 010 относительно сообщения u_2, и какое количество информации несет каждый двоичный символ об этом сообщении.

Таблица 6.2

Сообщение u_k	Вероятность сообщения $p(u_k)$	Двоичный код $y_1y_2y_3$	Вероятность появления на выходе k-го сообщения, при условии появления одного, двух и трех символов		
			$p(u_k / y_1^0)$	$p(u_k / y_1^0 y_2^1)$	$p(u_k / y_1^0 y_2^1 y_3^0)$
u_1	1/2	001	2/3	0	0
u_2	1/8	010	1/6	1/2	1
u_3	1/8	011	1/6	1/2	0
u_4	1/16	100	0	0	0
u_5	1/16	101	0	0	0
u_6	1/16	110	0	0	0
u_7	1/16	111	0	0	0

Если кодер работает идеально, без ошибок, то сообщение u_2 несет относительно кодовой последовательности или кодовая последовательность относительно сообщения следующее количество информации:

$$J(u_2; y_1^0 y_2^1 y_3^0) = \log_2 \frac{p(u_2 / y_1^0 y_2^1 y_3^0)}{p(u_2)} = \log_2 \frac{1}{p(u_2)} = 3 .$$

Посимвольный вклад, т.е. количество информации, которое несет каждый последующий символ о сообщении u_2, имеет вид

$$J(u_2; y_1^0) = \log_2 \frac{P(u_2 / y_1^0)}{P(u_2)} = \log_2 \frac{1/6}{1/8} = \log_2 4/3;$$

$$J(u_2; y_2^1 / y_1^0) = \log_2 \frac{P(u_2 / y_1^0 y_2^1)}{P(u_2 / y_1^0)} = \log_2 \frac{1/2}{1/6} = \log_2 3;$$

$$J(u_2; y_3^0 / y_1^0 y_2^1) = \log_2 \frac{P(u_2 / y_1^0 y_2^1 y_3^0)}{P(u_2 / y_1^0 y_2^1)} = \log_2 \frac{1}{1/2} = \log_2 2.$$

Используя свойство аддитивности, получаем

$$J(u_2; y_1^0 y_2^1 y_3^0) = J(u_2; y_1^0) + J(u_2; y_2^1 / y_1^0) + J(u_2; y_3^0 / y_1^0 y_2^1) = 3.$$

6.3. СОБСТВЕННАЯ ИНФОРМАЦИЯ

Введем теперь в рассмотрение понятие собственной информации. Очевидно, что $p(x_k / y_i) \le 1$. Следовательно,

$$J(x_k; y_i) = \log_2 \frac{p(x_k / y_i)}{p(x_k)} \le \log_2 \frac{1}{p(x_k)} = -\log_2 p(x_k) = J(x_k).$$

Назовем $J(x_k) = -\log_2 p(x_k)$ количеством ***собственной информации*** x_k.

Таким образом, информация какого-либо события измеряется логарифмом величины, обратной величине вероятности появления этого события (рис. 6.5). Отсюда следует, что чем меньше вероятность появления события, тем больше информации оно в себе содержит.

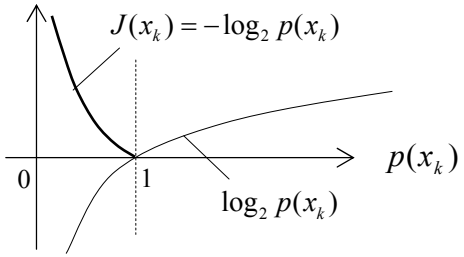


Рис. 6.5. Собственная информация

Если вернуться к примеру кодера с безошибочным кодированием, сообщение на его входе x_k требует для своего однозначного определения количества информации $J(x_k) = -\log_2 p(x_k)$. Взаимная информация между x_k и y_i становится равной собственной информации, т.е. $J(x_k) = J(x_k; y_i)$, когда $p(x_k / y_i) = 1$. Тогда y_i

однозначно определяет x_k, а $J(x_k) = -\log_2 p(x_k)$ можно трактовать как максимальное количество информации, которое нужно иметь о сообщении x_k.

Используя понятие собственной информации, можно несколько иначе интерпретировать взаимную информацию:

$$J(x_k; y_i) = \log_2 \frac{p(x_k / y_i)}{p(x_k)} = -\log_2 p(x_k) + \log_2 p(x_k / y_i) =$$
$$= J(x_k) - J(x_k / y_i),$$

где $J(x_k / y_i) = -\log_2 p(x_k / y_i)$. Информация, содержащаяся в событии y_i относительно x_k, равна разности между количествами информации, требуемой для идентификации x_k до и после того, как становится известным событие y_i.

Симметричное представление взаимной информации приводит к формулам:

$$J(x_k; y_i) = \log_2 \frac{p(x_k y_i)}{p(x_k) p(y_i)} = J(x_k) + J(y_i) - J(x_k y_i);$$
$$J(x_k y_i) = -\log_2 p(x_k y_i) = J(x_k) + J(y_i) - J(x_k; y_i).$$

6.4. МЕРА ИНФОРМАЦИИ КАК СЛУЧАЙНАЯ ВЕЛИЧИНА

Как можно видеть из предыдущих подразделов, введенные меры количества информации являются случайными величинами, т.е. числовыми функциями элементов выборочного пространства.

Мера информации является довольно необычной случайной величиной, так как ее значение зависит от вероятностной меры, однако с ней можно обращаться так же, как с любой случайной величиной.

Например, собственная информация является случайной величиной, так как зависит от распределения вероятностей событий. Мера информации $J(x_k), x_k \in X$ принимает на множестве X различные значения $-\log_2 p(x_k)$ с вероятностью $p(x_k)$ и имеет распределение $p\{J(x_k) = J\}$.

Аналогичным образом определяется и взаимная информация как случайная величина $J(x_k; y_i), (x_k, y_i) \in XY$, которая на множестве XY принимает различные значения $\log_2 \dfrac{p(x_k / y_i)}{p(x_k)}$ с вероятностью $p(x_k y_i)$ и имеет распределение $p\{J(x_k; y_i) = J\}$.

Собственная и взаимная информации имеют среднее значение, дисперсию и моменты всех порядков:

$$m[J(x_k)] = \sum_{x_k \in X} J(x_k) p(x_k) = J(X);$$

$$D[J(x_k)] = \sum_{x_k \in X} [J(x_k) - J(X)]^2 p(x_k);$$

$$m[J(x_k; y_i)] = \sum_{(x_k, y_i) \in XY} J(x_k; y_i) p(x_k y_i) = J(X; Y);$$

$$D[J(x_k; y_i)] = \sum_{(x_k, y_i) \in XY} [J(x_k; y_i) - J(X; Y)]^2 p(x_k y_i);$$

...

В качестве примера рассмотрим источник информации со статистикой, представленной в табл. 6.3.

Таблица 6.3

Сообщение u_k	Вероятность сообщения $p(u_k)$	Собственная информация $J(u_k)$	Распределение вероятностей $p\{J(u_k) = J\}$
u_1	1/2	1	$p\{J(u_k) = 1\} = 1/2$
u_2	1/8	3	
u_3	1/8	3	$p\{J(u_k) = 3\} = 1/4$
u_4	1/16	4	
u_5	1/16	4	$p\{J(u_k) = 4\} = 1/4$
u_6	1/16	4	
u_7	1/16	4	

Распределение количества собственной информации $p\{J(u_k) = J\}$ представлено на рис. 6.6. Функция распределения $F(J) = p\{J(u_k) < J\}$ (рис. 6.7) определяется обычным образом с

учетом дискретности. Математическое ожидание и дисперсия равны:

$$J(X) = 1\frac{1}{2} + 3\frac{1}{4} + 4\frac{1}{4} = 2{,}25;$$

$$D[J(x_k)] = \sum_{x_k \in X} J^2(x_k)\, p(x_k) - (2{,}25)^2 = 1{,}7.$$

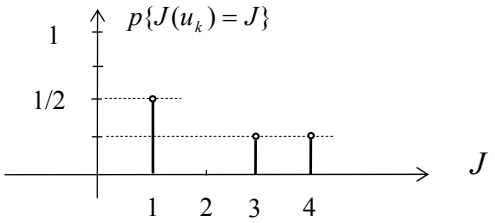


Рис. 6.6. Распределение вероятностей

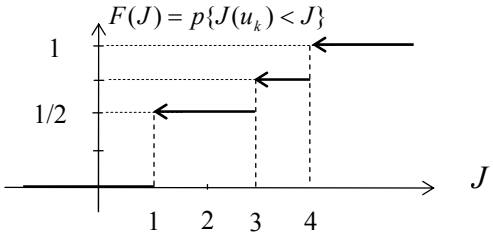


Рис. 6.7. Функция распределения вероятностей

6.5. ЭНТРОПИЯ

Собственная информация сообщения $J(x_k) = -\log_2 p(x_k)$ интерпретируется как количество информации, требуемое для однозначного определения этого сообщения. Среднее количество информации, необходимое для идентификации любого сообщения из множества X, носит название ***энтропии***

$$H(X) = -\sum_{x_k \in X} p(x_k) \log_2 p(x_k) = m[J(x_k)] = J(X).$$

Энтропию можно рассматривать как меру неопределенности сообщения до того, как это сообщение принято.

Для полного понимания этой важнейшей для практических приложений меры количества информации рассмотрим основные ее свойства.

1. $H(X)$ является непрерывной неотрицательной функцией своих аргументов $p(x_k) \in [0,1]$. Действительно, $H(X) \geq 0$, так как $\log_2 p(x_k) \leq 0$. Непрерывность энтропии вытекает из того факта, что бесконечно малое приращение вероятности в любой точке оси $p(x_k)$ приводит к бесконечно малому приращению энтропии.

2. Энтропия симметрична относительно своих аргументов $p(x_1), p(x_2),..., p(x_n),\ n = |X|$, т.е. она не изменяется при любой их перестановке.

3. $H(X) = 0$ при $p(x_k) = 1,\ p(x_j) = 0;\ j = \overline{1,n};\ k \neq j$. Справедливость этого утверждения следует из известного предела $\lim\limits_{p(x_k)\to 0} p(x_k)\log_2 p(x_k) = 0$.

4. Энтропия достигает максимального значения при равенстве $p(x_1) = p(x_2) = ... = p(x_n) = \dfrac{1}{n}$. Доказательство этого свойства проведем с использованием метода неопределенных множителей Лагранжа. Максимизацию $H(X)$ можно представить как решение экстремальной задачи с ограничением:

$$\begin{cases} \max\limits_{\{p(x_k)\}} H[p(x_1),..., p(x_n)] = -\dfrac{1}{\ln 2}\sum\limits_{x_k \in X} p(x_k)\ln p(x_k); \\ \sum\limits_{x_k \in X} p(x_k) = 1. \end{cases}$$

Поиск условного экстремума заменим на нахождение безусловного экстремума за счет введения неопределенного множителя λ и функционала

$$\Phi[p(x_1),..., p(x_n), \lambda] = -\frac{1}{\ln 2}\sum_{x_k \in X} p(x_k)\ln p(x_k) + \lambda[1 - \sum_{x_k \in X} p(x_k)].$$

Необходимое условие максимума для произвольного k можно представить в следующем виде:

$$\frac{\partial \Phi}{\partial p(x_k)} = -\frac{1}{\ln 2}[\ln p(x_k) + 1] - \lambda = 0,$$

откуда получаем, что $p(x_k) = e^{-(1+\lambda \ln 2)}$ и не зависит от k. Следовательно, все вероятности равны $p(x_1) = p(x_2) = ... = p(x_n) = p$, а так как $\sum_{x_k \in X} p(x_k) = np = 1$, то $p = \frac{1}{n}$. Таким образом, энтропия множества сообщений достигает наибольшего значения, когда все сообщения равновероятны. Другими словами, среднее значение информации, необходимое для однозначного определения сообщения из множества возможных сообщений, достигает своего максимума, когда все сообщения равновозможны.

5. Энтропия $H(X)$ удовлетворяет неравенству $H(X) \le \log_2 n$, а знак равенства достигается при $p(x_k) = p = \frac{1}{n}, k = \overline{1,n}$. Это утверждение можно доказать, воспользовавшись неравенством $\ln a \le a - 1$, которое следует из того факта, что функция $\ln a$ касается прямой $a - 1$ в точке $a = 1$ и ее наклон является монотонно убывающей функцией (рис. 6.8). Преобразования

$$H(X) - \log_2 n = \sum_{x_k \in X} p(x_k) \log_2 \frac{1}{p(x_k)} - \log_2 n \sum_{x_k \in X} p(x_k) =$$

$$= \frac{1}{\ln 2} \sum_{x_k \in X} p(x_k) \ln \frac{1}{np(x_k)} \le \frac{1}{\ln 2} \sum_{x_k \in X} p(x_k)[\frac{1}{np(x_k)} - 1] = 0$$

подтверждают утверждение $H(X) \le \log_2 n$. Знак равенства имеет место, когда $a = \frac{1}{np(x_k)} = 1$.

Для двумерного случая, когда $n = 2$, энтропию можно представить в виде $H(X) = -p \log_2 p - (1-p) \log_2 (1-p)$, где p – вероятность появления одного из двух сообщений (рис. 6.9).

Основной смысл формулы $H(X) \le \log_2 n$ может быть еще сформулирован в следующем виде. Для любого заданного алфавита X количество информации, которое в среднем может содер-

жаться в одном символе $x_k \in X$, достигает максимума, когда все символы используются равновероятно. Это максимально возможное среднее значение информации на символ называют ***информационной пропускной способностью алфавита***. Она равна $\log_2 |X|$. Например, пропускная способность русского алфавита равна $\log_2 33 \approx 5$ бит/симв.

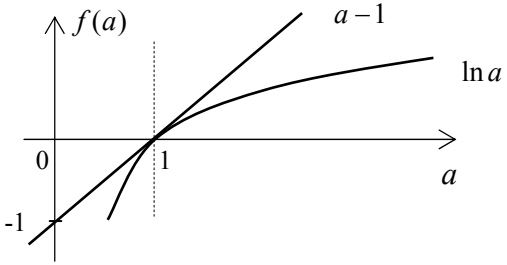


Рис. 6.8. Графическая иллюстрация неравенства $\ln a \le a - 1$

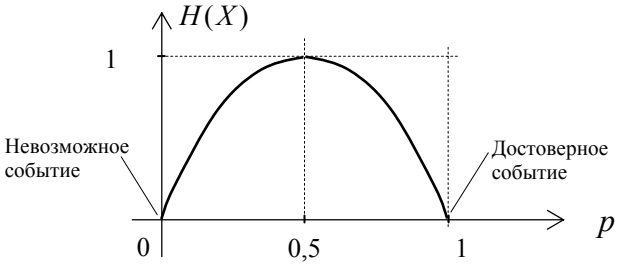


Рис. 6.9. Энтропия для случая $n = 2$

Энтропия максимальна, если буквы алфавита равновероятны и взаимно независимы. Если появление букв не равновероятно, то энтропия уменьшается. Еще меньшей будет энтропия при наличии статистической связи букв. Мерой количественной оценки избыточности алфавита является ***коэффициент избыточности***

$$K_И = \frac{H_{\max}(X) - H(X)}{H_{\max}(X)}.$$

Имеется определенная избыточность и в русском языке. Статистика появления букв алфавита в тексте приведена в табл. 6.4, а в

табл. 6.5 – оценка статистической связи букв [1]. Следовательно, избыточность русского языка составляет $K_И \approx \frac{5-2}{5} = 0{,}6$.

Таблица 6.4

Частота букв в русском языке			
Пробел 0,175	**Р** 0,040	**Я** 0,018	**Х** 0,009
О 0,090	**В** 0,038	**Ы** 0,016	**Ж** 0,007
Е,Ё 0,072	**Л** 0,035	**З** 0,016	**Ю** 0,006
А 0,062	**К** 0,028	**Ь,Ъ** 0,014	**Ш** 0,006
И 0,062	**М** 0,026	**Б** 0,014	**Ц** 0,003
Т 0,053	**Д** 0,025	**Г** 0,013	**Щ** 0,003
Н 0,053	**Г** 0,023	**Ч** 0,012	**Э** 0,003
С 0,045	**У** 0,021	**Й** 0,010	**Ф** 0,002

Таблица 6.5

Равновероятность и независимость букв	Учет различной вероятности	Учет зависимости между 2-мя буквами	Учет зависимости между 3-мя буквами	…	Учет зависимости между 8-ю буквами
5,04	4,39	3,52	3,05	…	2

Изучение последующих свойств энтропии требует введения понятий взаимной и условной энтропии для XY.

Взаимная энтропия интерпретируется как среднее количество взаимной информации, необходимой для определения событий $x_k \in X$ и $y_i \in Y$, взятых со своими вероятностями $p(x_k)$ и $p(y_i)$ соответственно:

$$H(X;Y) = \sum_{(x_k,y_i)\in XY} J(x_k;y_i)p(x_k y_i) = \sum_{(x_k,y_i)\in XY} p(x_k y_i)\log_2 \frac{p(x_k y_i)}{p(x_k)p(y_i)} =$$

$$= -\sum_{(x_k,y_i)\in XY} p(x_k y_i)\log_2 p(x_k) - \sum_{(x_k,y_i)\in XY} p(x_k y_i)\log_2 p(y_i) +$$

$$+ \sum_{(x_k,y_i)\in XY} p(x_k y_i)\log_2 p(x_k y_i) = H(X) + H(Y) - H(XY).$$

Аналогично выводятся формулы:

$$H(X;Y) = H(X) - H(X/Y) = H(Y) - H(Y/X);$$

$$H(XY) = H(X) + H(Y/X) = H(Y) + H(X/Y).$$

Условная энтропия вычисляется в соответствии со следующим выражением, что следует из определения энтропии:

$$H(Y/X) = -\sum_{(x_k, y_i)\in XY} p(x_k y_i)\log_2 p(y_i/x_k).$$

Отметим, если в системе нет потерь и искажений, то условные энтропии равны нулю.

6. Для заданного произведения множеств XY условная энтропия удовлетворяет неравенству $H(Y/X) \le H(Y)$, в котором знак равенства имеет место только тогда, когда события x_k и y_i статистически независимы, т.е. $p(y_i) = p(y_i/x_k)$. Данное неравенство показывает, что статистическая зависимость, например, символов алфавита не может увеличить количество информации на символ. Доказательство проведем уже известным способом:

$$H(Y/X) - H(Y) = -H(X;Y) = \sum_{(x_k, y_i)\in XY} p(x_k y_i)\log_2 \frac{p(y_i)}{p(y_i/x_k)} \le$$

$$\le \frac{1}{\ln 2}\sum_{(x_k, y_i)\in XY} p(x_k y_i)\left[\frac{p(y_i)}{p(y_i/x_k)} - 1\right] = \frac{1}{\ln 2}\left[\sum_{x_k\in X} p(x_k)\sum_{y_i\in Y} p(y_i) - 1\right] = 0.$$

7. Введем в рассмотрение ***условное среднее значение взаимной информации***:

$$J(X; y_i) = \sum_{x_k\in X} p(x_k/y_i) J(x_k; y_i) = \sum_{x_k\in X} p(x_k/y_i)\log_2 \frac{p(x_k/y_i)}{p(x_k)};$$

$$J(x_k; Y) = \sum_{y_i\in Y} p(y_i/x_k) J(x_k; y_i) = \sum_{y_i\in Y} p(y_i/x_k)\log_2 \frac{p(y_i/x_k)}{p(y_i)}.$$

8. Для заданного произведения множеств XY условное среднее значение взаимной информации величина неотрицательная, т.е. $J(X; y_i) \ge 0$, здесь знак равенства достигается при $p(x_k) = p(x_k/y_i)$. Данное свойство оказывается важным в том случае, когда y_i – символ, принятый на выходе канала с шумом, а x_k – какое-либо из сообщений. Неравенство устанавливает, что средняя информация, содержащаяся в принятом символе относительно переданного сообщения, всегда неотрицательна (рис. 6.10,а).

9. По аналогии $J(x_k;Y) \geq 0$. Это означает, что если x_k – сообщение, передаваемое по каналу с шумом, то средняя информация о сообщении x_k, полученная по этому каналу (рис. 6.10,б), неотрицательна.

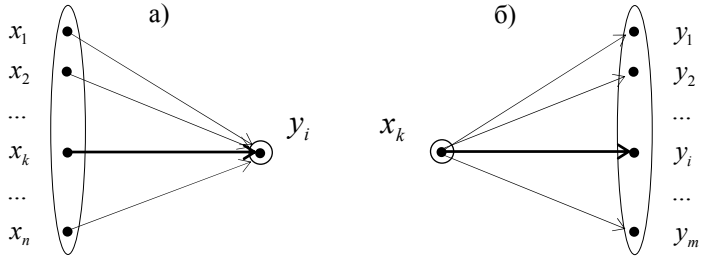


Рис. 6.10. Иллюстрация условного среднего значения взаимной информации:
а – $J(X;y_i)$; б – $J(x_k;Y)$

10. Взаимная энтропия всегда неотрицательна, т.е. $H(X;Y) \geq 0$.

Доказательство свойств 8, 9 и 10 проводится аналогично доказательству свойств 5 и 6.

Подведем некоторые итоги. Рассмотрим два различных выражения для взаимной энтропии:

$$H(X;Y) = H(X) - H(X/Y);$$
$$H(X;Y) = H(Y) - H(Y/X).$$

Эти два выражения особо интересны, когда x_k является сообщением, передаваемым по каналу с шумом, а y_i – соответствующим принятым символом.

При интерпретации первого соотношения мы понимаем:

– энтропию $H(X)$ как среднее количество переданной информации (энтропия источника);

– взаимную энтропию $H(X;Y)$ как среднее количество информации, полученной о переданном сообщении;

– условную энтропию $H(X/Y)$ как среднее количество информации, которое все еще потребуется для определения x_k после

приема символа y_i, т.е. потерянное вследствие шума, или короче, ***ненадежность передачи***.

Второе соотношение подчеркивает другой взгляд на среднее количество информации, а именно, как на разность между средним количеством информации, необходимым для определения принятого символа, и средним количеством информации, необходимым для определения того же сигнала, когда известно переданное сообщение. Следовательно, $H(Y/X)$ есть среднее количество информации, необходимое для определения помехи, имеющей место в канале, и мы можем понимать эту величину как энтропию шума в канале. Другими словами, $H(Y/X)$ является частью энтропии Y, которая возникает вследствие шума в канале.

ЗАДАЧИ

6.1. Урна содержит 5 черных и 10 белых шаров. Случайно, без возвращения, из урны выбирают три шара, и результат выбора передают по линии связи. Пусть выбрано: черный – черный – белый. Какое количество информации передается, если хотят сообщить:

а) о порядке шаров, их цвете, количестве белых и черных;

б) только о количестве черных и белых шаров?

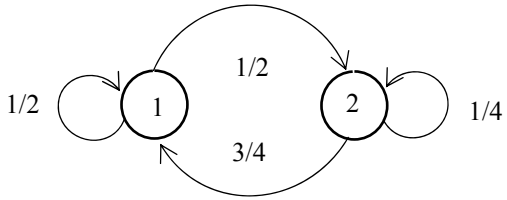
6.2. На борту самолета имеется дублированная вычислительная система, состоящая из двух одинаковых ЭВМ. Надежность одной ЭВМ $p = 0{,}96$. Сколько надо передать информации по радиоканалу, чтобы сообщить об отказе вычислительной системы?

6.3. Система описывается дискретным стационарным марковским процессом с матрицей переходов

$$\begin{array}{c|cc} & 1 & 2 \\ 1 & 1/2 & 1/2 \\ 2 & 3/4 & 1/4 \end{array}.$$

Сколько информации содержится в сообщении о том, что система находится в состоянии с номером "1"?

6.4. Система описывается дискретным стационарным марковским процессом, граф переходов которой имеет вид



Сколько информации содержится в сообщении о том, что система находится в состоянии с номером "2"?

6.5. Сколько информации содержится в сообщении о том, что сумма очков на двух брошенных игральных костях равна 7?

6.6. На двухэлементном множестве $X = (x_1, x_2)$ задано распределение вероятностей $p(x_1)$ и $p(x_2)$. Построить зависимость $H(X) = F[p(x_1), p(x_2)]$.

6.7. Ансамбли событий $X = (x_1, x_2, x_3)$ и $Y = (y_1, y_2)$ объединены. Вероятности совместных событий $p(x_i y_j)$ имеют вид

y_j	x_i		
	x_1	x_2	x_3
y_1	0,1	0,2	0,3
y_2	0,25	0	0,15

Определить:

а) энтропии ансамблей X, Y и XY;

б) условные энтропии $H(X/Y)$ и $H(Y/X)$;

в) взаимную энтропию $H(X;Y)$;

г) условные средние значения взаимной информации $J(X; y_i)$ и $J(x_k; Y)$;

д) частную условную энтропию $H(Y/x_k)$.

6.8. Пусть x_1, x_2 – буквы алфавита на входе канала, а y_1, y_2, y_3 – буквы выходного алфавита. Символы на входе канала появляются равновероятно, а последовательность букв во входной и выходной

последовательностях статистически независима. Условные вероятности имеют вид

$$\left\| p(y_i / x_k) \right\| = \begin{matrix} x_1 \\ x_2 \end{matrix} \begin{vmatrix} 1/32 & 61/64 & 1/64 \\ 1/32 & 1/64 & 61/64 \end{vmatrix} .$$

Построить функцию распределения для взаимной информации $I(x_k; y_i)$. Определить взаимную энтропию $H(X;Y)$.

6.9. Группа из четырех сообщений кодируется двоичным кодом

Сообщение u_k	Вероятность сообщения $p(u_k)$	Двоичный код
u_1	1/2	0
u_2	1/4	10
u_3	1/8	110
u_4	1/8	111

Найти энтропию множества сообщений и среднюю длину кодовых слов. Показать, что если сообщения статистически независимы, то при передаче последовательности кодовых слов символы “0” и “1” появляются равновероятно.

6.10. Среди группы ПВО ¼ составляют зенитные орудия (x_1), ½ – ракетные установки (x_2), ¼ – самолеты-перехватчики (x_3). Пусть ракетные установки с вероятностью ½ поражают цель, самолеты-перехватчики всегда поражают цель, а зенитные орудия всегда не попадают.

Какое количество информации содержится в сообщении *y*: «средство ПВО поразило цель относительно факта использования x_k»?

Какое количество информации содержится в сообщении *z*: «средство ПВО три раза подряд поразило цель относительно факта использования x_2»?

6.11. От источника информации поступают статистически независимые символы с вероятностями 1/4 и 3/4. Найти энтропию последовательности из 100 двоичных символов.

7. КОДИРОВАНИЕ СООБЩЕНИЙ ДИСКРЕТНОГО МНОЖЕСТВА

Перейдем к задаче ***кодирования***, т.е. представления сообщений из заданного дискретного множества последовательностью символов, принадлежащих заданному алфавиту. Цель при этом состоит в конструировании таких последовательностей символов, которые минимизируют среднее число символов, приходящихся на одно сообщение. Усреднение ведется по множеству сообщений. При решении поставленной задачи предполагается, что распределение вероятностей сообщений известно.

Рассмотрим множество сообщений $U = \{u_k\}, k = \overline{1, M}$, на котором задано распределение вероятностей $p(u_k)$. Каждое сообщение может быть представлено кодовым словом длиной n_k (рис. 7.1). Слово состоит из последовательности символов некоторого алфавита A кодирования. Число различных символов этого алфавита обозначим через $D = |A|$.

При этих обозначениях среднее число символов, приходящихся на одно сообщение, есть $\bar{n} = \sum_{k=1}^{M} n_k p(u_k)$. Информационная пропускная способность алфавита A равна $\log_2 D$.

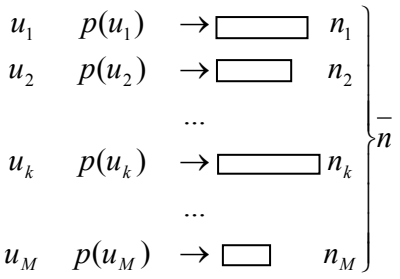


Рис. 7.1. Процесс кодирования

Энтропия $H(U)$ – это среднее количество информации, необходимое для однозначного определения, идентификации сообщения

из данного множества. Каждое сообщение содержит в среднем $\bar{n}$ символов, а каждый символ не может нести более чем $\log_2 D$ информации. Тогда энтропию множества сообщений U можно оценить следующим образом:

$$H(U) = -\sum_{k=1}^{M} p(u_k) \log_2 p(u_k) \le \bar{n} \log_2 D .$$

Следовательно, $\bar{n} \ge \dfrac{H(U)}{\log_2 D}$, т.е. ***среднее число символов на сообщение*** не может быть меньше энтропии множества сообщений, деленной на пропускную способность алфавита, используемого для кодирования.

7.1. МЕТОД КОДИРОВАНИЯ ШЕННОНА–ФАНО

Данный ***метод кодирования*** кодовых слов со средней длиной, достаточно близкой к ***нижней границе*** $\dfrac{H(U)}{\log_2 D}$, базируется на выполнении двух требований:

1. В каждой позиции кодового слова символы алфавита должны использоваться с равными вероятностями, что дает возможность максимизировать среднее количество информации на символ.

2. Вероятности появления символов в каждой позиции не должны зависеть от расположения всех предыдущих символов.

Если удастся построить код в точном соответствии с этими требованиями, то средняя длина кодовых слов будет в точности равна нижней границе.

Для реализации кодов с вышеописанными свойствами необходимо иметь соответствующие процедуры построения кодов. Рассмотрим ряд примеров. Имеется множество равновероятных сообщений $U = \{u_k\}, k = \overline{1,8}$ и алфавит $D = 2$. Предлагается следующая процедура кодирования:

– первоначальное множество сообщений разбивается на две равновероятные группы;

– первым символом кодовых слов для сообщений первой группы выбираем "0", а первым символом кодов для сообщений второй группы выбираем "1";

– далее процесс половинного деления с соответствующей кодировкой продолжается до логического конца, т.е. до тех пор, пока это возможно.

В результате получается двоичный код для всех сообщений. В табл. 7.1 показан описанный процесс.

Таблица 7.1

Сообщения u_k	Вероятности $p(u_k)$	Символы разрядов кодовых слов			Код Шеннона–Фано
		1	2	3	
u_1	1/8	0	0	0	000
u_2	1/8			1	001
u_3	1/8		1	0	010
u_4	1/8			1	011
u_5	1/8	1	0	0	100
u_6	1/8			1	101
u_7	1/8		1	0	110
u_8	1/8			1	111

Для построенного кода среднее число символов алфавита в коде равно $\overline{n} = 3$. Энтропия множества сообщений равна $H(U) = 8\frac{1}{8}\log_2 8 = 3$ бит. Пропускная способность двоичного алфавита равна $\log_2 2 = 1$ бит. В результате применения этой процедуры кодирования удалось достичь нижней границы $\overline{n} = \frac{H(U)}{\log_2 D} = 3$.

В дальнейшем под ***эффективностью кодирования*** будем понимать отношение нижней границы к средней длине кодового слова

$$Э_к = \frac{H(U)}{\overline{n}\log_2 D}.$$

Совершенно очевидно, что эффективность кодирования лежит в диапазоне $0 \le Э_к \le 1$, а равенству $\bar{n} = \dfrac{H(U)}{\log_2 D}$ соответствует эффективность $Э_к = 1$.

Рассмотрим случай, когда сообщения неравновероятны, но вероятности их появления пропорциональны отрицательным степеням двойки. При применении процедуры Шеннона–Фано целесообразно первоначальное множество расположить в порядке убывания вероятностей и именно в такой последовательности пытаться проводить разбиения на две примерно равновероятные группы. Как показывает практика построения кодов, это обеспечивает наибольшую эффективность кодирования, так как в каждой группе будут фигурировать сообщения с равновеликими вероятностями. Методика кодирования аналогична предыдущему примеру. Процесс кодирования сведен в табл. 7.2.

Из таблицы видно, что число символов в каждом слове кода равно собственной информации сообщения. Средняя длина кодовых слов равна $\bar{n} = 1\dfrac{1}{2} + 2\dfrac{1}{8}3 + 4\dfrac{1}{16}4 = 2{,}25$, а энтропия ансамбля сообщений $H(U) = -1\dfrac{1}{2}\log_2\dfrac{1}{2} - 2\dfrac{1}{8}\log_2\dfrac{1}{8} - 4\dfrac{1}{16}\log_2\dfrac{1}{16} = 2{,}25$ бит. Для построенного кода среднее число двоичных символов равно энтропии ансамбля сообщений, т.е. код также оптимален в смысле минимума средней длины ($Э_к = 1$).

Метод кодирования Шеннона–Фано легко обобщить на случай произвольного алфавита из D символов путем последовательного и равновероятного разбиения множества сообщений на D групп, подгрупп и т.д. Ясно также, что метод будет успешно работать, если $p(u_k) = D^{-\nu}$ (ν – целое число), т.е. вероятности являются отрицательными степенями числа символов алфавита, что обеспечивает равновероятность появления символов.

Если $p(u_k) \ne D^{-\nu}$, то группы не будут точно равновероятны. В этом случае метод Шеннона–Фано не обязательно привет к коду с наименьшей возможной средней длиной. Более того, описанный метод построения кода Шеннона–Фано не всегда приводит к одно-

значному построению кода. Ведь при разбиении на подгруппы можно сделать большей по вероятности как верхнюю, так и нижнюю подгруппу.

Таблица 7.2

Сообщения u_k	Вероятности $p(u_k)$	Символы разрядов кодовых слов				Код Шеннона–Фано
		1	2	3	4	
u_1	1/2	0				0
u_2	1/8	1	0	0		100
u_3	1/8			1		101
u_4	1/16		1	0	0	1100
u_5	1/16				1	1101
u_6	1/16			1	0	1110
u_7	1/16				1	1111

Наглядное представление множества кодовых слов можно получить с использованием ***кодового дерева***, метод образования которого для ранее приведенного примера ясен из рис. 7.2.

При построении кодового дерева устанавливается соответствие между сообщениями и концевыми узлами (вершинами). Требование, чтобы только концевые узлы дерева соответствовали сообщениям, эквивалентно выполнению свойства префикса.

Для того чтобы определить свойство префикса, допустим, что k-е кодовое слово представляется как $\overline{a_k} = (a_{k1},...,a_{kn_k})$, где $a_{k1},...,a_{kn_k}$ – отдельные символы, составляющие кодовое слово. Любая последовательность, составленная из начальной части $\overline{a_k}$, т.е. $a_{k1},...,a_{ki}$ для некоторого $i < n_k$, называется i-м префиксом $\overline{a_k}$.

ОПРЕДЕЛЕНИЕ 7.1. Код, обладающий ***свойством префикса***, определяется как код, в котором никакое кодовое слово не является i-м префиксом никакого другого кодового слова.

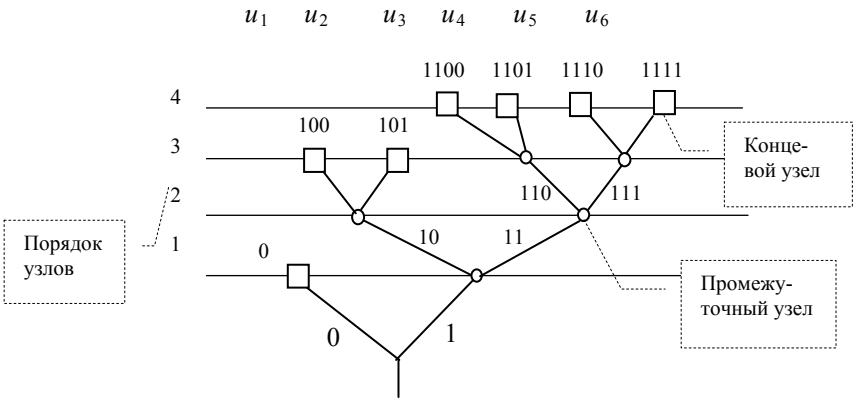


Рис. 7.2. Кодовое дерево кода Шеннона–Фано

Особенностью такого кода является то, что ни одна из более коротких кодовых комбинаций не повторяет начала более длинной комбинации. Это позволяет при передаче длинных последовательностей не использовать разделяющих интервалов или синхронизирующих импульсов для фиксации конца и начала кодовых слов. Всякий раз, когда последовательность двоичных цифр образует кодовую комбинацию, декодер приемника декодирует ее как сообщение и воспринимает следующую цифру, как начало нового сообщения. Пусть для примера передана двоичная последовательность 0010111010100. По таблице кода легко находим, что переданы сообщения $u_1 - u_1 - u_3 - u_5 - u_1 - u_2$.

ОПРЕДЕЛЕНИЕ 7.2. ***Полным кодовым деревом*** назовем такое конечное кодовое дерево, в котором из каждого промежуточного узла исходят D узлов следующего более высокого порядка.

УТВЕРЖДЕНИЕ 7.1. Наличие одного концевого узла порядка n_k исключает D^{n-n_k} узлов порядка n $(n_k \le n)$, где $n = \max\{n_1, n_2, ..., n_M\}$.

Рассмотрим принцип исключения узлов на примере частного случая полного двоичного кодового дерева (рис. 7.3). Наличие одного концевого узла второго порядка исключает $2^{4-2} = 4$ узлов четвертого порядка.

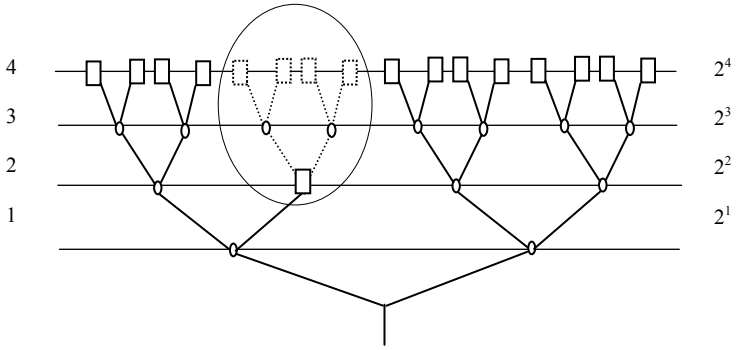


Рис. 7.3. Принцип исключения концевых узлов

7.2. ОСНОВОПОЛАГАЮЩИЕ ТЕОРЕМЫ ОПТИМАЛЬНОГО КОДИРОВАНИЯ

ТЕОРЕМА 7.1 (***неравенство Крафта***). Пусть D – число символов в кодовом алфавите, а $\{n_k\}, k=\overline{1,M}$ – заданное множество целых положительных чисел. Неравенство $\sum_{k=1}^{M} D^{-n_k} \le 1$ является необходимым и достаточным условием существования кодовых слов, соответствующих концевым узлам дерева, с длинами равными n_k.

ДОКАЗАТЕЛЬСТВО

НЕОБХОДИМОСТЬ. Построено кодовое дерево с узлами порядка $n_1, n_2, ..., n_M$. Кодовый алфавит состоит из D символов. Поэтому из каждого промежуточного узла дерева исходит не более D ветвей. Отсюда следует, что не может быть более чем D^n узлов порядка $n = \max\{n_1, n_2, ..., n_M\}$. Это максимальное число получается тогда, когда D ветвей, разветвляются из каждого узла, порядок которых меньше n. Наличие концевого узла порядка $n_k \le n$ исключает D^{n-n_k} возможных узлов порядка n. Следовательно, $\sum_{k=1}^{M} D^{n-n_k} \le D^n$. Делим полученное равенство на D^n и приходим к тому, что $\sum_{k=1}^{M} D^{-n_k} \le 1$.

ДОСТАТОЧНОСТЬ. Дано, что $\sum_{k=1}^{M} D^{-n_k} \le 1$. Необходимо показать, что можно построить кодовое дерево с заданными концевыми узлами, т.е. множество кодовых слов с заданными длинами $n_1, n_2, ..., n_M$. Пусть числа расположены в порядке возрастания, так что для любого k: $n_k \le n_{k+1}$. Предположим, что нам удалось построить часть кодового дерева, содержащего все заданные концевые узлы порядков $n_1, n_2, ..., n_j$, меньших m, и что дерево должно содержать N_m концевых узлов порядка m. Тогда из неравенства Крафта $\sum_{k=1}^{M} D^{-n_k} \le 1$ следует: $\sum_{k=1}^{j} D^{-n_k} + N_m D^{-m} + \sum_{k=j+N_m+1}^{M} D^{-n_k} \le 1$.

Умножаем последнее выражение на D^m и получаем неравенство $N_m + \sum_{k=1}^{j} D^{m-n_k} + \sum_{k=j+N_m+1}^{M} D^{m-n_k} \le D^m$. Из него видно, что число доступных узлов $D^m - \sum_{k=1}^{j} D^{m-n_k}$ порядка m не меньше заданного числа узлов N_m того же порядка, т.е. $D^m - \sum_{k=1}^{j} D^{m-n_k} \ge N_m + \sum_{k=j+N_m+1}^{M} D^{m-n_k}$, и поэтому все они могут быть включены в дерево. Поскольку в доказательстве m выбрано произвольно, дерево с требуемыми концевыми узлами всегда может быть построено.

ТЕОРЕМА 7.2. Пусть D – число символов в кодовом алфавите, а $\{n_k\}, k = \overline{1, M}$ – заданное множество целых положительных чисел. Равенство $\sum_{k=1}^{M} D^{-n_k} = 1$ является необходимым и достаточным условием того, чтобы заданное множество концевым узлов дерева было полным.

ДОКАЗАТЕЛЬСТВО

НЕОБХОДИМОСТЬ. Построено кодовое дерево с узлами порядка $n_1, n_2, ..., n_M$. Если множество концевых узлов полно, то они долж-

ны исключить все узлы порядка $n = \max\{n_1, n_2, ..., n_M\}$, т.е. $\sum_{k=1}^{M} D^{n-n_k} = D^n$. Откуда следует: $\sum_{k=1}^{M} D^{-n_k} = 1$.

ДОСТАТОЧНОСТЬ. Дано, что $\sum_{k=1}^{M} D^{-n_k} = 1$. Любое множество узлов, удовлетворяющее условию $\sum_{k=1}^{M} D^{-n_k} = 1$, удовлетворяет и неравенству $\sum_{k=1}^{M} D^{-n_k} \le 1$, т.е. существует, по крайней мере, одно дерево с заданными концевыми узлами порядков $n_1, n_2, ..., n_M$. Предположим, что это дерево имеет, кроме заданных, еще узлы. Если бы это было так, то мы могли бы, не нарушая неравенства $\sum_{k=1}^{M} D^{-n_k} \le 1$, добавить в сумму $\sum_{k=1}^{M} D^{-n_k}$ члены, соответствующие этим узлам. Очевидно, что это невозможно, так как дерево, удовлетворяющее условию $\sum_{k=1}^{M} D^{-n_k} = 1$, не может иметь дополнительных узлов.

ТЕОРЕМА 7.3 (***основная теорема кодирования***). При заданном множестве сообщений $\{u_k\}, k = \overline{1, M}$ с энтропией $H(U)$ и алфавитом, состоящим из D символов, существует такой способ кодирования сообщений множества посредством последовательностей символов, принадлежащих заданному алфавиту, что среднее число символов на сообщение $\bar{n} = \sum_{k=1}^{M} n_k p(u_k)$ будет удовлетворять неравенствам $\frac{H(U)}{\log_2 D} \le \bar{n} < \frac{H(U)}{\log_2 D} + 1$.

ДОКАЗАТЕЛЬСТВО. Докажем неравенство $H(U) - \bar{n}\log_2 D \le 0$. Пусть $p(u_1), ..., p(u_M)$ – вероятности сообщений источника, а $n_1, ..., n_M$ – длины кодовых слов. Представим $H(U) - \bar{n}\log_2 D$ в виде $\sum_{k=1}^{M} p(u_k)\log_2 \frac{1}{p(u_k)} - \sum_{k=1}^{M} p(u_k) n_k \log_2 D$.

Поместим $-n_k$ под знак логарифма, объединим слагаемые и получим $\sum_{k=1}^{M} p(u_k)\log_2 \frac{D^{-n_k}}{p(u_k)}$. Далее воспользуемся известным свойством натурального логарифма $\ln a \le a-1$:

$$H(U)-\bar{n}\log_2 D \le \frac{1}{\ln 2}\sum_{k=1}^{M} p(u_k)[\frac{D^{-n_k}}{p(u_k)}-1] =$$

$$= \frac{1}{\ln 2}[\sum_{k=1}^{M} D^{-n_k} - \sum_{k=1}^{M} p(u_k)] \le 0.$$

Последнее неравенство следует из неравенства Крафта, что и доказывает $\frac{H(U)}{\log_2 D} \le \bar{n}$. Заметим, что равенство имеет место тогда и только тогда, когда $p(u_k) = D^{-n_k}$. Это условие совпадает с ранее полученным условием оптимального кодирования по Шеннону–Фано.

Докажем неравенство $\bar{n} < \frac{H(U)}{\log_2 D} + 1$. Если бы длины кодовых слов не обязательно были целыми числами, то можно было бы просто подобрать n_k^*, чтобы удовлетворить условию $p(u_k) = D^{-n_k^*}, k = \overline{1,M}$. Однако удовлетворить это условие можно только приближенно, выбирая целые числа n_k так, чтобы выполнялись неравенства

$$D^{-n_k} \le p(u_k) < D^{-n_k+1}, \quad k = \overline{1,M}$$

(рис. 7.4). Суммирование по k превращает левое неравенство в неравенство Крафта, и, следовательно, существует префиксный код с этими длинами. Логарифмируя правое неравенство, получаем $\log_2 p(u_k) < (-n_k+1)\log_2 D$ или $n_k < \frac{-\log_2 p(u_k)}{\log_2 D} + 1$. Далее полученное неравенство умножаем на $p(u_k)$ и суммируем по k

$\sum_{k=1}^{M} n_k p(u_k) < \sum_{k=1}^{M} p(u_k)[\frac{-\log_2 p(u_k)}{\log_2 D} + 1]$, $\bar{n} < \frac{H(U)}{\log_2 D} + 1$, что и завершает доказательство теоремы.

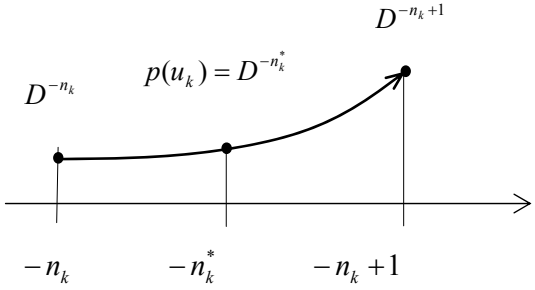


Рис. 7.4. Графическая интерпретация неравенств $D^{-n_k} \le p(u_k) < D^{-n_k+1}$

7.3. МЕТОД ХАФФМАНА (ОПТИМАЛЬНОЕ КОДИРОВАНИЕ)

Метод кодирования и построения кодового дерева, основанный на равновероятном использовании символов алфавита и использующий статистическую независимость символов, как было показано выше, приводит к коду с минимальной средней длиной слова $\bar{n} = \frac{H(U)}{\log_2 D}$ не всегда, а только в случае выполнения равенства $p(u_k) = D^{-n_k}$.

Существует систематический метод оптимального кодирования, который приводит к оптимальному множеству кодовых слов в том смысле, что никакое множество других кодовых слов не имеет меньшего среднего числа символов на сообщение. Соответствующее кодирование предложено Хаффманом, и получаемый по этому методу код носит название кода Хаффмана. В процедуре Хаффмана все шаги строго регламентированы и никаких неоднозначностей кодирования нет. Произвольное присвоение символов алфавита кодировки различным ветвям дерева не нарушает общую структуру оптимального кода.

Процедура ***оптимального кодирования*** сообщений заданного множества и произвольного алфавита представлена на рис. 7.5 в виде алгоритма, который не требует дополнительных пояснений.

Рассмотрим пример построения кода Хаффмана для множества сообщений, представленных в табл. 7.3. Будем использовать алфавит кодировки, состоящий из $D = 2$ символов, для которого m_0 всегда равно 2. Соответствующая процедура формирования кодовых слов приведена на рис. 7.6. Эффективность кодирования составила $Э_к = \frac{H(U)}{\bar{n}\log_2 D} = \frac{2{,}55}{2{,}60} = 0{,}98$. Построенное кодовое дерево является полным.

Таблица 7.3

Сообщения u_k	Вероятности $p(u_k)$	Код Хаффмана
u_1	0,3	00
u_2	0,2	10
u_3	0,2	11
u_4	0,1	011
u_5	0,1	0100
u_6	0,05	01010
u_7	0,05	01011
$H(U) = 2{,}55$	$\sum_U p(u_k) = 1$	$\bar{n} = 2{,}6$

Рассмотрим другой пример построения кода Хаффмана для множества сообщений, представленных в табл. 7.4. Будем использовать алфавит кодировки, состоящий из трех символов. В этом случае процедура Хаффмана предполагает на первом шаге объединение $m_0 = 2$ сообщений с наименьшими вероятностями и $D = 3$ на последующих шагах (рис. 7.7). Эффективность кодирования составила $Э_к = \frac{H(U)}{\bar{n}\log_2 D} = \frac{2{,}45}{1{,}7 \cdot 1{,}58} = 0{,}91$. Построенное кодовое дерево является неполным.

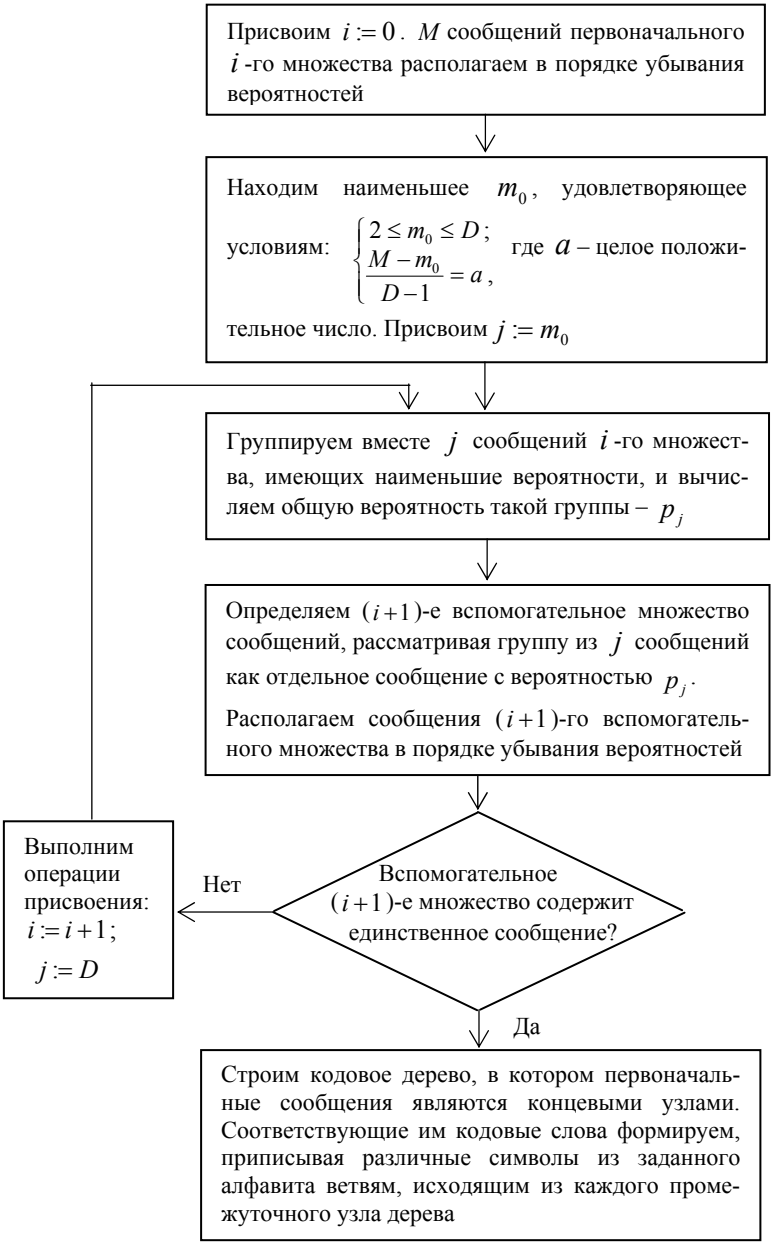


Рис. 7.5. Алгоритм процедуры Хаффмана

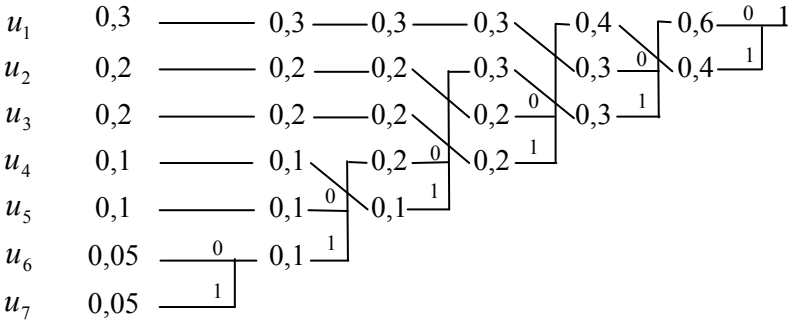


Рис. 7.6. Процедура оптимального кодирования для $D = 2$

Таблица 7.4

Сообщения u_k	Вероятности $p(u_k)$	Код Хаффмана
u_1	0,3	1
u_2	0,2	2
u_3	0,2	00
u_4	0,1	02
u_5	0,1	010
u_6	0,1	011
$H(U) = 2{,}45$	$\sum_U p(u_k) = 1$	$\bar{n} = 1{,}7$

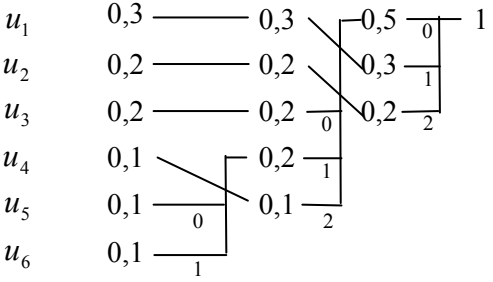


Рис. 7.7. Процедура оптимального кодирования для $D = 3$

ЗАДАЧИ

7.1. Для заданного ансамбля сообщений построить код Шеннона–Фано $(D = 2)$:

x_k	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$p(x_k)$	0,3	0,2	0,2	0,1	0,1	0,05	0,05

Оценить эффективность кодирования. Построить кодовые деревья.

7.2. Найти оптимальный троичный код для ансамбля сообщений

x_k	x_1	x_2	x_3	x_4	x_5	x_6
$p(x_k)$	0,3	0,25	0,13	0,12	0,11	0,09

Построить кодовое дерево. Оценить эффективность кодирования.

7.3. Задано множество сообщений

x_k	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$p(x_k)$	0,4	0,17	0,11	0,09	0,08	0,06	0,05	0,04

Закодировать сообщения с использованием кодов Хаффмана и Шеннона–Фано $(D = 3)$. Сравнить эффективности кодирования.

8. ДИСКРЕТНЫЕ ИСТОЧНИКИ ИНФОРМАЦИИ

Под дискретным источником информации будем понимать источник, который порождает конечное число сообщений. К этому типу источников можно отнести книжный текст, множество телеграфных сообщений и т.п. Соответственно, источники, на выходе которых появляется непрерывно меняющаяся величина, естественно называть непрерывным источником информации.

До сих пор рассматривались сообщения $u_k \in U$ безотносительно к тому, что на выходе источника во времени появляется последовательность сообщений, а энтропия вычислялась как среднее по ансамблю $U = \{u_k\}, k = \overline{1, M}$, т.е.

$$H(U) = -\sum_{k=1}^{M} p(u_k) \log_2 p(u_k) \le \log_2 M .$$

8.1. ДИСКРЕТНЫЕ ИСТОЧНИКИ, ПОРОЖДАЮЩИЕ СТАТИСТИЧЕСКИ НЕЗАВИСИМЫЕ СООБЩЕНИЯ

Разделим порождаемую источником последовательность сообщений на подпоследовательности равной длины, каждая из которых содержит ν сообщений (рис. 8.1). Число различных последовательностей из ν сообщений равно M^{ν}. Формула автоматически следует из выражения для числа перестановок из M по ν с учетом повторений $P_M^{\nu} = M^{\nu}$. Каждая ν - я последовательность является элементом декартового произведения множеств U^{ν}, причем $\left|U^{\nu}\right| = M^{\nu}$.

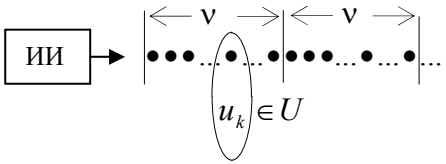


Рис. 8.1. Разбиение последовательности независимых сообщений на группы

ТЕОРЕМА 8.1. При любом заданном как угодно малом положительном числе $\varepsilon > 0$ можно найти натуральное число ν и соответствующее множество кодовых слов мощности M^{ν}, такое, что среднее число символов алфавита кодировки на сообщение удовлетворяет неравенству $\bar{n} < \dfrac{H(U)}{\log_2 D} + \varepsilon$.

ДОКАЗАТЕЛЬСТВО. Используя символы заданного алфавита кодировки, закодируем последовательности сообщений длиной ν кодовыми словами с длинами, удовлетворяющими неравенству Крафта (рис. 8.2). Пусть $\overline{n_\nu}$ – среднее число символов на кодовое слово, а $\bar{n} = \dfrac{\overline{n_\nu}}{\nu}$ – среднее число символов на сообщение. Тогда из основной теоремы кодирования следует неравенство для $\overline{n_\nu}$: $\dfrac{H(U^\nu)}{\log_2 D} \le \overline{n_\nu} < \dfrac{H(U^\nu)}{\log_2 D} + 1$. В силу статистической независимости сообщений, порождаемых источником, $H(U^\nu) = H(U \times U \times \ldots \times U) = \nu H(U)$ имеем:

$$\frac{\nu H(U)}{\log_2 D} \le \overline{n_\nu} < \frac{\nu H(U)}{\log_2 D} + 1;$$

$$\frac{H(U)}{\log_2 D} \le \bar{n} = \frac{\overline{n_\nu}}{\nu} < \frac{H(U)}{\log_2 D} + \frac{1}{\nu}.$$

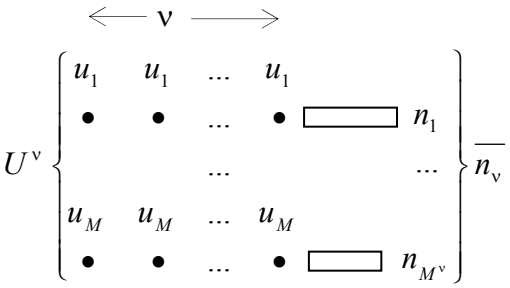


Рис. 8.2. Кодирование последовательностей сообщений длиной ν

Положим $\varepsilon \geq \frac{1}{\nu}$, тогда $\bar{n} < \frac{H(U)}{\log_2 D} + \varepsilon$.

Основной же вывод состоит в том, что если кодируется последовательность статистически независимых сообщений вместо кодировки отдельных сообщений, то это дает возможность уменьшить среднюю величину $\bar{n}$ не более чем на один символ.

Докажем соотношение $H(U^\nu) = \nu H(U)$, использованное выше, для простого случая $|U| = 2$. Пусть $p(u_k) = p_k, k \in \{1,2\}$. Образуем последовательности из ν сообщений, общее число которых составит 2^ν. Ниже приводятся подробные соответствующие преобразования, не представляющие серьезных трудностей для понимания:

$$H(U^\nu) = -\sum_{k=0}^{\nu}\binom{\nu}{k} p_1^{\nu-k} p_2^k \log_2 p_1^{\nu-k} p_2^k = -\sum_{k=0}^{\nu}\binom{\nu}{k} p_1^{\nu-k} p_2^k (\nu-k)\log_2 p_1 -$$

$$-\sum_{k=0}^{\nu}\binom{\nu}{k} p_1^{\nu-k} p_2^k k \log_2 p_2 = -p_1 \log_2 p_1 \sum_{k=0}^{\nu}\binom{\nu}{k} p_1^{\nu-k-1} p_2^k (\nu-k) -$$

$$-p_2 \log_2 p_2 \sum_{k=0}^{\nu}\binom{\nu}{k} p_1^{\nu-k} p_2^{k-1} k = -p_1 \log_2 p_1 \sum_{k=0}^{\nu}\frac{\nu!(\nu-k)}{k!(\nu-k)!} p_1^{\nu-k-1} p_2^k -$$

$$-p_2 \log_2 p_2 \sum_{k=0}^{\nu}\frac{\nu! k}{k!(\nu-k)!} p_1^{\nu-k} p_2^{k-1} = -\nu p_1 \log_2 p_1 \sum_{k=0}^{\nu-1}\binom{\nu-1}{k} p_1^{\nu-k-1} p_2^k -$$

$$-\nu p_2 \log_2 p_2 \sum_{k=1}^{\nu}\binom{\nu-1}{k-1} p_1^{\nu-k} p_2^{k-1} = -\nu[(p_1+p_2)^{\nu-1} p_1 \log_2 p_1 +$$

$$+(p_1+p_2)^{\nu-1} p_2 \log_2 p_2] = \nu[-p_1 \log_2 p_1 - p_2 \log_2 p_2] = \nu H(U).$$

8.2. СЛУЧАЙНЫЕ ДИСКРЕТНЫЕ ИСТОЧНИКИ

Пусть $\alpha_1\alpha_2 ... \alpha_i ... \alpha_\nu$ – последовательность статистически зависимых событий на выходе источника. Каждое конкретное событие α_i может быть одной из L букв $a_1 a_2 ... a_k ... a_L$ множества $A_i = \{a_k\}, k = \overline{1, L}$ (рис. 8.3). Уменьшение избыточности, обуслов-

ленной наличием статистических связей между буквами, может быть достигнуто кодированием не отдельных элементарных букв, а составленных из них групп длиной ν. Подобный метод уменьшения межбуквенных статистических связей часто называют ***декорреляцией***.

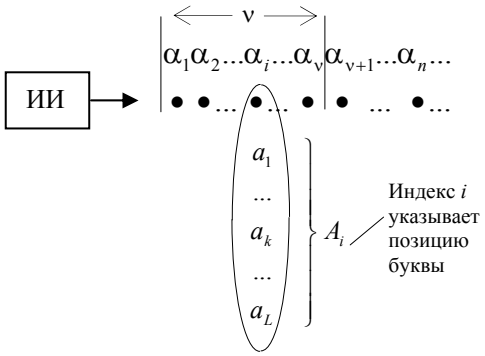


Рис. 8.3. Разбиение последовательности зависимых событий на группы

Прежде чем переходить к кодированию, дадим математическое описание случайных источников. Множество, порождаемых источником последовательностей, описывается совокупностью условных вероятностей:

$$p(\alpha_1) = p\{\alpha_1 = a_k\} \to L;$$

$$p(\alpha_2 / \alpha_1) = p\{\alpha_2 = a_j / \alpha_1 = a_k\} \to L^2;$$

$$p(\alpha_3 / \alpha_1\alpha_2) = p\{\alpha_3 = a_d / \alpha_1 = a_k, \alpha_2 = a_j\} \to L^3;$$

...

$$p(\alpha_i / \alpha_1\alpha_2...\alpha_{i-1}) = p\{\alpha_i = a_m / \alpha_1 = a_k, \alpha_2 = a_j, ..., \alpha_{i-1} = a_s\} \to L^i.$$

Совместная вероятность появления на выходе источника первых ν событий есть $p(\alpha_1\alpha_2...\alpha_\nu) = p(\alpha_1)p(\alpha_2/\alpha_1)...p(\alpha_\nu/\alpha_1...\alpha_{\nu-1})$, а совместное распределение $(\nu-1)$-го и ν-го событий равно

$$p(\alpha_{\nu-1}\alpha_\nu) = \sum_{A_1}...\sum_{A_{\nu-2}} p(\alpha_1...\alpha_{\nu-2}\alpha_{\nu-1}\alpha_\nu).$$

Описание последовательностей событий через условные вероятности приводит к следующим формулам определения энтропии события на выходе источника:

$$\underset{\alpha_1\in A_1}{\alpha_1} \rightarrow H(A_1) = -\sum_{\alpha_1\in A_1} p(\alpha_1)\log_2 p(\alpha_1);$$

$$\underset{\alpha_1\in A_1}{\alpha_1}\ \underset{\alpha_2\in A_2}{\alpha_2} \rightarrow H(A_2/A_1) = -\sum_{\alpha_1\in A_1}\sum_{\alpha_2\in A_2} p(\alpha_1\alpha_2)\log_2 p(\alpha_2/\alpha_1);$$

...

$$\underset{\alpha_1\in A_1}{\alpha_1}\ \underset{\alpha_2\in A_2}{\alpha_2}\ \ldots\ \underset{\alpha_i\in A_i}{\alpha_i} \rightarrow H(A_i/A_1\ldots A_{i-1}) =$$

$$= -\sum_{\alpha_1\in A_1}\sum_{\alpha_2\in A_2}\ldots\sum_{\alpha_i\in A_i} p(\alpha_1\alpha_2\ldots\alpha_i)\log_2 p(\alpha_i/\alpha_1\alpha_2\ldots\alpha_{i-1}).$$

Обозначим через $H(A_i/A^{\nu})$ условную энтропию i-го события при условии, что ν предшествующих событий уже произошло (рис. 8.4):

$$H(A_i/A^{\nu}) = -\sum_{A_{i-\nu}}\ldots\sum_{A_{i-1}}\sum_{A_i} p(\alpha_{i-\nu}\ldots\alpha_{i-1}\alpha_i)\log_2 p(\alpha_i/\alpha_{i-\nu}\ldots\alpha_{i-1}) =$$

$$= H(A_i/A_{i-1}\ldots A_{i-\nu}).$$

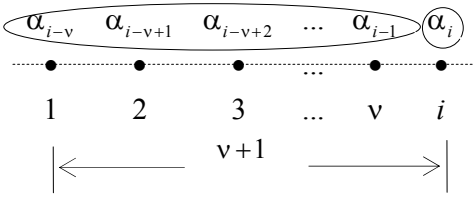


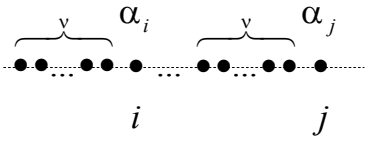
Рис. 8.4. Иллюстрация условной энтропии $H(A_i/A^{\nu})$

ОПРЕДЕЛЕНИЕ 8.1. Источник информации называют ***стационарным***, если условная вероятность появления события α_i при условии задания ν предшествующих событий не зависит от номера i для всех целых ν, т.е.

$$p(\alpha_i/\alpha_{i-1}\ldots\alpha_{i-\nu}) = p(\alpha_j/\alpha_{j-1}\ldots\alpha_{j-\nu});\ i, j > \nu.$$

Таким образом, для стационарных источников все вероятности инвариантны относительно любых сдвигов вдоль последовательности событий $\alpha_1\alpha_2...\alpha_i...\alpha_\nu$ или, говоря другими словами, статистические характеристики процесса появления событий на выходе источника не зависят от выбора начала отсчета времени или наблюдения (рис. 8.5).

Для стационарных источников справедливо равенство $H(A_i / A^\nu) = H(A_j / A^\nu)$ для всех $i, j > \nu$.



$$p\{\alpha_i = a_k / \alpha_{i-1} = a_m,..., \alpha_{i-\nu} = a_s\} =$$
$$= p\{\alpha_j = a_k / \alpha_{j-1} = a_m,..., \alpha_{j-\nu} = a_s\}$$

Рис. 8.5. Иллюстрация свойства стационарности

ТЕОРЕМА 8.2. Для стационарного источника условные энтропии событий $H(A_{\nu+1} / A^\nu)$ при условии, что заданы все предшествующие события, образуют монотонную невозрастающую последовательность.

ДОКАЗАТЕЛЬСТВО. При рассмотрении свойств энтропии доказано, что

$$H(A_i / A^\nu) \le H(A_i / A^{\nu-1}) \le H(A_i / A^{\nu-2}) \le ... \le H(A_i / A^1) \le H(A_i).$$

В силу стационарности число i можно заменить на любое целое, лишь бы это целое было больше числа предшествующих событий. Положим в первом члене неравенства $i = \nu + 1$, во втором $i = \nu$, в третьем $i = \nu - 1$, …, в последнем $i = 1$ и получим

$$H(A_{\nu+1} / A^\nu) \le H(A_\nu / A^{\nu-1}) \le H(A_{\nu-1} / A^{\nu-2}) \le ... \le H(A_2 / A^1) \le H(A_1),$$

что и доказывает справедливость теоремы.

ТЕОРЕМА 8.3. Для условных энтропий событий стационарного источника существует предел $\lim\limits_{\nu\to\infty} H(A_{\nu+1}/A^{\nu}) = H(A/A^{\infty})$, где $H(A/A^{\infty})$ – некоторое число, характеризующее источник и представляющее собой среднее количество информации на событие, достаточно удаленное от момента начала наблюдения.

ДОКАЗАТЕЛЬСТВО. Как всякая энтропия, $H(A_{\nu+1}/A^{\nu})$ удовлетворяет ограничениям $0 \le H(A_{\nu+1}/A^{\nu}) \le \log_2 L$, т.е. энтропия – неотрицательная величина, ограниченная снизу величиной 0. Последовательность $H(A_{\nu+1}/A^{\nu})$ не возрастает по аргументу ν и ограничена снизу, откуда сразу же следует существование предела $\lim\limits_{\nu\to\infty} H(A_{\nu+1}/A^{\nu}) = H(A/A^{\infty})$. Энтропия $H(A/A^{\infty})$ есть предельная энтропия, необходимая для определения события на выходе стационарного источника.

ТЕОРЕМА 8.4. При любом заданном сколько угодно малом положительном числе $\varepsilon > 0$ последовательность событий, порождаемых дискретным стационарным источником, можно закодировать в последовательность символов, выбираемых из заданного алфавита, таким образом, чтобы среднее по ансамблю число символов на событие удовлетворяло неравенству $\bar{n} < \dfrac{H(A/A^{\infty})}{\log_2 D} + \varepsilon$.

ДОКАЗАТЕЛЬСТВО. Разобьем последовательность событий на группы по ν. Тогда каждая последовательность будет являться элементом множества A^{ν}, состоящего из ν букв, принадлежащих A (рис. 8.6). Энтропия этого множества в соответствии с известным свойством равна

$$H(A^{\nu}) = H(A_1) + H(A_2/A_1) + \ldots + H(A_{\nu}/A^{\nu-1}) = \sum_{j=1}^{\nu} H(A_j/A^{j-1}).$$

Для множества образованных последовательностей букв можно построить оптимальный код, причем среднее число символов алфавита кодировки на одну последовательность $\overline{n_{\nu}}$ должно удовлетворять, по ранее доказанной теореме 7.3, соотношению

$\frac{H(A^\nu)}{\log_2 D} \le \overline{n}_\nu < \frac{H(A^\nu)}{\log_2 D} + 1$. Делим неравенство на ν и получаем новое соотношение $\frac{\frac{1}{\nu}H(A^\nu)}{\log_2 D} \le \bar{n} = \frac{\overline{n}_\nu}{\nu} < \frac{\frac{1}{\nu}H(A^\nu)}{\log_2 D} + \frac{1}{\nu}$. Далее более детально рассмотрим величину $\frac{1}{\nu}H(A^\nu)$, представив ее в виде суммы $\frac{1}{\nu}[H(A_1) + ... + H(A_i / A^{i-1}) + H(A / A^\infty)(\nu - i)]$. Первые i слагаемых суммы отражают переходной процесс условных энтропий событий, а последнее слагаемое – стационарное состояние. После перегруппировки слагаемых можно получить выражение $\frac{H(A_1) - H(A/A^\infty)}{\nu} + ... + \frac{H(A_i / A^{i-1}) - H(A/A^\infty)}{\nu} + H(A/A^\infty)$, из которого видно, что с увеличением ν дробные слагаемые становятся бесконечно малыми величинами. Следовательно, $\lim_{\nu\to\infty} \frac{1}{\nu} H(A^\nu) = H(A/A^\infty)$. Таким образом, всегда можно найти такое целое ν, что $\varepsilon \ge \frac{1}{\nu}$ и $\bar{n} < \frac{H(A/A^\infty)}{\log_2 D} + \varepsilon$. Полученный результат завершает доказательство теоремы.

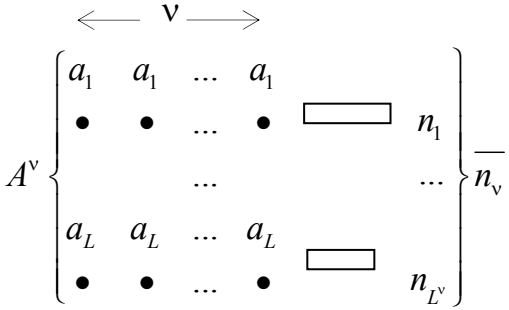


Рис. 8.6. Кодирование группы событий длиной ν

Отметим, что эффективность процедур кодирования для случайных источников оценивается по формуле

$$Э_к = \frac{H(A/A^\infty)}{\overline{n}\log_2 D},$$

происхождение которой достаточно очевидно.

8.3. СРЕДНЕЕ ПО АНСАМБЛЮ И СРЕДНЕЕ ПО ПОСЛЕДОВАТЕЛЬНОСТИ

В доказательствах, касающихся стационарных источников, все средние величины, в частности $\overline{n_v}$, $H(A^v)$, вычислялись как среднее по множеству.

На практике обычно интересуются средней длиной последовательности кодовых слов во времени, а не математическим ожиданием длины кодового слова в некоторый момент.

Кроме того, особое значение имеют процессы, в которых усреднение по множеству и усреднение по времени совпадают. Такое совпадение имеет место для эргодических стационарных источников, которые являются подклассом стационарных источников информации.

Для более детального изучения усреднения по множеству и усреднения по последовательности рассмотрим пример марковского источника.

Марковский источник характеризуется состояниями, в которых он может находиться, и правилами, управляющими его переходом из одного состояния в какое-либо другое. Буквы порождаются источником с вероятностями, зависящими только от состояния, в котором находится источник. После порождения буквы новое состояние однозначно определяется старым состоянием и этой буквой.

Действие источника удобно иллюстрировать граф-схемами. На рис. 8.7 приведена граф-схема разложимого марковского источника (режимы E_1 и E_2) [6]. Событие α_i, порождаемое источником, может быть одной из четырех букв множества $A = \{a_1, a_2, a_3, a_4\}$. Каждый узел представляет собой состояние

S_j, $j=\overline{0,4}$, а каждая ветвь указывает изменение состояния, происшедшее в результате порождения буквы a_k. Состояние S_0 является начальным, имеющим место при порождении первой буквы.

Матрица условных вероятностей полностью описывает источник, а потому и ансамбль сообщений, образованный порожденными им последовательностями букв:

$$\left\| p(a_k / S_j) \right\| = \begin{array}{c|cccc} & a_1 & a_2 & a_3 & a_4 \\ S_0 & 0{,}3 & 0{,}2 & 0{,}3 & 0{,}2 \\ S_1 & 1/2 & 1/2 & 0 & 0 \\ S_2 & 3/4 & 1/4 & 0 & 0 \\ S_3 & 0 & 0 & 2/3 & 1/3 \\ S_4 & 0 & 0 & 1/2 & 1/2 \end{array}.$$

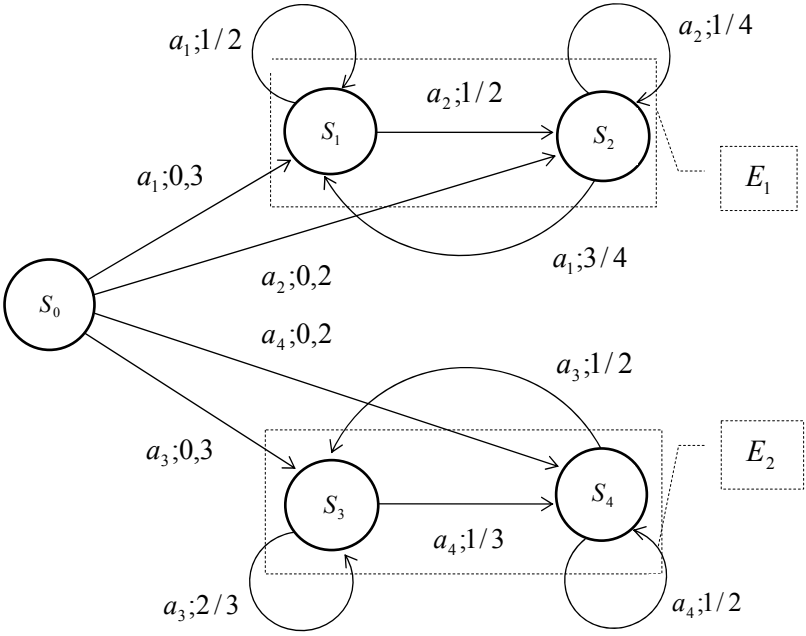


Рис. 8.7. Граф-схема марковского источника

Как видно из граф-схемы, распределение вероятностей $p(\alpha_i)$ для всех натуральных i равно:

$$p\{\alpha_i = a_1\} = 0{,}3; \quad p\{\alpha_i = a_2\} = 0{,}2;$$
$$p\{\alpha_i = a_3\} = 0{,}3; \quad p\{\alpha_i = a_4\} = 0{,}2.$$

Условное распределение вероятностей $p(\alpha_i / \alpha_{i-1})$ можно получить из представленной граф-схемы

$$\left\| p(\alpha_i / \alpha_{i-1}) \right\| = \begin{array}{c|cccc} & a_1 & a_2 & a_3 & a_4 \\ a_1 & 1/2 & 1/2 & 0 & 0 \\ a_2 & 3/4 & 1/4 & 0 & 0 \\ a_3 & 0 & 0 & 2/3 & 1/3 \\ a_4 & 0 & 0 & 1/2 & 1/2 \end{array}.$$

Условные распределения вероятностей более высоких порядков задаются соотношением

$$p(\alpha_i / \alpha_{i-1}, \alpha_{i-2}, ..., \alpha_{i-j}) = p(\alpha_i / \alpha_{i-1}) \quad , \quad j > 1.$$

Другими словами, условные распределения вероятностей зависят только от непосредственно предшествующего события, если это событие задано.

Источник стационарен, так как все распределения вероятностей не зависят от i. Легко убедиться в справедливости соотношения

$$\sum_{\alpha_{i-1} \in A_{i-1}} p(\alpha_{i-1}) p(\alpha_i / \alpha_{i-1}) = p(\alpha_i).$$

Предположим, что последовательность на выходе источника разбивается на группы событий по два. Возможные комбинации символов на выходе марковского источника и необходимые количественные оценки приведены в табл. 8.1, а распределение вероятностей случайной величины $J(\alpha^2)$ на рис. 8.8.

Энтропия ансамбля $H(A^2)$ как среднее по множеству имеет вид

$$H(A^2) = m[J(\alpha^2)] = 2{,}88.$$

Поскольку вероятность реализации режимов E_1 и E_2 одинакова, то по истечении некоторого времени источник будет генерировать либо последовательность $a_1 a_2 a_2 a_1 ...$ с энтропией ансамбля

$H(A^2/E_1)$ как среднее по временной последовательности

$$H(A^2/E_1) = m[J(\alpha^2)/E_1] = 2{,}86,$$

либо последовательность $a_3a_3a_4a_4...$ с энтропией ансамбля $H(A^2/E_2)$ как среднее по временной последовательности

$$H(A^2/E_2) = m[J(\alpha^2)/E_2] = 2{,}90.$$

Таблица 8.1

Комбинации $\alpha_{i-1}\alpha_i$	Режимы работы	Вероятности $p(\alpha_{i-1}\alpha_i) = p(\alpha_{i-1})p(\alpha_i/\alpha_{i-1})$	Собственная информация $J(\alpha^2) = -\log_2 p(\alpha_{i-1}\alpha_i)$	Распределение вероятностей $P\{J(\alpha^2) = J\}$
a_1a_1	E_1	0,15	2,70	0,45
a_1a_2		0,15		
a_2a_1		0,15		
a_2a_2		0,05	4,30	0,05
a_3a_3	E_2	0,20	2,30	0,20
a_3a_4		0,10	3,30	0,30
a_4a_3		0,10		
a_4a_4		0,10		

Последние два выражения получены с учетом того, что распределения вероятностей принимают иной вид (табл. 8.2).

Таблица 8.2

$p\{J(\alpha^2) = J/E_1\}$	$J = 2{,}7$	0,9
	$J = 4{,}3$	0,1
$p\{J(\alpha^2) = J/E_2\}$	$J = 2{,}3$	0,4
	$J = 3{,}3$	0,6

$p\{J(\alpha^2) = 2{,}7\} = 0{,}45;$

$p\{J(\alpha^2) = 4{,}3\} = 0{,}05;$

$p\{J(\alpha^2) = 2{,}3\} = 0{,}20;$

$p\{J(\alpha^2) = 3{,}3\} = 0{,}30;$

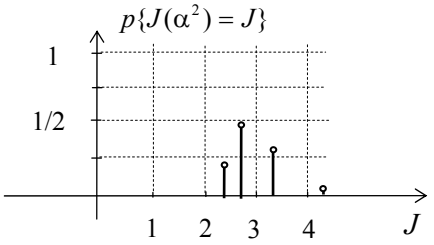


Рис. 8.8. Распределение вероятностей $J(\alpha^2)$

В этом случае среднее по ансамблю 2,88 не совпадает со средними по временной последовательности 2,86 или 2,90.

Стационарный источник, обладающий более чем одним режимом работы, называется разложимым или приводимым. Неприводимый стационарный источник всегда эргодический. Можно показать, что любой стационарный источник можно представить как комбинацию эргодических источников.

8.4. КОДИРОВАНИЕ СОБЫТИЙ, ПОРОЖДАЕМЫХ ИСТОЧНИКОМ С ФИКСИРОВАННОЙ СКОРОСТЬЮ

До сих пор рассматривались методы кодирования для источников с управляемой скоростью, т.е. источников, в которых моменты порождения событий могут изменяться кодером (рис. 8.9).

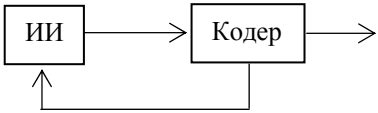


Рис. 8.9. Источник с управляемой скоростью

Очевидно, что для ***источника с фиксированной скоростью*** кодер не может передавать и кодировать слова с той же скоростью, что и источник, так как разные сообщения требуют разного числа символов. Буфер, как следует из теории массового обслуживания,

также не решает вопроса. Он будет переполняться с некоторой отличной от нуля вероятностью и вносить искажения в передаваемую последовательность за счет потерь.

В этом случае используется следующий метод кодирования. Множество последовательностей событий кодируется так, чтобы все коды имели равную длину и могли бы быть переданы за одно и то же время. Если каждое событие имеет L возможных букв, а ν – число событий в последовательности, то множество возможных комбинаций букв составит L^{ν}. Каждой комбинации букв ставится в соответствие кодовое слово длиной n_{ν}. При этом должно удовлетворяться соотношение $D^{n_{\nu}} \geq L^{\nu}$, где число букв алфавита кодировки обозначено, как и прежде, через D. Тогда $n_{\nu} = \nu \left\lceil \frac{\log_2 L}{\log_2 D} \right\rceil$, где $\rceil .. \lceil$ – наименьшее большее целое. В этом суть кодировки. Однако эффективная процедура кодирования невозможна.

Оказывается, что ***асимптотически эффективное кодирование*** можно осуществить, если выделить подмножество маловероятных комбинаций букв и приписать всем комбинациям этого подмножества одно и то же кодовое слово. Естественно, при этом вводится неоднозначность кодирования.

Рассмотрим простой случай статистически независимых событий на выходе источника [10]. Пусть $H(A)$ – энтропия ансамбля букв на выходе источника информации. Количество информации, которое несет каждая комбинация из ν букв, равно

$$J(\alpha^{\nu}) = -\log_2 p(\alpha^{\nu}) = J(\alpha_1) + ... + J(\alpha_j) + ... + J(\alpha_{\nu}),$$

где $J(\alpha_j) = -\log_2 p(\alpha_j),\ j = \overline{1,\nu}$.

ТЕОРЕМА 8.5. Пусть $\varepsilon > 0$ и $\delta > 0$ – сколь угодно малые числа, а T – множество последовательностей α^{ν}, для которых $J(\alpha^{\nu}) \leq \nu[H(A) + \varepsilon]$, причем $M = |T|$, а $p(T)$ – вероятность появления последовательностей, принадлежащих множеству T. Тогда существует такое целое число μ, что для любого $\nu > \mu$ справедливы неравенства: $p(T) > 1 - \delta$; $M < 2^{\nu[H(A)+\varepsilon]}$.

ДОКАЗАТЕЛЬСТВО. $J(\alpha^{\nu})$ есть сумма одинаково распределенных и статистически независимых случайных величин со средним $H(A)$. Согласно закону больших чисел, который гласит, что для любых $\varepsilon > 0$ и $\delta > 0$ существует такое целое μ, что для любого $\nu > \mu$ $p\left\{\left|\frac{J(\alpha^{\nu})}{\nu} - H(A)\right| > \varepsilon\right\} < \delta$. Другими словами, вероятность того, что среднее арифметическое будет отличаться от математического ожидания больше, чем на любое конечное число, может быть сделана сколь угодно малой, если ν достаточно велико. Представим неравенство, соответствующее закону больших чисел, в более удобном виде

$$1 - p\left\{\left|\frac{J(\alpha^{\nu})}{\nu} - H(A)\right| > \varepsilon\right\} > 1 - \delta,$$

$$p\left\{\left|\frac{J(\alpha^{\nu})}{\nu} - H(A)\right| \le \varepsilon\right\} > 1 - \delta.$$

Вероятности $p(T)$ соответствует неравенство $\frac{J(\alpha^{\nu})}{\nu} \le H(A) + \varepsilon$, которое имеет большую область определения, чем диапазон, указанный в неравенствах $H(A) - \varepsilon \le \frac{J(\alpha^{\nu})}{\nu} \le H(A) + \varepsilon$. Отсюда следует, что $p(T) > p\left\{\left|\frac{J(\alpha^{\nu})}{\nu} - H(A)\right| \le \varepsilon\right\} > 1 - \delta$, т.е. $p(T) > 1 - \delta$.

Верхняя граница числа последовательностей в T следует из условия теоремы $J(\alpha^{\nu}) = -\log_2 p(\alpha^{\nu}) \le \nu[H(A) + \varepsilon]$. Далее $p(\alpha^{\nu}) \ge 2^{-\nu[H(A)+\varepsilon]}$, которое можно представить в следующем виде:

$$M 2^{-\nu[H(A)+\varepsilon]} \le \sum_{\alpha^{\nu} \in T} p(\alpha^{\nu}) < \sum_{\alpha^{\nu} \in A^{\nu}} p(\alpha^{\nu}) = 1.$$

Следовательно, $M < 2^{\nu[H(A)+\varepsilon]}$.

Доказанная теорема дает следующий способ кодирования. M последовательностей из ν букв кодируются различными кодовыми словами одинаковой длины n_ν, а все остальные последовательности – одним и тем же кодовым словом также длины n_ν. Возникающая при этом вероятность неоднозначного кодирования равна вероятности появления последовательностей из множества $\alpha^\nu \notin T$.

Наконец, асимптотически эффективный алгоритм кодирования для источников с фиксированной скоростью состоит из шагов:

– находим наибольшее M, удовлетворяющее неравенствам

$$\begin{cases} M < 2^{\nu[H(A)+\varepsilon]}; \\ J(\alpha^\nu) \le \nu[H(A)+\varepsilon]; \end{cases}$$

– из условия $D^{n_\nu} \ge M+1$ определяем длину кодовых слов

$$n_\nu \ge \left\lceil \frac{\log_2(M+1)}{\log_2 D} \right\rceil;$$

– формируем $M+1$ различных кодовых слов длины n_ν с помощью символов алфавита кодировки.

ЗАДАЧИ

8.1. От источника информации поступают статистически независимые символы с вероятностями 1/3 и 2/3 $\left(H(U)=0{,}918\right)$. Построить оптимальные двоичные слова для сообщений, состоящих из трех последовательных двоичных символов. Оценить эффективность кодирования.

8.2. Источник информации порождает статистически независимые символы с вероятностями 1/4 и 3/4 соответственно. Построить оптимальные троичные слова для сообщений, построенных из трех последовательных двоичных символов источника. Построить кодовое дерево. Оценить эффективность кодирования.

8.3. Стационарный источник с фиксированной скоростью порождает последовательность статистически независимых сообщений. Ансамбль сообщений имеет вид

Сообщение u_k	Вероятность сообщения $p(u_k)$
u_1	1/2
u_2	1/3
u_3	1/6

Провести асимптотически эффективное кодирование при $D = 2$, $\nu = 2$ и $\varepsilon = 0{,}01$.

8.4. Стохастический источник порождает последовательность событий $\alpha_1, \alpha_2, \alpha_3, \ldots$. Каждое событие α_i – одна из букв алфавита $A = \{a_1, a_2, a_3, a_4\}$. Источник описывается распределением вероятностей $p(\alpha_i) = 1/4$ и

$$\left\| p(\alpha_i / \alpha_{i-1}) \right\| = \begin{vmatrix} 1/2 & 1/4 & 1/8 & 1/8 \\ 1/4 & 1/8 & 1/8 & 1/2 \\ 1/8 & 1/8 & 1/2 & 1/4 \\ 1/8 & 1/2 & 1/4 & 1/8 \end{vmatrix}.$$

Условные распределения вероятностей более высоких порядков задаются соотношением

$$p(\alpha_i / \alpha_{i-1}, \alpha_{i-2}, \ldots, \alpha_{i-j}) = p(\alpha_i / \alpha_{i-1}), \quad j > 1.$$

Вычислить $H(A), H(A/A^1), \ldots, H(A/A^\infty)$.

8.5. Стохастический источник имеет параметры аналогичные тем, которые приведены в задаче 8.4. Разбить последовательность событий на выходе источника на сообщения, образованные двумя последовательными событиями. Для полученного ансамбля построить оптимальный двоичный код и оценить его эффективность.

8.6. Источник информации порождает статистически независимые буквы с энтропией $H(U) = 2$. Формируются двоичные коды для всех возможных последовательностей из 10 букв. Найти верхнюю и нижнюю границу для требуемого среднего числа двоичных символов на букву.

8.7. Последовательность, порожденная двоичным дискретным источником, имеет вид

1	2	3	4	5	6	7	8
0	1	1	0	1	1	0	1

.

Используя один из эффективных методов универсального кодирования (УК – метод Бабкина), провести ее кодирование и декодирование с использованием того же алфавита кодировки.

8.8. Стационарный марковский источник двоичных сообщений с матрицей условных вероятностей

$$\left\| p(\alpha_i / \alpha_{i-1}) \right\| = \begin{array}{c|cc} & 0 & 1 \\ 0 & 1/2 & 1/2 \\ 1 & 3/4 & 1/4 \end{array}$$

достаточно долго генерирует сообщения. Сколько информации содержится в сообщении о том, что на выходе источника появилась "1"?

9. ДИСКРЕТНЫЕ КАНАЛЫ СВЯЗИ

Под ***каналом связи*** или передачи информации принято понимать совокупность средств, предназначенных для передачи сигналов. Обобщенная схема канала передачи информации представлена на рис. 9.1.

Канал связи преобразует последовательность входных событий $\xi_1, \xi_2, ..., \xi_j, ...$, каждое из которых представляется точкой $\xi_j = x_k \in X$ входного пространства X, в выходные события $\eta_1, \eta_2, ..., \eta_j, ...$, каждое из которых представляется точкой $\eta_j(\xi_j) = y_i \in Y$ выходного пространства Y. Преобразование управляется условным распределением вероятностей $p(y_i / x_k)$, которое описывает случайные помехи, действующие в канале.

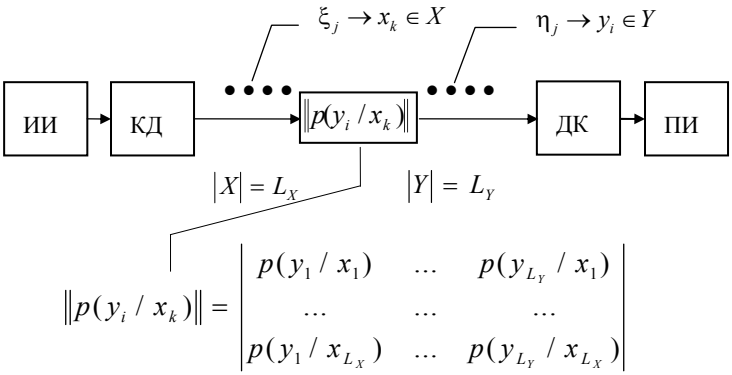


Рис. 9.1. Структура дискретного канала связи: ИИ – источник информации; КД – кодер; ДК – декодер; ПИ – потребитель информации

9.1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Обычно каналы классифицируются в соответствии с характеристиками входного и выходного пространств и распределением ве-

роятностей, управляющих преобразованием одного пространства в другое.

Канал называется ***дискретным***, если входное и выходное пространства дискретны. Он называется ***непрерывным***, если входное и выходное пространства непрерывны. Если одно из пространств дискретно, а другое непрерывно, то канал называется соответственно ***дискретно-непрерывным*** или ***непрерывно-дискретным***.

Канал, в котором условное распределение вероятностей $p(\eta_j / \xi_j) = p(y_i / x_k)$ одно и то же для всех последовательных входных и выходных событий j, называют ***стационарным каналом без памяти***. Если же, наоборот, это условное распределение вероятностей является функцией предыдущих событий на входе или выходе, то говорят, что канал обладает памятью. Память называют конечной, если эта зависимость распространяется только на конечное число предшествующих событий.

Дальнейшее разделение каналов на классы можно произвести в соответствии с особыми свойствами симметрии, которыми могут обладать условные распределения вероятностей. Мы вернемся к этому вопросу при вычислении пропускной способности каналов. В этих вопросах свойства симметрии оказываются существенными.

Одним из основных вопросов в теории передачи информации является определение зависимости скорости передачи информации и пропускной способности от параметров канала и характеристик сигналов и помех. Эти вопросы были впервые глубоко исследованы К. Шенноном.

Прежде чем переходить к изучению дискретных каналов, попробуем получить более ясное понимание природы взаимной энтропии. Проинтерпретируем взаимную энтропию при передаче информации по каналу с шумом. Пусть передаются сообщения двоичным равноплотным кодом со ***скоростью передачи символов*** $\upsilon_x = \dfrac{1}{\tau_x} = 1000$ [симв/с], где τ_x – длительность передачи одного символа. За одну секунду на вход канала передается количество информации $\upsilon_x H(X) = 1000[-\dfrac{1}{2}\log_2\dfrac{1}{2} - \dfrac{1}{2}\log_2\dfrac{1}{2}] = 1000$ [бит/с].

Пусть в процессе передачи из-за искажений 1 % символов принимается неправильно. Ясно, что передача информации с полной точностью будет при этом вестись с меньшей скоростью. Кажется, что скорость передачи информации в этом случае можно вычислить просто как $1000-10=990$ [бит/с]. Однако это не совсем верно, так как на приемном конце не известно, где именно произошла ошибка. Действительно, рассмотрим крайний случай, когда 50 % символов искажаются при передаче. Пришлось бы считать в этом случае, что скорость передачи информации равна $1000-500=500$ [бит/с]. На самом же деле передачи информации нет, канал вообще не нужен, можно просто на приемном конце разыгрывать прием символов "0" или "1", подбрасывая монету.

Мерой переданной по каналу информации, естественно считать взаимную энтропию между передаваемыми и принимаемыми символами $H(X;Y)=H(Y)-H(Y/X)$. Выполним необходимые вычисления для рассматриваемого случая:

$$H(Y)=-\sum_{Y} p(y_i)\log_2 p(y_i) = 1;$$

$$H(Y/X)=-\sum_{X}\sum_{Y} p(x_k)p(y_i/x_k)\log_2 p(y_i/x_k) = 0{,}081;$$

$$p(y_i)=\sum_{X} p(x_k)p(y_i/x_k)=0{,}5;\ i=\overline{1,2}\ .$$

Количественные оценки получены с учетом того, что исходные вероятности равны

$$p(x_k)=p(0)=p(1)=0{,}5;$$

$$p(y_i/x_k)=\begin{cases} p(0/0)=p(1/1)=0{,}99; \\ p(0/1)=p(1/0)=0{,}01. \end{cases}$$

Тогда скорость передачи информации по каналу можно вычислить как $\upsilon_x H(X;Y)=1000(1-0{,}081)=919$ [бит/с]. Следовательно, на приемном конце получено 919 [бит/с], а не 990 [бит/с]. В дальнейших рассуждениях множитель υ_x мы будем просто опускать.

Таким образом, средняя ***скорость передачи информации*** по каналу с шумом может быть определена как $H(X;Y)$. Особенность данной оценки состоит в том, что имеется в виду передача информации с полной точностью.

9.2. ДИСКРЕТНЫЕ СТАЦИОНАРНЫЕ КАНАЛЫ БЕЗ ПАМЯТИ

Обозначим через $\xi^n = (\xi_1 \xi_2 ... \xi_n)$ произвольную входную последовательность событий, а через $\eta^n = (\eta_1 \eta_2 ... \eta_n)$, соответствующую выходную последовательность событий.

Для стационарных каналов без памяти условное распределение $p(\eta_j / \xi_j)$ остается одним и тем же для каждого целого j независимо от всех предшествующих событий и равно $p(y_i / x_k)$ при $\xi_j = x_k$, $\eta_j = y_i$. Следовательно, $p(\eta^n / \xi^n) = \prod_{j=1}^{n} p(\eta_j / \xi_j)$.

Если $p(\xi^n)$ – совместное распределение событий на входе, то совместное распределение выходной последовательности событий легко вычисляется как

$$p(\eta^n) = \sum_{X^n} p(\xi^n \eta^n) = \sum_{X^n} p(\xi^n) p(\eta^n / \xi^n) = \sum_{X^n} p(\xi^n) \prod_{j=1}^{n} p(\eta_j / \xi_j).$$

Взаимная энтропия между выходными и входными последовательностями событий по определению имеет вид

$$H(X^n; Y^n) = H(Y^n) - H(Y^n / X^n).$$

Ранее мы уже получали

$$H(Y^n) = \sum_{j=1}^{n} H(Y_j / Y^{j-1}) = f_1[n, p(\xi^n)],$$

где $H(Y_j / Y^{j-1}) = -\sum_{Y^j} p(\eta^j) \log_2 p(\eta_j / \eta^{j-1})$.

Вычислим условную энтропию

$$H(Y^n / X^n) = -\sum_{X^n Y^n} p(\xi^n \eta^n) \log_2 p(\eta^n / \xi^n) =$$

$$= -\sum_{X^n Y^n} p(\xi^n \eta^n) \sum_{j=1}^{n} \log_2 p(\eta_j / \xi_j) =$$

$$= -\sum_{j=1}^{n} \sum_{X_j Y_j} \dots \sum_{X_n Y_n} p(\xi_j \eta_j \dots \xi_n \eta_n) \log_2 p(\eta_j / \xi_j) =$$

$$= -\sum_{j=1}^{n} \sum_{X_j Y_j} p(\xi_j \eta_j) \log_2 p(\eta_j / \xi_j) = \sum_{j=1}^{n} H(Y_j / X_j) = f_2[n, p(\xi^n)].$$

Наконец, взаимная энтропия может быть представлена в виде

$$H(X^n; Y^n) = \sum_{j=1}^{n} [H(Y_j / Y^{j-1}) - H(Y_j / X_j)].$$

Средняя взаимная информация $H(X^n; Y^n)$ является мерой количества информации, содержащейся в среднем по множеству последовательностей в n выходных событиях относительно соответствующих n событий на входе. Из предыдущих выражений видно, что $H(X^n; Y^n)$ зависит от длины последовательностей n и распределения событий на входе $p(\xi^n)$ для заданного канала с матрицей $\left\| p(y_i / x_k) \right\|$. Распределение на входе канала можно варьировать путем соответствующего выбора метода кодирования.

Тогда естественно определить ***информационную пропускную способность канала*** C как максимально возможное значение

$$C = \max_{\{n, p(\xi^n)\}} \left\{ \frac{H(X^n; Y^n)}{n} \right\}, \qquad (9.1)$$

взятое по всем возможным n и распределениям $p(\xi^n)$.

Скорость передачи информации с полной точностью в общем случае зависит от статистических свойств сообщения и метода кодирования. Информационная пропускная способность – это характеристика канала. Она не зависит от фактической скорости передачи информации.

ТЕОРЕМА 9.1. Для любого дискретного стационарного канала справедливо неравенство $H(X^n; Y^n) \le \sum_{j=1}^{n} H(X_j; Y_j)$. Знак равенства имеет место, когда события на выходе канала статистически независимы.

ДОКАЗАТЕЛЬСТВО. Известно, что условная энтропия удовлетворяет неравенству

$$H(Y_j / Y^{j-1}) \le H(Y_j) = -\sum_{Y_j} p(\eta_j) \log_2 p(\eta_j).$$

Следовательно,

$$H(X^n; Y^n) = \sum_{j=1}^{n} [H(Y_j / Y^{j-1}) - H(Y_j / X_j)] \le \sum_{j=1}^{n} [H(Y_j) - H(Y_j / X_j)],$$

что соответствует условию теоремы $H(X^n; Y^n) \le \sum_{j=1}^{n} H(X_j; Y_j)$.

Знак равенства имеет место, когда $H(Y_j) = H(Y_j / Y^{j-1})$, т.е. события на выходе канала статистически независимы.

ТЕОРЕМА 9.2. Пусть $p(y_i / x_k)$ – условное распределение для дискретного стационарного канала, а $p(x_k)$ – произвольное распределение на входном пространстве X. Тогда пропускная способность канала есть максимальное среднее значение взаимной информации $C = \max_{\{p(x_k)\}} \{H(X;Y)\}$, вычисленное по всем возможным распределениям $p(x_k)$.

ДОКАЗАТЕЛЬСТВО. Максимум информационной пропускной способности канала (9.1) достижим, если $H(X^n; Y^n) = \sum_{j=1}^{n} H(X_j; Y_j)$. В свою очередь, взаимные энтропии $H(X_j; Y_j)$ отличаются индексом j и зависят от $p(\eta_j / \xi_j)$, которые для стационарного канала одни и те же для всех целых j. Следовательно, $H(X_j; Y_j) = H(X;Y)$, $j = \overline{1,n}$. Это позволяет вычислять информационную пропускную способность канала C по более простой формуле, так как $H(X^n; Y^n) = nH(X;Y)$ и

$$C = \max_{\{n, p(\xi^n)\}} \left\{ \frac{H(X^n; Y^n)}{n} \right\} = \max_{\{p(x_k)\}} \{H(X;Y)\}.$$

9.3. СИММЕТРИЧНЫЕ СТАЦИОНАРНЫЕ КАНАЛЫ

Вычисления информационной пропускной способности дискретного стационарного канала C еще более просты, если матрица канала обладает специфическими свойствами симметрии.

Говорят, что ***канал симметричен по входу***, когда все строки матрицы условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \left| \begin{array}{cccc} & & \dots & \\ p(y_1 / x_k) = p_1 & p(y_2 / x_k) = p_2 & \dots & p(y_{L_Y} / x_k) = p_{L_Y} \\ & & \dots & \end{array} \right|$$

являются перестановками одного и того же множества чисел $p_1, p_2, ..., p_{L_Y}$, для которых $\sum_{i=1}^{L_Y} p_i = 1$. В этом случае частные условные энтропии имеют одинаковое значение для всех $x_k \in X$:

$$H(Y / x_k) = -\sum_{Y} p(y_i / x_k) \log_2 p(y_i / x_k) = -\sum_{i=1}^{L_Y} p_i \log_2 p_i .$$

Таким образом, $H(Y / x_k) = H(Y / x_m)$ для всех $x_k, x_m \in X$, т.е. для канала симметричного по входу передача искажает все символы в равной степени.

Канал ***симметричен по выходу***, когда все столбцы матрицы условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \left| \begin{array}{ccc} & p(y_1 / x_1) = p_1 & \\ & p(y_1 / x_2) = p_2 & \\ \dots & \dots & \dots \\ & p(y_1 / x_{L_X}) = p_{L_X} & \end{array} \right|$$

являются перестановками одного и того же множества чисел $p_1, p_2, ..., p_{L_X}$, для которых $\sum_{k=1}^{L_X} p_k = \text{Const}$. Симметричный по выходу канал обладает свойством инвариантности относительно равномерного распределения. Равномерное распределение на входе $p(x_k) = \dfrac{1}{L_X}$ приводит к равномерному распределению на выходе

$$p(y_i)=\sum_X p(x_k)p(y_i/x_k)=\sum_X \frac{1}{L_X}p(y_i/x_k)=\frac{1}{L_X}\sum_{k=1}^{L_X}p_k=\frac{\text{Const}}{L_X}=\frac{1}{L_Y}.$$

Симметричный канал – это канал симметричный по входу и выходу.

ТЕОРЕМА 9.3. Пропускная способность дискретного стационарного канала симметричного по входу удовлетворяет неравенству $C\le \log_2 L_Y+\sum_{i=1}^{L_Y} p_i \log_2 p_i$, где p_i – элементы матрицы $\left\|p(y_i/x_k)\right\|, i=\overline{1,L_Y}$. Знак равенства имеет место для симметричного канала, т.е. максимальная пропускная способность канала реализуется, если распределение событий на выходе равномерное.

ДОКАЗАТЕЛЬСТВО. Представим взаимную энтропию с учетом свойства симметрии:

$$H(X;Y)=H(Y)-H(Y/X);$$

$$H(Y/X)=\sum_X H(Y/x_k)p(x_k)=-\sum_Y p_i\log_2 p_i\sum_X p(x_k)=-\sum_{i=1}^{L_Y}p_i\log_2 p_i;$$

$$H(X;Y)=H(Y)+\sum_{i=1}^{L_Y}p_i\log_2 p_i.$$

Известно, что энтропия ансамбля не превосходит логарифма его мощности $H(Y)\le \log_2 L_Y$. Тогда для взаимной энтропии имеем неравенство $H(X;Y)\le \log_2 L_Y+\sum_{i=1}^{L_Y} p_i \log_2 p_i$. Поскольку $C=\max\limits_{\{p(x_k)\}}\{H(X;Y)\}$, то пропускная способность дискретного стационарного канала симметричного по входу также удовлетворяет неравенству $C\le \log_2 L_Y+\sum_{i=1}^{L_Y} p_i \log_2 p_i$. Энтропия $H(Y)=\log_2 L_Y$ тогда и только тогда, когда все события на выходе равновероятны, т.е. $p(y_i)=\frac{1}{L_Y}$. Распределение $p(x_k)$, при котором $p(y_i)=\frac{1}{L_Y}$, может в общем случае и не существовать. Если же такое $p(x_k)$

существует, то канал обязательно симметричный и $C = \log_2 L_Y + \sum_{i=1}^{L_Y} p_i \log_2 p_i$.

В качестве примера рассмотрим дискретный ***двоичный симметричный канал*** (ДСК), который задается матрицей условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 1-p & p \\ p & 1-p \end{vmatrix}.$$

Матрице соответствует граф переходных вероятностей, приведенный на рис. 9.2. Пропускная способность ДСК вычисляется по формуле $C = \log_2 |Y| + \sum_{i=1}^{L_Y} p_i \log_2 p_i = 1 + p \log_2 p + (1-p)\log_2 (1-p)$.

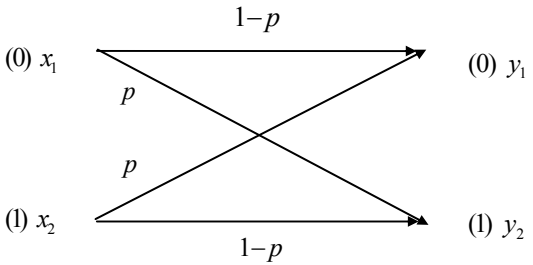


Рис. 9.2. Граф-схема ДСК

Графическая иллюстрация пропускной способности C приведена на рис. 9.3.

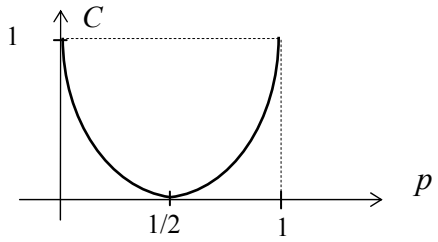


Рис. 9.3. График C для ДСК

9.4. ВЫЧИСЛЕНИЕ ИНФОРМАЦИОННОЙ ПРОПУСКНОЙ СПОСОБНОСТИ СТАЦИОНАРНОГО КАНАЛА

Матрица условных вероятностей дискретного канала $\|p(y_i / x_k)\|$ по существу представляет собой матрицу $\|p_{ik}\|$ неизменных параметров реального канала. Поэтому взаимную энтропию можно представить в следующем виде:

$$H(X;Y) = f_1[p(x_k)] - f_2[p(x_k)],$$

где $f_1[p(x_k)] = H(Y) = -\sum_Y \left[\sum_X p(x_k)p_{ik}\right] \log_2 \left[\sum_X p(x_k)p_{ik}\right]$;

$f_2[p(x_k)] = H(Y/X) = -\sum_X \sum_Y p(x_k)\, p_{ik} \log_2 p_{ik}$.

В общем случае, при определении информационной пропускной способности канала, приходится решать следующую экстремальную задачу с ограничением:

$$\begin{cases} C = \max\limits_{\{p(x_k)\}} \{f_1[p(x_k)] - f_2[p(x_k)]\}; \\ \sum\limits_X p(x_k) = 1. \end{cases}$$

Самый простой способ нахождения условного экстремума состоит в применении метода неопределенных множителей Лагранжа, суть которого состоит в нахождении безусловного экстремума функционала $\Phi[p(x_k), \lambda] = \{f_1[p(x_k)] - f_2[p(x_k)]\} + \lambda[1 - \sum_X p(x_k)]$:

$\frac{\partial \Phi}{\partial p(x_k)} = 0\Big|_{k=\overline{1,L_X}}$, $\frac{\partial \Phi}{\partial \lambda} = 0$, где λ – неопределенный множитель. Как правило, данный метод позволяет достичь желаемого результата.

Однако следует отметить, что при формулировке экстремальной задачи совершенно не учитывается условие $p(x_k) \geq 0$. Поэтому при нахождении экстремума можно получить вероятности отрицательные и больше единицы при выполнении в целом условия $\sum_X p(x_k) = 1$. В этом случае поиск экстремума необходимо проводить на границах области $H(X;Y)$. Встречаются такие случаи крайне редко, и на практике такой подход используется достаточно

широко. Тем более, что введение дополнительного ограничения необоснованно усложняет решение задачи в целом.

Найдем пропускную способность для дискретного ***двоичного симметричного канала со стиранием*** (ДСКС). ДСКС или канал с контролем задается матрицей условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 1-p-q & p & q \\ p & 1-p-q & q \end{vmatrix} .$$

Матрице соответствует граф переходных вероятностей, приведенный на рис. 9.4.

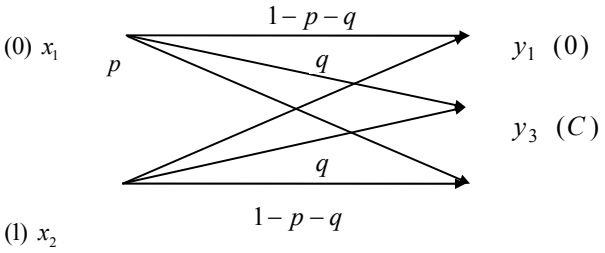


Рис. 9.4. Граф-схема ДСКС

Информационная пропускная способность ДСКС, который является каналом симметричным по входу, вычисляется по более простой схеме:

$$\begin{cases} C = \max\limits_{\{p(x_k)\}} H(Y) + \sum\limits_{i=1}^{L_Y} p_i \log_2 p_i ; \\ \sum\limits_{X} p(x_k) = 1. \end{cases}$$

Распределение $p(x_1) = z$, $p(x_2) = 1 - z$, максимизирующее $H(Y)$, максимизирует $H(X;Y)$ и определяет пропускную способность канала C. Определим безусловные вероятности

$$p(y_1) = z(1-p-q) + (1-z)p = z(1-2p-q) + p;$$

$$p(y_2) = zp + (1-z)(1-p-q) = (1-p-q) - z(1-2p-q);$$
$$p(y_3) = zq + (1-z)q = q.$$

Тогда

$$H(Y) = -\frac{1}{\ln 2}\{[z(1-2p-q)+p]\ln[z(1-2p-q)+p] +$$
$$+[(1-p-q)-z(1-2p-q)]\ln[(1-p-q)-z(1-2p-q)]+q\ln q\}.$$

Безусловный экстремум в данном случае будет иметь вид

$$\frac{\partial H(Y)}{\partial z} = -\frac{1}{\ln 2}\{[(1-2p-q)]\ln[z(1-2p-q)+p]+(1-2p-q)-$$
$$-(1-2p-q)]\ln[(1-p-q)-z(1-2p-q)]-(1-2p-q)\} = 0.$$

Откуда находим $z = 1/2$, так как $2z(1-2p-q) = 1-2p-q$. Наконец,

$$C = H(Y)\big|_{z=1/2} + \sum_{i=1}^{3} p_i \log_2 p_i =$$
$$= -\left\{2\frac{1-q}{2}\log_2\frac{1-q}{2} + q\log_2 q\right\} + \sum_{i=1}^{3} p_i \log_2 p_i =$$
$$= (1-q)[1-\log_2(1-q)] + (1-p-q)\log_2(1-p-q) + p\log_2 p.$$

Поверхность $C = F[p,q]$ приведена на рис. 9.5 [9].

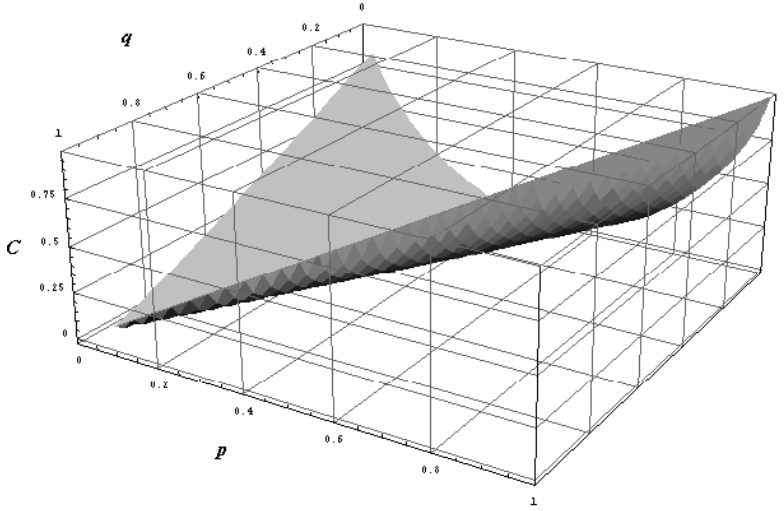


Рис. 9.5. Поверхность C для ДСКС

9.5. КОДИРОВАНИЕ ПРИ НАЛИЧИИ ШУМОВ

Подведем некоторые итоги. Была введена модель канала передачи информации. Под скоростью передачи информации понимаем среднее количество информации, приходящееся на один канальный символ. Обозначим скорость передачи информации через $R = H(X;Y) = H(X) - H(X/Y)$. Очевидно, что $R \le H(X)$, ибо нельзя принять больше информации, чем посылается.

Под пропускной способностью канала $C = \max\limits_{\{p(x_k)\}} \{H(X;Y)\}$ понимается максимальное среднее значение информации на символ, которое можно передать по данному каналу. Пропускную способность канала при передаче можно достичь надлежащим выбором источника информации и соответствующим выбором метода кодирования.

Ненадежность передачи по каналу определяется как $H(X/Y) = H(X) - R$ и соответствует неопределенности, которая имеется на приемном конце по отношению к переданному сообщению после оценки принятого сигнала. Ненадежность удовлетворяет соотношению $0 \le H(X/Y) \le H(X)$.

ТЕОРЕМА 9.4 (***основная теорема Шеннона***). Если дискретный источник создает сообщения со скоростью $H(X)$ и если $H(X) \le C$, то существует такая система кодирования, что сообщения источника могут быть переданы по каналу с произвольно малой ненадежностью. Если $H(X) > C$, то можно закодировать сообщения таким образом, что ненадежность передачи по каналу станет меньше, чем $H(X) - C + \varepsilon$, где $\varepsilon > 0$ сколь угодно мало, и не существует способа кодирования, обеспечивающего ненадежность передачи, меньше чем $H(X) - C$.

Примем теорему без доказательства, но дадим ее более подробную интерпретацию. Итак, теорема утверждает справедливость неравенств $H - C \le H(X/Y) < H(X) - C + \varepsilon$. Существование вполне определенной пропускной способности для канала с шумом на первый взгляд кажется нереальным. Однако ясно, что можно уменьшить вероятность ошибки, посылая информацию в избыточной форме. Так, например, многократно повторяя передачу сооб-

щений и статистически обрабатывая принятые сообщения, вероятность ошибки может быть сделана очень малой. Естественно ожидать, что для того чтобы приблизить вероятность ошибки к нулю, нужно неограниченно увеличивать избыточность кодирования и, следовательно, уменьшать скорость передачи к нулю.

На самом деле это не так. Введенная пропускная способность C имеет вполне определенное значение. Основная теорема для дискретного канала с шумом утверждает, что при надлежащем кодировании по каналу можно передавать информацию со скоростью C со сколь угодно малой частотой ошибок или со сколь угодно малой ненадежностью. Графическая иллюстрация основной теоремы приведена на рис. 9.6. Любая точка заштрихованной области может быть получена при надлежащем выборе способа кодирования, остальные точки, в принципе, получены быть не могут. Наибольший интерес, конечно, представляют точки, лежащие на выделенном горизонтальном отрезке и принадлежащие ε окрестности выделенной наклонной линии. Именно эти точки соответствуют основной теореме Шеннона.

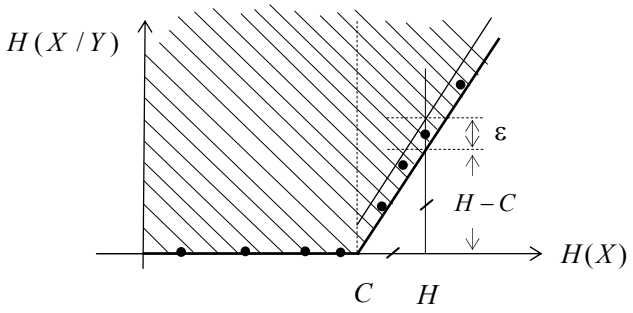


Рис. 9.6. Область допустимых значений ненадежности

Вернемся к рассмотрению уже разобранного примера, когда на вход ДСК поступает равноплотный двоичный код и 1 % символов искажается при передаче.

Информационная пропускная способность такого канала $\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 0{,}99 & 0{,}01 \\ 0{,}01 & 0{,}99 \end{vmatrix}$ была определена через $H(X;Y)$. Хотя

теперь вполне определенно можно записать, что его пропускная способность равна $C = \log_2 2 + \sum_{i=1}^{2} p_i \log_2 p_i = 0{,}919$, а ненадежность передачи – $H(X/Y) = H(X) - C = 0{,}081$.

Для того чтобы уменьшить $H(X/Y)$ до нуля (рис. 9.7), необходимо уменьшить энтропию источника $H(X)$, что неизбежно приведет к неравновероятному использованию двоичных символов и, следовательно, к избыточности. Именно эта избыточность в дальнейшем и должна использоваться для помехоустойчивого кодирования.

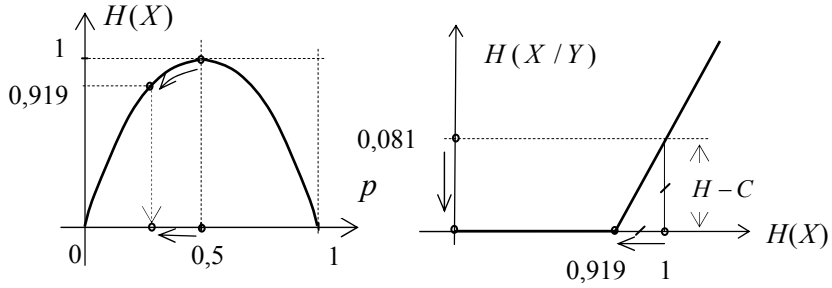


Рис. 9.7. Иллюстрация основной теоремы Шеннона

Основная теорема Шеннона не дает конкретного метода построения кода, обеспечивающего передачу со скоростью, равной пропускной способности канала. Однако это не уменьшает фундаментального значения теоремы, которая в области помехоустойчивого кодирования стимулирует поиск новых оптимальных корректирующих кодов и дает оценку их эффективности.

ЗАДАЧИ

9.1. Вывести формулы оценки информационной пропускной способности дискретного двоичного канала (ДК) (рис. 9.8) без памяти. Построить поверхность $C = F[p_1, p_2]$.

9.2. Определить информационную пропускную способность канала, состоящего из двух последовательно соединенных двоичных симметричных каналов (ДСК).

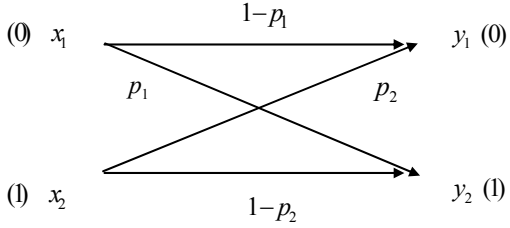


Рис. 9.8. Граф-схема ДК

9.3. Сообщения равноплотным двоичным кодом передаются на вход канала. В процессе передачи из-за искажений 1 % символов принимается неправильно. Вычислить информационную пропускную способность двоичного симметричного стационарного канала.

9.4. Два одинаковых двоичных симметричных канала соединены параллельно, т.е. один и тот же символ передается по обоим каналам. Какова информационная пропускная способность такой системы, если информация стирается, когда на выходе символы не совпадают?

9.5. Два одинаковых двоичных симметричных канала соединены параллельно, т.е. один и тот же символ передается по обоим каналам. Какова информационная пропускная способность такой системы, если в случае искажений ошибка по обоим каналам не обнаруживается?

9.6. Определить информационную пропускную способность канала, заданного матрицей условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \frac{1}{2} \begin{vmatrix} 1-p & 1-p & p & p \\ p & p & 1-p & 1-p \end{vmatrix} .$$

9.7. Определить информационную пропускную способность канала, заданного матрицей условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 1-p & p & 0 & 0 \\ p & 1-p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 0 & 0 & p & 1-p \end{vmatrix}.$$

9.8. Определить информационную пропускную способность канала, заданного матрицей условных вероятностей

$$\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 1-p & p & 0 \\ p & 1-p & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

10. ПОМЕХОУСТОЙЧИВОЕ КОДИРОВАНИЕ

Основная цель ***помехоустойчивого кодирования*** передаваемой информации может быть сформулирована следующим образом: если при передаче последовательность символов была искажена, то на выходе канала возникшие ошибки должны быть обнаружены и исправлены. Помехоустойчивый код отличается от обычного кода тем, что передаются в канал не все кодовые комбинации, которые можно сформировать из имеющегося количества разрядов, а лишь некоторые из них, обладающие определенным свойством и называемые разрешенными. Остальные неиспользуемые кодовые комбинации называют запрещенными. Помехоустойчивые или корректирующие коды классифицируют в зависимости от правил, по которым формируются разрешенные и запрещенные кодовые комбинации, а также от методов и способов введения избыточности (рис. 10.1) [4].

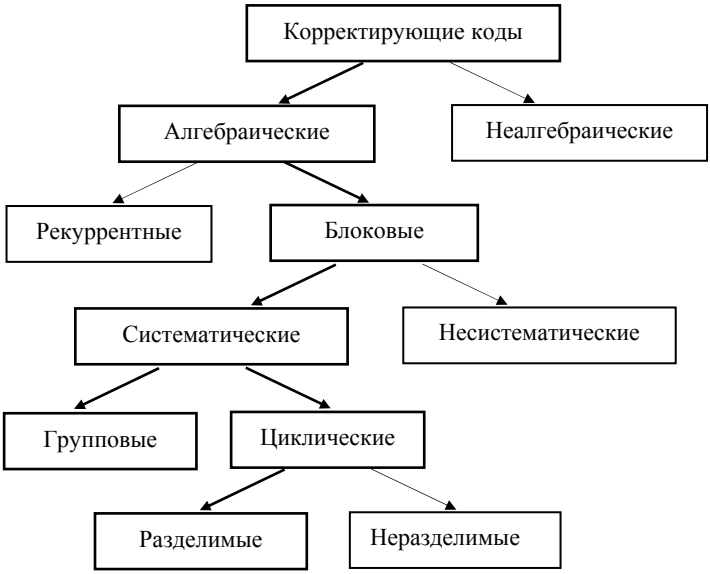


Рис. 10.1. Классификация корректирующих кодов

В ***рекуррентных кодах*** процесс кодирования и декодирования носит непрерывный характер. В ***блоковых кодах*** каждому сообщению сопоставляется кодовая комбинация из n символов. Блоки кодируются и декодируются отдельно друг от друга. В ***систематических кодах*** применяются линейные операции над символами. ***Групповые*** и ***циклические*** являются частным случаем систематических кодов и в своей основе имеют алгебраические структуры – группы и кольца. В ***разделимых кодах*** происходит простое разделение информационных разрядов и разрядов избыточных (контрольных). Коды, получившие наиболее широкое развитие, в схеме выделены жирными линиями.

10.1. ОСНОВНЫЕ ПРИНЦИПЫ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ

Общая блок-схема системы передачи цифровой информации, использующей помехоустойчивое кодирование, представлена на рис. 10.2.

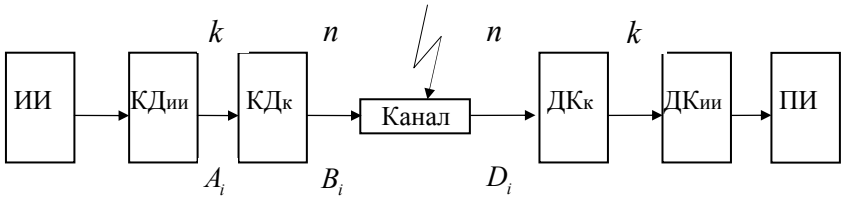


Рис. 10.2. Средства помехоустойчивого кодирования

Способность кода обнаруживать и исправлять ошибки, допущенные при передаче по каналу, обусловлена наличием избыточных символов. На вход кодирующего устройства канала ($КД_к$) поступает последовательность из k информационных двоичных символов. На выходе ей соответствует последовательность из n двоичных символов, причем $n > k$ (рис. 10.3).

Всего может быть 2^k различных входных последовательностей B_i и 2^n различных выходных последовательностей D_i. Из общего числа 2^n выходных последовательностей D_i только 2^k соответствуют входным и носят название разрешенных кодовых ком-

бинаций. Остальные D_i последовательности, число которых $2^n - 2^k$, для передачи полезной информации не используются и носят название запрещенных кодовых комбинаций.

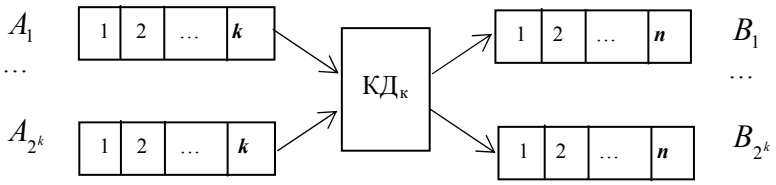


Рис. 10.3. Введение избыточных символов

Искажение информации в процессе передачи сводится к тому, что часть символов заменяется другими (ошибки типа сбоя 1 в 0 или наоборот). В результате разрешенные кодовые комбинации после передачи заменяются или трансформируются в другие. Возможные трансформации разрешенных кодовых комбинаций при прохождении по каналу с помехами представлены на рис. 10.4.

Отметим, что существует $2^k 2^n$ возможных переходов комбинаций из одной в другие. При этом число случаев переходов из разрешенных B_i в разрешенные D_i – $2^k 2^k$, среди которых правильных переходов – 2^k. Число случаев, когда переход осуществляется из разрешенных комбинаций в запрещенные, – $2^k(2^n - 2^k)$.

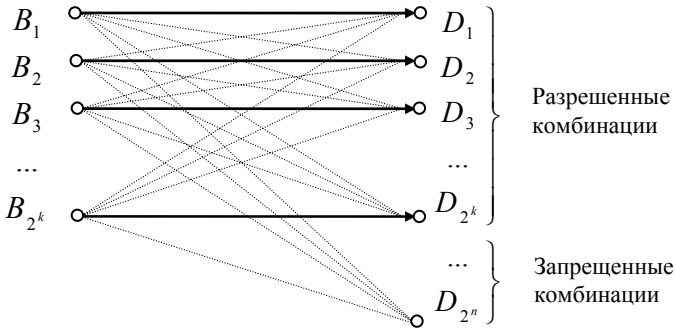


Рис. 10.4. Комбинации возможных переходов вследствие ошибок

Тогда относительная часть распознанных ошибок на выходе канала после декодирования или ***условная вероятность обнаружения ошибки*** при условии, что она имеет место, равна

$$p_{\text{обн}} = \frac{2^k(2^n - 2^k)}{2^k(2^n - 1)} = 1 - \frac{2^k - 1}{2^n - 1} \approx 1 - 2^{-(n-k)}.$$

Предполагается, что все комбинации имеют равную вероятность появления.

Рассмотрим случай исправления ошибок. Любой метод декодирования может быть представлен как правило разбиения множества полученных комбинаций D_i на непересекающиеся подмножества W_i, $i = \overline{1,2^k}$ или классы (рис. 10.5). При получении кодовой последовательности D_i, принадлежащей W_i, принимается решение о том, что передавалась по каналу последовательность B_i. Ошибка может быть исправлена правильно в том случае, если действительно передавалась B_i.

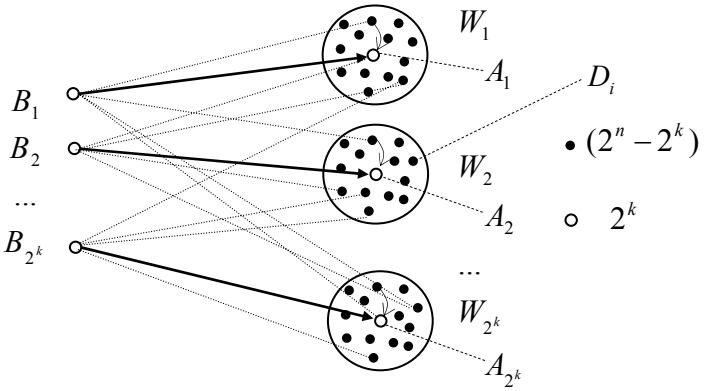


Рис. 10.5. Принцип исправления ошибок

Суть исправления ошибки состоит в том, что каждому классу W_i сопоставляется одна из входных последовательностей A_i длины k. Последовательность A_i длины k можно так закодировать в последовательность B_i длины n, что воздействие ошибок в канале

выводит ее из соответствующего ей класса W_i с вероятностью сколь угодно малой. Это позволяет определить на приемной стороне канала, к какому классу W_i принадлежит искаженная ошибками принятая последовательность D_i длины n, и тем самым точно восстановить исходную последовательность символов A_i длины k.

Условная вероятность правильного исправления ошибки при условии, что она имеет место, в этом случае равна

$$p_{\text{исп}} = \frac{2^n - 2^k}{2^k(2^n - 1)} \approx \frac{1}{2^k} - \frac{1}{2^n} = 2^{-k} - 2^{-n}.$$

Разбиение выходных последовательностей на подмножества зависит от того, какие ошибки должен исправить конкретный код. Существуют коды, предназначенные для обнаружения и исправления взаимно-независимых ошибок и пачек (пакетов) ошибок.

Взаимно-независимые – это ошибки в каждом разряде, которые не зависят от искажений в других разрядах и имеют постоянную вероятность искажения q. Типичным примером подобного искажения служит модель ДСК.

Будем называть количество искаженных разрядов r кратностью ошибки. Вероятность искажения r символов в кодовой комбинации длины n равна $q_r = \binom{n}{r} q^r (1-q)^{n-r}$. При $q << 1$ строятся коды для исправления ошибок малой кратности. Этот случай наиболее интересен практически с учетом достаточно высокой надежности передачи цифровых данных по современным каналам связи.

Пакетом ошибок длины l называется любая комбинация ошибок, число разрядов в которой между первой и последней ошибками, включая и эти ошибки, равно l. Например, пакет ошибок длиной 4 имеет вид 0…00110100.

Степень отличия двух кодовых слов оценивают через ***расстояние Хэмминга*** $d(X,Y)$. Это расстояние равно числу позиций, в которых различаются два кода $X = (x_1, x_2, ..., x_n)$ и $Y = (y_1, y_2, ..., y_n)$. Определим ***вес кода*** W как число единиц в

данном коде. Тогда расстояние по Хэммингу равно $d(X,Y)=W(x_1\oplus y_1,x_2\oplus y_2,...,x_n\oplus y_n)=W(X\oplus Y)$, где знак $\oplus$ означает суммирование по модулю 2.

Слово на выходе канала можно рассматривать как входное слово X, искаженное при передаче по каналу с шумом $Y=X\oplus E$, где $E=(e_1,e_2,...,e_n)$ – ***вектор ошибки***, содержащий единицы на тех позициях, которые были искажены при передаче.

Например, $Y=X\oplus E=(10100101)\oplus(10001000)=(00101101)$. Учитывая $d(X,Y)=W(X\oplus Y)$, получаем $d(X,Y)=d(X,X\oplus E)=$ $=W(X\oplus X\oplus E)=W(E)$, т.е. расстояние между переданным и принятым словами равно весу вектора ошибки.

Свяжем корректирующие способности кода, т.е. способность обнаруживать или исправлять ошибки определенной кратности, с ***минимальным расстоянием Хэмминга*** d этого кода. Под d понимают минимальное расстояние, взятое по всем парам кодовых комбинаций данного кода.

Рассмотрим трехразрядные кодовые последовательности. Если $d=1$, то все кодовые слова образуют равнодоступный код: 000, 001,…, 111. Этот код не обладает обнаруживающими и корректирующими способностями. Возьмем $d=2$ и образуем, например, такие две группы:

$$\underbrace{000\quad 011\quad 101\quad 110}_{\text{Разрешенные комбинации}}\qquad\qquad \underbrace{001\quad 100\quad 010\quad 111}_{\text{Запрещенные комбинации}}\ .$$

Этот код обнаруживает все ошибки кратности 1.

Для обнаружения ошибки кратности $r_{\text{о}}$ необходимо и достаточно, чтобы минимальное кодовое расстояние разрешенных кодовых комбинаций удовлетворяло неравенству $d\geq r_{\text{о}}+1$.

Возьмем $d=3$ и представим запрещенные комбинации в виде непересекающихся подмножеств (рис. 10.6). Этот код обнаруживает однократные и двухкратные ошибки, а также исправляет все однократные.

Для исправления любой ошибки кратности $r_и$ необходимо, чтобы минимальное расстояние кода удовлетворяло неравенству $d \geq 2r_и + 1$.

Очевидно, что код, который исправляет ошибки кратности $r_и$, может обнаружить ошибки кратности $r_о = 2r_и$.

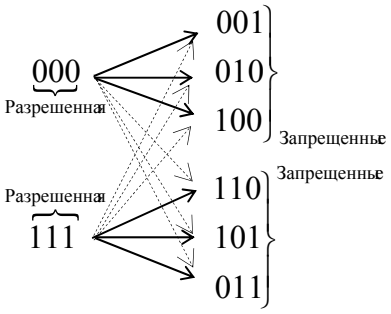


Рис. 10.6. Переходы разрешенных комбинаций в запрещенные

Справедливость введенных неравенств $d \geq r_о + 1$ и $d \geq 2r_и + 1$ может быть наглядно продемонстрирована с помощью рис. 10.7.

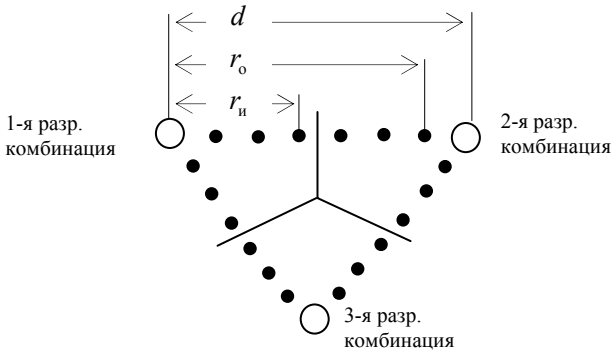


Рис. 10.7. К определению зависимости кратности обнаруживаемых и исправляемых ошибок от d

В весьма общем виде содержание современной алгебры можно охарактеризовать так: изучение алгебраических операций над элементами произвольной природы [17].

ОПРЕДЕЛЕНИЕ 10.1. Под ***алгебраической операцией***, определенной в данном множестве M, понимается соответствие, в силу которого произвольным элементам $x, y \in M$, взятым в определенном порядке (x, y), соотносится вполне определенный элемент $z \in M$.

Такое определение операции приводит к понятию ***внутреннего закона композиции***. Если же второй элемент берется из другого множества $y \in N$ и $(x, y) \to z \in M$, то такая алгебраическая операция носит название ***внешнего закона композиции***.

Задание на множестве M одного или нескольких законов композиции (внутренних или внешних) определяет на множестве M ***алгебраическую систему*** (структуру). Часто используемыми алгебраическими системами в соответствии с усложнением их математической структуры являются: группа; кольцо; поле.

Теория групп, колец и полей является математической основой построения корректирующих кодов, нашедших широкое применение на практике.

10.2. МАТЕМАТИЧЕСКОЕ ВВЕДЕНИЕ К ГРУППОВЫМ КОДАМ

ОПРЕДЕЛЕНИЕ 10.2. ***Группой*** G называется множество элементов, если для каждой его пары определена операция сложения и выполняются следующие аксиомы:

1) замкнутость $x + y = z;\ \ x, y \in G \Rightarrow z \in G$;
2) закон ассоциативности $(x + y) + z = x + (y + z)$;
3) существование ***нейтрального элемента*** $0 + x = x$;
4) существование ***обратного элемента*** $x \in G \Rightarrow -x \in G$; $x + (-x) = 0$.

Ясно, что группа обладает единственным нейтральным элементом. Если операция, определенная на множестве G, коммутативна $x + y = y + x$, то группа называется ***коммутативной*** (абелевой). Группа конечна, если число ее элементов конечно. ***Порядок группы*** – число ее элементов.

Например, множество всех n разрядных двоичных чисел образует абелеву группу порядка 2^n относительно операции поразрядного сложения. Нейтральный элемент группы: 00…000. Обратным элементом является сам элемент.

ОПРЕДЕЛЕНИЕ 10.3. Подмножество A множества G называется ***подгруппой группы*** G, если оно само является группой относительно операции, определенной в G.

При проверке того, является ли A подгруппой группы G, достаточно проверить следующее:

– содержится ли в A сумма любых двух элементов из A;

– содержит ли A вместе со всяким своим элементом и его обратный.

ОПРЕДЕЛЕНИЕ 10.4. Пусть $A-$ подгруппа группы G. Тогда множество всех элементов вида $x+x_A$, получающееся при сложении $x \in G$ с каждым элементом $x_A \in A$, называют ***левым смежным классом группы*** G по подгруппе A, определяемым элементом x, и обозначают $x+A$.

Аналогично определяется ***правый смежный класс группы*** G по подгруппе A: $A+x$. В абелевой группе левый смежный класс совпадает с правым смежным классом и называется просто смежным классом.

Например, множество $G=\{000,001,010,011,100,101,110,111\}$ является группой. Подмножество $A=\{000,001,010,011\}$ – подгруппа, так как имеет место замкнутости относительно операции поразрядного сложения $000 \oplus 001=001,...,010 \oplus 011=001.$

Смежные классы группы G по подгруппе A можно образовать, беря последовательно элементы из $x \in G$:

$$A \oplus 000=\{000,001,010,011\}, \quad A \oplus 100=\{100,101,110,111\};$$
$$A \oplus 001=\{001,000,011,010\}, \quad A \oplus 101=\{101,100,111,110\};$$
$$A \oplus 010=\{010,011,000,001\}, \quad A \oplus 110=\{110,111,100,101\};$$
$$A \oplus 011=\{011,010,001,000\}, \quad A \oplus 111=\{111,110,101,100\}.$$

Мы показали, что любые два смежных класса группы G по подгруппе A либо совпадают, либо не имеют ни одного общего элемента. Одним из смежных классов является сама подгруппа $A = A + 0$. Кроме того, $A + x_A = A$, так как $x_A \in A$.

ОПРЕДЕЛЕНИЕ 10.5. Число различных смежных классов группы G по подгруппе A называют ***индексом подгруппы*** A в группе G и обозначают $index_{AG}$.

Разложение группы по подгруппе приводит к разбиению группы G на непересекающиеся подмножества, а порядок группы равен порядку подгруппы умноженному на индекс этой подгруппы в группе: $|G| = |A| \cdot index_{AG}$.

Таким образом, беря последовательно некоторые элементы группы $x \in G$, которые не вошли в уже образованные классы, можно разложить всю группу G на смежные классы по подгруппе A. Элементы x, по которым производится разложение, называют ***образующими элементами*** смежных классов.

Например, дадим таблицу разложения группы четырехразрядных двоичных чисел $G = \{0000,0001,0010,0011,0100,0101,0110,$ $0111,1000,1001,1010,1011,1100,1101,1110,1111\}$ по подгруппе двухразрядных чисел $A = \{0000,0001,0010,0011\}$, представленных в едином формате:

$$\begin{vmatrix} 0000 \\ 0001 \\ 0010 \\ 0011 \end{vmatrix} A + 0000 \quad \begin{vmatrix} 0100 \\ 0101 \\ 0110 \\ 0111 \end{vmatrix} A + 0100 \quad \begin{vmatrix} 1000 \\ 1001 \\ 1010 \\ 1011 \end{vmatrix} A + 1000 \quad \begin{vmatrix} 1100 \\ 1101 \\ 1110 \\ 1111 \end{vmatrix} A + 1100 \ .$$

Группа G разложилась на четыре смежных классов по подгруппе A. В общем случае, число смежных классов при разложении группы n-разрядных чисел по подгруппе k-разрядных двоичных чисел равно:

$$index_{AG} = \frac{2^n}{2^k} = 2^{n-k} \ .$$

10.3. ПОСТРОЕНИЕ ГРУППОВОГО КОДА

Предлагается следующая процедура построения кода. Пусть требуемый объем кода Q равен необходимому числу передаваемых сообщений. Из соотношения

$$2^k - 1 \geq Q$$

определяется число информационных символов k. Каждой из $2^k - 1$ ненулевых кодовых комбинаций k-разрядного неизбыточного кода ставится в соответствие комбинация из n символов. Значение символов в проверочных разрядах, число которых $n - k$, устанавливается в зависимости от суммы по модулю 2 определенных информационных разрядов.

Множество 2^k k-разрядных комбинаций информационных разрядов образует подгруппу группы, состоящей из 2^n n-разрядных кодов. Множество 2^k n-разрядных комбинаций, состоящих из $n - k$ проверочных символов и k информационных разрядов, носит название разрешенных кодовых комбинаций и также представляет собой подгруппу группы 2^n n-разрядных кодов.

Множество разрешенных кодовых комбинаций и будет ***групповым кодом*** (n, k).

Разложим группу G n-разрядных чисел на смежные классы по подгруппе A разрешенных кодовых комбинаций. В этом разложении будет 2^{n-k} классов. Элементы каждого класса представляют собой сумму по модулю 2 образующих элементов данного класса и разрешенных кодов:

$$\underset{\substack{\text{Образующие} \\ \text{элементы}}}{x} \oplus \underset{\substack{\text{Разрешенные} \\ \text{коды}}}{A}.$$

За образующие элементы классов возьмем наиболее вероятные для данного канала вектора ошибок E_i, которые должны исправляться:

Опознаватели Образующие элементы

$$
\begin{array}{lccccc}
 & & & & & \\
x \in A & O_0 & \to & 0 & \oplus & A \\
x \notin A & O_1 & \to & E_1 & \oplus & A \\
 & O_2 & \to & E_2 & \oplus & A \\
 & \dots & & \dots & & \\
 & O_{2^{n-k}-1} & \to & E_{2^{n-k}-1} & \oplus & A
\end{array}
\left.\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right\} 2^{n-k} .
$$

При таком выборе образующих элементов в каждом смежном классе сгруппируются кодовые комбинации, получающиеся в результате воздействия на все разрешенные комбинации данного кода вектора E_i. Для исправления любой полученной на выходе канала кодовой комбинации теперь достаточно будет определить, к какому классу смежности относится полученная комбинация. Складывая ее затем по модулю 2 с образующим элементом класса, получаем истинную комбинацию.

Чтобы иметь возможность получить информацию о том, к какому смежному классу относится полученная кодовая комбинация, каждому смежному классу ставится в соответствие ***опознаватель***.

Исправление ошибки возможно при наличии взаимнооднозначного соответствия между множеством опознавателей и множеством смежных классов, т.е. множеством ошибок, подлежащих исправлению. Таким образом, количество подлежащих исправлению ошибок является определяющим для выбора числа избыточных символов $n-k$. Их должно быть достаточно для того, чтобы обеспечить необходимое число опознавателей.

Так, если мы хотим исправлять все одиночные ошибки, векторы которых 000…01, 000…10,…, 100…00 и число которых n, то число опознавателей (смежных классов) должно быть не меньше n, т.е.

$$2^{n-k}-1 \geq \binom{n}{1} = n .$$

В общем случае для исправления всех независимых ошибок кратности s включительно необходимо выполнение неравенства

$$2^{n-k}-1 \geq \sum_{i=1}^{s} \binom{n}{i}.$$

В приведенном соотношении указывается теоретический минимальный предел числа возможных символов, который далеко не всегда достижим на практике.

Рассмотрим пример. Допустим, что надо закодировать $Q = 15$ команд, которые передаются по каналу, в котором наиболее вероятны одиночные ошибки. Число информационных разрядов вычисляется из соотношения $2^k - 1 \geq 15$. Поскольку оно удовлетворяется при $k \geq 4$, выбираем наименьшее $k = 4$. Для исправления одиночных ошибок необходимо, чтобы $2^{n-4} - 1 \geq n$. Следовательно, наименьшая длина $n = 7$.

Соотношение $2^{n-k} - 1 \geq n$, которое справедливо для устранения всех одиночных ошибок, позволяет составить табл. 10.1 параметров группового кода (n, k).

Итак, строим групповой код (7,4). Число разрядов кода $n = 7$, число ошибок, подлежащих исправлению, тоже 7. Три избыточных разряда $n - k = 3$ позволяют использовать в качестве опознавателей трехразрядные числа:

Векторы ошибок		Опознаватели 3р 2р 1р
$E_0 = 0000000$		$O_0 = 000$
$E_1 = 0000001$		$O_1 = 001$
$E_2 = 0000010$		$O_2 = 010$
$E_3 = 0000100$		$O_3 = 011$
$E_4 = 0001000$	$\Rightarrow$	$O_4 = 100$
$E_5 = 0010000$		$O_5 = 101$
$E_6 = 0100000$		$O_6 = 110$
$E_7 = 1000000$		$O_7 = 111.$

При таком сопоставлении каждый опознаватель представляет собой двоичное число, указывающее номер разряда, в котором

произошла ошибка. Коды, в которых используется такой принцип сопоставления, известны как ***групповые коды Хэмминга***.

Таблица 10.1

Длина кода n	3	4	5	6	7	8	9	10	11	12	
Число информационных разрядов k	1	1	2	3	4	4	5	6	7	8	…
Число проверочных разрядов $n-k$	2	3	3	3	3	4	4	4	4	4	

Каждый разряд опознавателя определяется на приемном конце в результате проверки справедливости некоторых равенств. Проверочные символы подбираются так, чтобы сумма по модулю 2 всех символов, включая проверочные, входящие в каждое такое равенство, равнялась 0. Поэтому операцию определения символов опознавателя называют проверками на четность.

Предположим, что в результате первой проверки на четность для первого разряда опознавателя будет получена "1". Это может быть следствием ошибки в одном из разрядов, опознаватели которых в первом разряде имеют "1":

$$\underset{\text{Опознавателя}}{1\text{р}} \quad \rightarrow \quad x_1 \oplus x_3 \oplus x_5 \oplus x_7 = 0 .$$

Единица во втором разряде опознавателя может быть следствием ошибки в разрядах, опознаватели которых во втором разряде имеют "1":

$$\underset{\text{Опознавателя}}{2\text{р}} \quad \rightarrow \quad x_2 \oplus x_3 \oplus x_6 \oplus x_7 = 0 .$$

Аналогично, находим и третье равенство:

$$\underset{\text{Опознавателя}}{3\text{р}} \quad \rightarrow \quad x_4 \oplus x_5 \oplus x_6 \oplus x_7 = 0 .$$

Чтобы эти равенства удовлетворялись при отсутствии ошибок для любых значений информационных символов, в нашем распоряжении имеются три проверочных разряда. Мы должны так выбрать номера этих проверочных разрядов, чтобы каждый из них входил только в одно из равенств. Это обеспечит однозначное определение значений символов в проверочных разрядах при кодировании. Указанному условию удовлетворяют разряды, опознава-

тели которых имеют по одной единице: x_1, x_2, x_4. Тогда правила кодирования для кода (7,4) принимают вид

$$x_1 = x_3 \oplus x_5 \oplus x_7 ,$$
$$x_2 = x_3 \oplus x_6 \oplus x_7 ,$$
$$x_4 = x_5 \oplus x_6 \oplus x_7 .$$

Таким образом, информационные разряды: 3, 5, 6, 7. Контрольные: 1, 2, 4. Таблицу разрешенных кодовых комбинаций группового кода (7,4) можно представить в виде:

7	6	5	(4)	3	(2)	(1)
0	0	0	0	0	0	0
0	0	0	0	1	1	1
0	0	1	1	0	0	1
			...			
1	1	1	1	1	1	1

.

При больших n и k таблица оказывается слишком громоздкой.

Для полного определения группового кода достаточно записать только линейно независимые строки. Среди $2^k - 1$ ненулевых кодов линейно независимых строк всего k. Следовательно, ***образующая матрица*** кода (n,k) может быть представлена следующим образом:

$$M_{n,k} = \left| I_k^T : B_{n-k,k} \right| = \begin{vmatrix} 0 & 0 & \ldots & 0 & 1 & : & \\ 0 & 0 & \ldots & 1 & 0 & : & \\ & & \ldots & & & : & B_{n-k,k} \\ 0 & 1 & \ldots & 0 & 0 & : & \\ 1 & 0 & \ldots & 0 & 0 & : & \end{vmatrix},$$

где $B_{n-k,k}$ – ***проверочная матрица***.

Суммируя строки единичной матрицы в различных сочетаниях можно получить все $2^k - 1$ ненулевые комбинации k-разрядного

неизбыточного кода. Нулевая комбинация получается при суммировании любой строки с самой собой.

Для группового кода (7,4) образующая матрица имеет вид

$$M_{7,4} = \begin{array}{c} \overset{\text{Разряды}}{7\,6\,5\,3} \quad \overset{\text{кода}}{4\,2\,1} \\ \left| \begin{array}{l} 0001:011 \\ 0010:101 \\ 0100:110 \\ 1000:111 \end{array} \right| \end{array}.$$

Пример аппаратной реализации группового кода (7,4) приведен на рис. 10.8. Логическая схема кодирующего и декодирующего устройств выделена пунктиром. Простота реализации является одним из существенных факторов построения корректирующих кодов Хемминга.

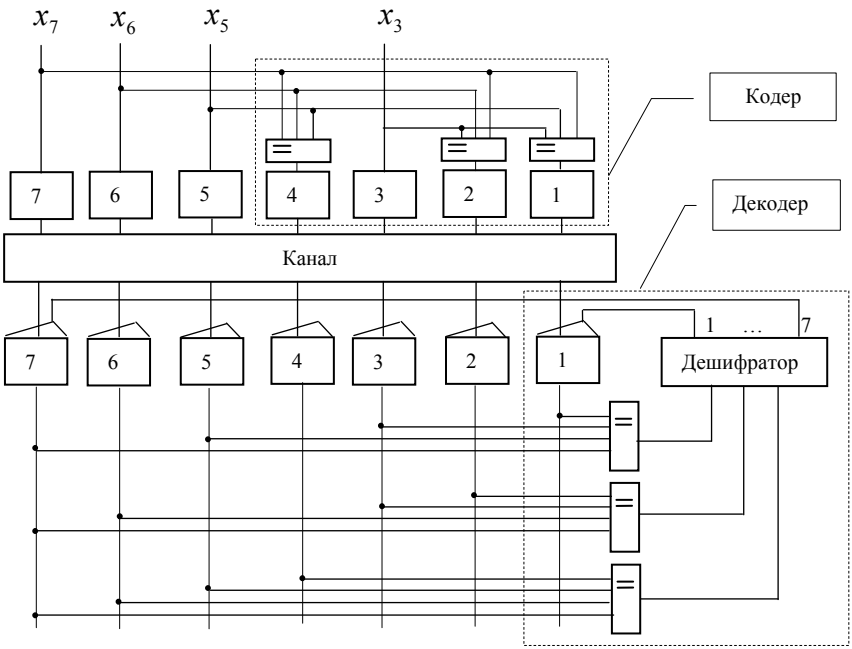


Рис. 10.8. Кодер и декодер группового кода (7,4)

10.4. МАТЕМАТИЧЕСКОЕ ВВЕДЕНИЕ К ЦИКЛИЧЕСКИМ КОДАМ

Среди многообразия алгебраических кодов выделяется класс кодов, образованных аналогично групповым, но на образующую матрицу которых наложено дополнительное условие цикличности (все строки матрицы получаются в результате циклического сдвига некоторой кодовой комбинации) и используется алгебраическая структура – кольцо.

Для описания циклических кодов удобна запись любого двоичного числа в виде многочлена от фиктивной переменной x:

$$100101 \to G_j(x) = 1 \cdot x^5 + 0 \cdot x^4 + 0 \cdot x^3 + 1 \cdot x^2 + 0 \cdot x^1 + 1 \cdot x^0 = x^5 + x^2 + 1.$$

В этом случае действия над кодами сводятся к операциям над многочленами. В множестве n - разрядных чисел или в множестве многочленов $G_j(x)$, степень которых не превышает $n-1$, введем операции сложения по модулю 2 и символического умножения. Операция символического умножения выполняется обычным образом, как в алгебре для многочленов, но с приведением подобных членов суммированием по модулю 2:

$$G_1(x) = x^3 + x + 1 \Rightarrow 1011,\ G_2(x) = x + 1 \Rightarrow 0011,$$

$$G_1(x)G_2(x) = x^4 + x^3 + x^2 + x + x + 1 = x^4 + x^3 + x^2 + 1 \Rightarrow 11101.$$

ОПРЕДЕЛЕНИЕ 10.6. ***Кольцом*** R называется множество элементов, если для каждой его пары определены операции сложения и умножения, а также выполняются следующие аксиомы:

1) относительно сложения R является абелевой группой;
2) замкнутость $a, b \in R \Rightarrow ab \in R$;
3) закон ассоциативности $a, b, c \in R \Rightarrow (ab)c = a(bc)$;
4) закон дистрибутивности $a, b, c \in R \Rightarrow (a+b)c = ac + bc$.

Сложение в кольце всегда коммутативно, а умножение не обязательно должно быть коммутативным. Коммутативное кольцо – это кольцо, в котором умножение коммутативно, т.е. $ab = ba$ для всех $a, b \in R$.

ОПРЕДЕЛЕНИЕ 10.7. ***Полем*** F называется множество элементов, если для каждой его пары определены операции сложения и умножения, а также выполняются следующие аксиомы:

1) множество F образует абелеву группу по сложению;
2) множество ненулевых элементов F образует абелеву группу по умножению (нейтральный элемент $1 \cdot a = a$, обратный элемент $a^{-1} \cdot a = 1$);
3) закон дистрибутивности $a, b, c \in R \Rightarrow (a+b)c = ac + bc$.

ОПРЕДЕЛЕНИЕ 10.8. Поле, состоящее из конечного числа элементов q, называют конечным или ***полем Галуа*** и обозначают $GF(q) = \{0,1,...,q-1\}$.

ОПРЕДЕЛЕНИЕ 10.9. Пусть F – некоторое поле. Подмножество $F_0 \subseteq F$ называется ***подполем***, если оно само является полем относительно наследуемых из F операций сложения и умножения. В этом случае исходное поле F называется ***расширением*** F_0.

Итак, двоичные кодовые комбинации $a_{n-1}...a_1a_0$ представляются в виде многочленов степени не более $n-1$ от переменной x с двоичными (из поля $GF(2)$) коэффициентами a_i, $i = \overline{0, n-1}$: $G_j(x) = a_{n-1}x^{n-1} + ... + a_1x + a_0$.

Уточним операцию символического умножения. Если произведение двух многочленов $G_1(x)G_2(x)$ имеет степень большую или равную n, то условие замкнутости относительно введенной операции символического умножения нарушается. Поэтому специально подбирают некоторый многочлен $P_0(x)$ степени n и результатом операции символического умножения считают остаток от деления $\left[\frac{G_1(x)G_2(x)}{P_0(x)}\right]$, степень которого не более $n-1$.

Для того чтобы найти вид многочлена $P_0(x)$, произведем умножение, например, многочлена степени $n-1$ на x:

$$G_j(x) = (x^{n-1} + x^{n-2} + ... + x + 1) \cdot x = x^n + x^{n-1} + ... + x^2 + x.$$

Чтобы результат умножения соответствовал кодовой комбинации n - разрядного числа, необходимо заменить x^n на 1, т.е.

$x^n + x^{n-1} + \ldots + x^2 + x \to x^{n-1} + \ldots + x^2 + x + 1$. Такая замена эквивалентна получению остатка от деления $G_j(x)$ на $x^n + 1$:

$$\begin{array}{l|l} x^n + x^{n-1} + \ldots + x^2 + x & x^n + 1 \\ \hline x^n + 1 & 1 \\ \hline x^{n-1} + \ldots + x^2 + x + 1 & \end{array}.$$

Поэтому естественно в качестве $P_0(x)$ взять $x^n + 1$.

При заданных операциях поразрядного суммирования и символического умножения все множество n - разрядных чисел или многочленов степени не более $n-1$ образует ***коммутативное кольцо*** R.

ОПРЕДЕЛЕНИЕ 10.10. Кольцо многочленов по модулю $P(x)$ является ***полем*** $GF(2^m)$ тогда и только тогда, когда $P(x)$ – ***неприводимый многочлен*** (простой) и $\deg P(x) = m$.

Например, поле $GF(2^2)$ состоит из элементов $\{0, 1, x, x+1\}$ для неприводимого многочлена $P(x) = x^2 + x + 1$ и является расширением поля $GF(2)$, а $GF(2)$, в свою очередь, подполем $GF(2^2)$.

Следует отметить, что при выбранных нами операциях сложения и умножения образованное кольцо R одновременно является конечным полем, так как оно содержит единичный элемент по умножению и обратные элементы.

Все множество n - разрядных кодовых комбинаций при выбранных операциях может также рассматриваться как алгебра многочленов над полем $GF(2)$ по модулю $x^n + 1$.

Выделим в R подмножество таких многочленов, которые кратны некоторому многочлену – $g(x) \in R$, т.е. делятся на него без остатка: $G_i(x) = 0 \bmod g(x)$. Такое подмножество многочленов называют ***идеалом***, а $g(x)$ именуют ***порождающим многочленом идеала***.

Количество различных элементов в идеале определяется видом

порождающего многочлена. Если $g(x)=0$, то весь идеал пуст; если $g(x)=1$, то в идеал войдут все многочлены кольца; если же $g(x)$ – неприводимый многочлен степени $n-k$, то число элементов идеала равно 2^k.

ОПРЕДЕЛЕНИЕ 10.11. ***Циклический код*** – идеал в кольце многочленов R, образованный неприводимым порождающим многочленом $g(x)$ степени $n-k$.

Остается только выбрать $g(x)$, способный породить код с заданными свойствами. Все многочлены циклического кода, соответствующие его кодовым комбинациям, должны делиться на $g(x)$ без остатка. Для этого достаточно, чтобы на $g(x)$ без остатка делились все многочлены, составляющие образующую матрицу циклического кода. Многочлены $G_i(x)$ кода получаются циклическим сдвигом, что соответствует последовательному умножению $g(x)$ на x с приведением по модулю x^n+1:

$$G_i(x)=\left[\frac{g(x)\cdot x^i}{x^n+1}\right].$$

Например, образующая матрица циклического кода (7,4), образованного порождающим многочленом $g(x)=x^3+x+1$, имеет вид

$$M_{7,4}=\begin{vmatrix}0&0&0&1&:&0&1&1\\0&0&1&0&:&1&1&0\\0&1&0&0&:&1&1&1\\1&0&0&0&:&1&0&1\end{vmatrix}.$$

Итак, многочлен $G_i(x)$ циклического кода является остатком от деления, для которого справедливы следующие выражения:

$$G_i(x)=g(x)\cdot x^i \bmod(x^n+1),$$

$$g(x)\cdot x^i=A_i(x)\cdot(x^n+1)+G_i(x).$$

Следовательно, все многочлены матрицы, т.е. все многочлены кода, будут делиться без остатка на $g(x)$, если на $g(x)$ будет делиться без остатка x^n+1. Предполагать, что многочлен $A_i(x)$,

соответствующий целой части, должен делиться на $g(x)$, никаких оснований нет.

Таким образом, многочлен $g(x)$ должен быть делителем x^n+1, чтобы породить идеал или циклический код:

$$x^n+1=0 \bmod g(x).$$

Поскольку для кольца справедливы все свойства группы, а для идеала – подгруппы, то R можно разложить на смежные классы, которые в данном случае называются ***классами вычетов по идеалу***.

Рассмотрим пример. Пусть имеется кольцо многочленов R со степенью меньше или равной 2, соответствующее кольцу трехразрядных двоичных чисел (табл. 10.2). Делителями $P_0(x)=x^3+1$ являются $x+1$ и x^2+x+1, так что в качестве порождающего многочлена можно выбрать

$$g_1(x)=x+1,$$
$$g_2(x)=x^2+x+1.$$

Проведем теперь разложение кольца многочленов R по идеалам с порождающими многочленами $g_1(x)$ и $g_2(x)$.

Таблица 10.2

Двоичные числа	Многочлен $G_j(x)$	Остатки от деления $G_j(x)$ на	
		$g_1(x)=x+1$	$g_2(x)=x^2+x+1$
000	0	0	0
001	1	1	1
010	x	1	x
011	$x+1$	0	$x+1$
100	x^2	1	$x+1$
101	x^2+1	0	x
110	x^2+x	0	1
111	x^2+x+1	1	0

Первую строку разложения образует идеал, причем нулевой элемент располагается крайним слева. В качестве образующего

первого класса вычетов можно выбрать любой многочлен, не принадлежащий идеалу. Остальные элементы данного класса вычетов образуются путем суммирования образующего многочлена с каждым многочленом идеала.

Разложение R по идеалу с $g_1(x)$ дает два класса вычетов:

0	$x+1$	x^2+1	x^2+x
(1)	x	x^2	x^2+x+1

Разложение R по идеалу с $g_2(x)$ дает четыре класса вычетов:

0	x^2+x+1
(1)	x^2+x
(x)	x^2+1
($x+1$)	x^2

В общем случае число элементов идеала, порожденного многочленом $g(x)$ степени $n-k$, равно 2^k, а число классов вычетов по этому идеалу равно $\frac{2^n}{2^k}=2^{n-k}$.

Корректирующая способность циклического кода будет тем выше, чем больше остатков будет образовано при делении кодовых многочленов на многочлен $g(x)$.

Наибольшее число остатков, число классов вычетов по идеалу, равное $2^{n-k}-1$, исключая нулевой многочлен, может обеспечить только неприводимый многочлен с $\deg g(x)=n-k$.

10.5. ЦИКЛИЧЕСКИЙ КОД, ОБНАРУЖИВАЮЩИЙ И ИСПРАВЛЯЮЩИЙ ОШИБКИ

Любая принятая по каналу связи кодовая комбинация $G_j'(x)$, возможно, содержащая ошибку (рис. 10.9), может быть представлена в виде суммы по модулю 2 неискаженной комбинации циклического кода $G_j(x)$ и вектора ошибки $E_j(x)$:

$$G_j'(x)=G_j(x)\oplus E_j(x).$$

При делении $G_j'(x)$ на образующий многочлен $g(x)$ остаток,

указывающий на наличие ошибки, будет обнаружен только в том случае, если многочлен $E_j(x)$ не делится на $g(x)$, так как неискаженная комбинация $G_j(x)$ на $g(x)$ делится без остатка.

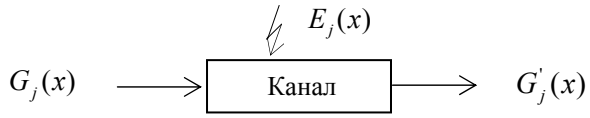


Рис. 10.9. Канал с шумом

По заданному объему кода однозначно определяется число информационных разрядов k. Далее необходимо найти наименьшее n, обеспечивающее обнаружение или исправление ошибок заданной кратности. В случае циклического кода эта проблема сводится к нахождению нужного многочлена $g(x)$.

Начнем рассмотрение с простейшего циклического кода, обнаруживающего все одиночные ошибки.

10.5.1. Циклический код, обнаруживающий одиночные ошибки

Вектор единичной ошибки будет иметь 1 в искаженном разряде и 0 во всех остальных разрядах. Следовательно, вектору единичной ошибки $E_j = (0,0,...,\underset{i\text{ разряд}}{1},...,0)$ соответствует многочлен $E_j(x) = x^{i-1}$. Он не должен делиться на $g(x)$. Среди неприводимых многочленов в разложении $x^n + 1$ минимальная степень у многочлена $x+1$. Остаток от деления на $x+1$ равен 0 или 1. В этом случае кольцо разбивается на два смежных класса: идеал с порождающим многочленом и класс вычетов. Таким образом, при любом k необходим только один проверочный разряд и $n = k+1$.

Значение символа в этом контрольном разряде обеспечивает четность числа единиц в любой разрешенной кодовой комбинации, а, следовательно, и делимость ее многочлена на $x+1$. Полученный циклический код способен обнаруживать не только все одиночные ошибки, но и ошибки в любом нечетном числе разрядов.

Например, при $n = 4$ и $k = 3$ образующая матрица циклического кода имеет вид

$$M_{4,3} = \begin{vmatrix} 0 & 0 & 1 & : & 1 \\ 0 & 1 & 0 & : & 1 \\ 1 & 0 & 0 & : & 1 \end{vmatrix}.$$

10.5.2. Циклический код, исправляющий одиночные ошибки и обнаруживающий двойные

Ошибку можно исправить, если каждой единичной ошибке в любом разряде сопоставить свой класс вычетов и свой опознаватель. В циклическом коде опознаватель – это остаток от деления многочлена $E_j(x)$ на $g(x)$. Поэтому $g(x)$ должен обеспечить требуемое число различных остатков или классов вычетов. Наибольшее число остатков дает неприводимый многочлен степени $n-k$. В этом случае число смежных классов, исключая идеал, равно $2^{n-k}-1$.

Необходимое условие исправления любой одиночной ошибки есть $2^{n-k}-1 \geq n$. Отсюда находится $\deg g(x) = n-k \geq \log_2(n+1)$. Несколько значений n, k сведены в табл. 10.3, учитывающие условие $x^n + 1 = 0 \bmod g(x)$.

Таблица 10.3

$m = n-k$	2	3	4	5	6	7	8	9	10	…
n	3	7	15	31	63	127	255	511	1023	…
k	1	4	11	26	57	120	247	502	1013	…

Можно показать, что любой двучлен $x^n + 1 = x^{2^m - 1} + 1$ может быть представлен произведением всех без исключения неприводимых многочленов, степени которых являются делителями числа m, от 1 до m включительно.

Так для $n = 15, k = 11, m = 4$ двучлен $x^{15} + 1$ можно записать в виде произведения всех неприводимых многочленов, степени которых являются делителями числа $m = 4$, т.е. 1, 2, 4. Существует таблица неприводимых многочленов. В ней находим один многочлен первой степени $x+1$, один второй $x^2 + x + 1$, три четвертой: $x^4 + x + 1, x^4 + x^3 + 1, x^4 + x^3 + x^2 + x + 1$. Таким образом,

$$x^{15}+1=(x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1).$$

Один из этих сомножителей может быть принят за образующий многочлен кода (порождающий многочлен идеала). Возьмем $x^4+x^3+1\to 11001$. Необходимо убедиться, что каждому вектору одиночных ошибок соответствует свой остаток. Проще всего найти все остатки от деления на $g(x)$, если взять код 1000…00 и начать делить его на 11001, выписывая все промежуточные остатки:

1000000000000000... | 11001	Остатки:
⊕11001	1.1001
10010	2.1011
⊕11001	3.1111
10110	4.0111
⊕11001	5.1110
11110	6.0101
⊕11001	...
...	15.1000

При последующем делении остатки повторяются.

С тем же успехом можно использовать многочлен x^4+x+1. Многочлен же $x^4+x^3+x^2+x+1$ использовать нельзя, так как при проверке числа различных остатков обнаруживается, что их у него не 15, а только 5:

1000000000000000... | 11111	Остатки:
⊕11111	1.1111
11110	2.0001
⊕11111	3.0010
10	4.0100
100	5.1000
1000	
10000	
⊕11111	
11110	
...	

Этот факт объясняется тем, что многочлен $x^4+x^3+x^2+x+1$ входит не только в разложение $x^{15}+1$, но и в разложение $x^5+1=(x+1)(x^4+x^3+x^2+x+1)$.

Поэтому целесообразно выбирать такой многочлен $g(x)$, который, являясь делителем x^n+1, не входит ни в одно разложение ни одного двучлена x^r+1, для которого $r<n$. Другими словами, многочлен $g(x)$ должен быть не только неприводимым, но еще и ***примитивным***. Если $g(x)$ неприводим и примитивен, то говорят, что он принадлежит максимальному показателю $e=2^{n-k}-1$:

$$x^e+1=0 \bmod g(x).$$

В табл. 10.4 приведены основные характеристики некоторых циклических кодов Хэмминга, способных исправлять одиночные ошибки или обнаруживать все одиночные и двойные ошибки.

Таблица 10.4

(n,k)	$m=n-k$	$g(x)$
(3,1)	2	x^2+x+1
(7,4)	3	$x^3+x+1,\ x^3+x^2+1$
(15,11)	4	$x^4+x+1,\ x^4+x^3+1$
(31,26)	5	$x^5+x^2+1,\ x^5+x^3+1$

Многочлен, соответствующий вектору двойной ошибки, имеет вид $E(x)=x^i+x^j=x^i(x^{j-i}+1),\ j>i$. Поскольку $j-i<n$, а $g(x)$ не кратен x и принадлежит показателю степени n, то $E(x)$ не делится на $g(x)$, что и позволяет обнаружить двойные ошибки.

10.5.3. Обнаружение ошибок высокой кратности

Для определения порождающего многочлена циклического кода, который обнаруживает ошибки более высокой кратности, можно воспользоваться правилом Хэмминга. Если $g_{s-1}(x)$ – порождающий многочлен кода, обнаруживающего ошибки кратности

$s-1$, то порождающий многочлен кода, обнаруживающего ошибки кратности s, равен $g_s(x) = g_{s-1}(x)(x+1)$.

Умножение $g_{s-1}(x)$ на $x+1$ равносильно введению дополнительной проверки на четность. При этом число символов в комбинациях кода за счет добавления еще одного проверочного символа увеличивается до $n+1$.

10.5.4. Исправление ошибок высокой кратности

Процедуру отыскания порождающего многочлена кода, исправляющего s - кратные ошибки, можно представить как итеративный процесс:

1. По заданному числу информационных символов k_1 подбирается n_1, т.е. определяются параметры (n_1, k_1) кода, исправляющего одиночные ошибки;

2. Полученное значение n_1 интерпретируется как новое значение k_2, и п.1 повторяется;

3. Повторяя п.1 и п.2 s раз, получим длину кодовой комбинации n_s циклического кода, исправляющего ошибки кратности до s включительно;

4. По заданному значению k_1 и значению n_s можно определить степень порождающего многочлена $m = n_s - k_1$, а по таблице неприводимых и примитивных многочленов выбрать подходящий $g_s(x)$.

Найденный таким образом $g_s(x)$ порождает неоптимальный, с точки зрения числа проверочных символов, циклический код (n_s, k_1). Более того, s - кратные ошибки исправляются только в информационных разрядах k_1. Это не всегда приемлемо.

10.5.5. Обнаружение пакетов ошибок

Любой циклический код, образованный с помощью полинома степени $n-k$, обнаруживает любой пакет ошибок длиной

$l \le n-k$.

Пакет ошибок в $n-k$ разрядов описывается многочленом степени $x^i(x^{n-k-1}+x^{n-k-2}+...+1)$. Многочлен же степени $n-k-1$ на многочлен степени $n-k$ не делится.

10.5.6. Параметры некоторых циклических кодов

Циклический код длиной $n=2^{n-k}-1$, порождаемый неприводимым многочленом степени $n-k$, исправляет однократные ошибки и называется ***циклическим кодом Хэмминга***. Таблица примитивных многочленов различной степени частично приводилась.

Наилучшими из известных циклических кодов для исправления независимых ошибок произвольной кратности являются ***коды Боуза–Чоудхури–Хоквингема*** (БЧХ) [7]. В этих кодах число проверочных разрядов $n-k=\deg g(x)\le mt$ при длине кода $n=2^m-1$, где t количество независимых исправляемых ошибок, а m любое целое число.

Образующий полином кода БЧХ является наименьшим общим кратным (НОК) так называемых минимальных многочленов $F_i(x)$, максимальный порядок которых равен $\phi \le d-2$:

$$g(x)=\text{НОК}\{F_1(x),...,F_i(x),...,F_\phi(x)\},$$

где $d \ge 2t+1$, а $i=1,3,5,...,\phi$.

ОПРЕДЕЛЕНИЕ 10.12. Пусть $GF(q)$ – некоторое поле, $GF(Q)$ – расширение поля $GF(q)$, построенное по неприводимому многочлену $P(z)$ степени n, и β – элемент $GF(Q)$. Тогда простой многочлен $F(x)$ наименьшей степени, для которого $F(\beta) \bmod P(z)=0$, называется ***минимальным многочленом*** элемента β.

ОПРЕДЕЛЕНИЕ 10.13. ***Примитивный элемент*** поля $GF(Q)$ – это такой элемент α поля, все элементы которого, за исключением нуля, могут быть представлены в виде степени элемента α.

В табл. 10.5 задано представление поля $GF(2^4)$ как расширение поля $GF(2)$, построенное по примитивному многочлену $P(z) = z^4 + z + 1$. Очевидно, что таблица сформирована для случая, когда $m = 4$. В нее включены также минимальные многочлены над *GF*(2) для всех элементов из поля $GF(2^4)$.

Таблица 10.5

Поле $GF(2^4)$ в виде			Минимальные многочлены при α^i
степени	многочлена	кода	
0	0	0000	
α	1	0001	$x+1$
α^1	z	0010	$F_1(x) = x^4 + x + 1$
α^2	z^2	0100	$x^4 + x + 1$
α^3	z^3	1000	$F_3(x) = x^4 + x^3 + x^2 + x + 1$
α^4	$z+1$	0011	$x^4 + x + 1$
α^5	$z^2 + z$	0110	$F_5(x) = x^2 + x + 1$
α^6	$z^3 + z^2$	1100	$x^4 + x^3 + x^2 + x + 1$
α^7	$z^3 + z + 1$	1011	$F_7(x) = x^4 + x^3 + 1$
α^8	$z^2 + 1$	0101	$x^4 + x + 1$
α^9	$z^3 + z$	1010	$x^4 + x^3 + x^2 + x + 1$
α^{10}	$z^2 + z + 1$	0111	$x^2 + x + 1$
α^{11}	$z^3 + z^2 + z$	1110	$x^4 + x^3 + 1$
α^{12}	$z^3 + z^2 + z + 1$	1111	$x^4 + x^3 + x^2 + x + 1$
α^{13}	$z^3 + z^2 + 1$	1101	$x^4 + x^3 + 1$
α^{14}	$z^3 + 1$	1001	$x^4 + x^3 + 1$

Пусть $\beta = z^3$ – элемент $GF(2^4)$. Тогда $F(x) = x^4 + x^3 + x^2 + x + 1$ – минимальный многочлен элемента z^3,

так как $F(\beta)\bmod P(z)=[(z^3)^4+(z^3)^3+(z^3)^2+(z^3)+1]\bmod(z^4+z+1)\equiv 0$.

Коды БЧХ строятся по заданным n и t. Значение k (число информационных разрядов) неизвестно, пока не найден порождающий многочлен кода $g(x)$.

Значения минимальных многочленов для степени $m=2\div 6$ приведены в табл. 10.6 [8].

Таблица 10.6

m		2	3	4	5	6 (t)
	1	111 (1)	1011 (1)	10011 (1)	100101 (1)	1000011 (1)
	3		1101 (3)	11111 (2)	111101 (2)	1010111 (2)
	5		1	111 (3)	110111 (3)	1100111 (3)
	7			11001 (7)	101111 (5)	1001001 (4)
i	9			1	101001 (7)	1101 (5)
	11			1	111011 (15)	1101101 (6)
	13			1	1	1011011 (7)
	15				1	1110101 (10)
	17					1100111 (10)
	19					1011011 (10)

Например, порождающий многочлен кода длины $n=15$, служащий для исправления трехкратных ошибок $t=3$, имеет вид

$$g(x)=\text{НОК}\{F_1(x)\to 10011, F_3(x)\to 11111, F_5(x)\to 111\}=$$
$$=x^{10}+x^8+x^5+x^4+x^2+x+1.$$

Циклический код БЧХ (15,5) имеет $k=5$ информационных разрядов и полностью определяется порождающей матрицей:

$$M_{15,5}=\begin{vmatrix} 00001 & : & 0100110111 \\ 00010 & : & 1001101110 \\ 00100 & : & 0111101011 \\ 01000 & : & 1111010110 \\ 10000 & : & 1010011011 \end{vmatrix}.$$

Циклические коды Файра позволяют исправить одиночную вспышку (пакет) ошибок. Порождающий многочлен этих кодов имеет вид $g(x)=p(x)(x^c+1)$, где $p(x)$ – неприводимый много-

член степени m. Длина кода n равна наименьшему общему кратному e и c:

$$n = \text{НОК}\{e, c\},$$

где e – максимальный показатель для $p(x)$, т.е. наименьшее число, при котором $x^e + 1 = 0 \bmod p(x)$. Причем c не должно делиться нацело на $e = 2^m - 1$. В частности, если $p(x)$ является примитивным многочленом, то $e = 2^m - 1$. Число проверочных разрядов равно $n - k = m + c$.

Код позволяет исправить вспышку ошибок длиной l или менее и одновременно обнаружить любую вспышку ошибок длиной $t \geq l$, если $c \geq t + l - 1$ и $m \geq l$.

Например, код Файра, генерируемый $g(x) = (x^4 + x + 1)(x^7 + 1)$, содержит $n = 105$ разрядов, из которых 11 являются проверочными, и позволяет обнаружить и исправить произвольную вспышку ошибок длиной 4 и менее.

В табл. 10.7 приведены параметры некоторых циклические коды Файра (n, k).

Таблица 10.7

(n, k)	$l = m$
(9,4)	2
(35,27)	3
(105,94)	4
(279,265)	5
(693,676)	6
(1651,1631)	7

10.6. МЕТОДЫ ПОСТРОЕНИЯ ЦИКЛИЧЕСКОГО КОДА ХЭММИНГА

Существуют различные методы построения циклического кода. Рассмотрим основные из них. При классификации циклических кодов были введены понятия неразделимых и разделимых циклических кодов.

10.6.1. Неразделимый циклический код

Разрешенные n-разрядные комбинации циклического кода получаются умножением многочленов, соответствующих k-разрядным информационным комбинациям, на порождающий многочлен: $\widetilde{a}_i(x) = a_i(x) \cdot g(x)$. Кодер при этом легко реализуется, однако при декодировании возникают трудности.

Пусть необходимо построить код для исправления одиночных ошибок. Остатки от деления многочлена ошибки $E_j(x) = x^{j-1}$ на $g(x)$ являются признаком ошибки в соответствующем j-м разряде. Декодирующее устройство должно осуществлять деление принятого кода $\widetilde{a}_i(x)$ на $g(x)$. В случае равенства остатка от деления нулю делается вывод о правильной передаче. Если же остаток не равен нулю, то его величина указывает разряд, где произошла ошибка. Недостатком такого кодирования является также то, что информационные разряды не содержатся в явном виде.

Рассмотрим пример кода (7,4) с порождающим многочленом $g(x) = x^3 + x^2 + 1$ (табл. 10.8).

Таблица 10.8

Информационный код	$a_i(x)$	$\widetilde{a}_i(x) = a_i(x) \cdot g(x)$	Циклический код
0001	1	$x^3 + x^2 + 1$	0001101
0010	x	$x^4 + x^3 + x$	0011010
0100	x^2	$x^5 + x^4 + x^2$	0110100
1000	x^3	$x^6 + x^5 + x^3$	1101000

Пусть передается код 0111, после кодирования на выходе кодера будет код

$$(x^2 + x + 1)g(x) = (x^2 + x + 1)(x^3 + x^2 + 1) = x^5 + x + 1 \to 0100011.$$

Допустим, что на приемном конце принимается кодовая комбинация с искажением четвертого разряда: $x^5 + x^3 + x + 1 \to 0101011$. Эта комбинация делится на $g(x) = x^3 + x^2 + 1 \to 1101$:

$$
\begin{array}{l|l}
x^5+x^3+x+1 & x^3+x^2+1 \\
\underline{x^5+x^4+x^2} & \underline{} \\
x^4+x^3+x^2+x+1 & x^2+x \\
\underline{x^4+x^3+x} & \\
x^2+1 \to 101 &
\end{array}.
$$

По таблице ошибок (табл. 10.9) находим, что искажение произошло в четвертом разряде. Этот разряд исправляется и деление повторяется, при этом получается правильный код:

$$
\begin{array}{l|l}
x^5+x+1 & x^3+x^2+1 \\
\underline{x^5+x^4+x^2} & \underline{} \\
x^4+x^2+x+1 & x^2+x+1 \to 0111 \\
\underline{x^4+x^3+x} & \\
x^3+x^2+1 & \\
\underline{x^3+x^2+1} & \\
0 &
\end{array}.
$$

Таблица 10.9

Разряд j	$E_j(x)$	$\left[\frac{E_j(x)}{g(x)}\right]$	Код опознаватель
1	1	1	001
2	x	x	010
3	x^2	x^2	100
④	x^3	x^2+1	101
5	x^4	x^2+x+1	111
6	x^5	$x+1$	011
7	x^6	x^2+x	110

10.6.2. Разделимый циклический код

Более широкое распространение на практике получили разделимые циклические коды. Под информационные символы принято отводить старшие k разрядов, а под проверочные $n-k$ младших разрядов. Чтобы получить такой код, применяется следующая процедура:

$$\widetilde{a}_i(x) = a_i(x) \cdot x^{n-k} \oplus r_i(x).$$

Разделим левую и правую части этого уравнения на $g(x)$:

$$\frac{\widetilde{a}_i(x)}{g(x)} = \frac{a_i(x) \cdot x^{n-k}}{g(x)} \oplus \frac{r_i(x)}{g(x)} = 0.$$

Поскольку $\deg r_i(x) < \deg g(x)$, то $\frac{r_i(x)}{g(x)} = r_i(x)$. Тогда многочлен, соответствующий проверочным разрядам, примет вид

$$r_i(x) = \left[\frac{a_i(x)x^{n-k}}{g(x)} \right].$$

Таким образом, циклический код можно строить, приписывая к каждой неизбыточной комбинации информационных символов остаток от деления этой комбинации на $g(x)$. Для кодов, число k для которых удовлетворяет условию $k > n-k$, рассмотренный метод особенно прост.

Рассмотрим опять пример кода (7,4), но с порождающим многочленом $g(x) = x^3 + x + 1$ (табл. 10.10).

Таблица 10.10

Код	$a_i(x)x^3$	$r_i(x) = \left[\frac{a_i(x)x^3}{g(x)} \right]$	Циклический код $\widetilde{a}_i(x) = a_i(x)x^3 + r_i(x)$	
0001 000	x^3	$x+1$	$x^3 + x + 1$	0001:011
0010 000	x^4	$x^2 + x$	$x^4 + x^2 + x$	0010:110
0100 000	x^5	$x^2 + x + 1$	$x^5 + x^2 + x + 1$	0100:111
1000 000	x^6	$x^2 + 1$	$x^6 + x^2 + 1$	1000:101

На приемном конце при декодировании $\widetilde{a}_i(x)$ делится на $g(x)$. Если результат деления – нуль, то, отбрасывая три младших разряда, получаем передаваемый код. Если же результат деления не нулевой, то остаток от деления указывает номер разряда, где произошла ошибка в соответствии с табл. 10.11.

Пусть передавалась комбинация 1001. После кодера имеем 1001110. Пусть на приемном конце получена последовательность 1011110, т.е. ошибка произошла в пятом разряде. Производим деление:

$$\begin{array}{l|l} x^6+x^4+x^3+x^2+x & x^3+x+1 \\ \hline x^6+x^4+x^3 & x^3 \\ \hline x^2+x \rightarrow 110. & \end{array}$$

Таблица 10.11

Разряд j	$E_j(x)$	$\left[\frac{E_j(x)}{g(x)}\right]$	Код - опознаватель
1	1	1	001
2	x	x	010
3	x^2	x^2	100
4	x^3	$x+1$	011
(5)	x^4	x^2+x	110
6	x^5	x^2+x+1	111
7	x^6	x^2+1	101

Остаток 110 указывает, что ошибка произошла в пятом разряде. Этот разряд исправляется, а три младших проверочных разряда отбрасываются.

Однако не исключается возможность возникновения в кодовых комбинациях многократных ошибок, что может привести к ложным исправлениям или необнаружению ошибок при трансформации одной разрешенной комбинации в другую.

Разделимый циклический код так же, как и групповой, удобно записывать в матричной форме:

$$M_{n,k} = \left| I_k^T : B_{n-k,k} \right| = \begin{vmatrix} 0001 & : & 011 \\ 0010 & : & 110 \\ 0100 & : & 111 \\ 1000 & : & 101 \end{vmatrix},$$

где $B_{n-k,k} = \left\| \dfrac{a_i(x) \cdot x^{n-k}}{g(x)} \right\|$ – матрица остатков (проверочных символов).

10.7. ПОСТРОЕНИЕ ОДНОКАНАЛЬНЫХ КОДИРУЮЩИХ И ДЕКОДИРУЮЩИХ УСТРОЙСТВ ЦИКЛИЧЕСКОГО КОДА ХЭММИНГА

Для кодирования и декодирования циклических кодов применяются многотактные линейные схемы, называемые часто ***фильтрами*** или сдвигающими регистрами с обратными связями. Основными компонентами этих схем являются элементы задержки на один такт и сумматоры по модулю 2 (рис. 10.10).

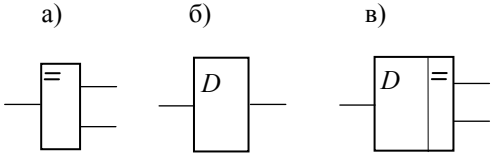


Рис. 10.10. Компоненты фильтра:
а – сумматор по модулю 2; б – элемент задержки (*D* триггер);
в – объединенный элемент

Структура фильтра описывается матрицей связей между его элементами

$$M = \left\| m_{ij} \right\|, \quad i, j = \overline{1, r},$$

где r – количество элементов задержки в фильтре, а $m_{ij} = 1$, если выход j-го элемента задержки связан с входом i-го элемента задержки, в противном случае $m_{ij} = 0$.

Например, матрица $M = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{vmatrix}$ соответствует фильтру, изображенному на рис. 10.11.

Вектор-столбцам матрицы связей M поставим в соответствие полиномы $\mu_j(x)$ степени не более $r-1$ таким образом, что

$$\mu_j(x) = \sum_{i=1}^{r} m_{ij} x^{i-1}, \quad j = \overline{1,r}.$$

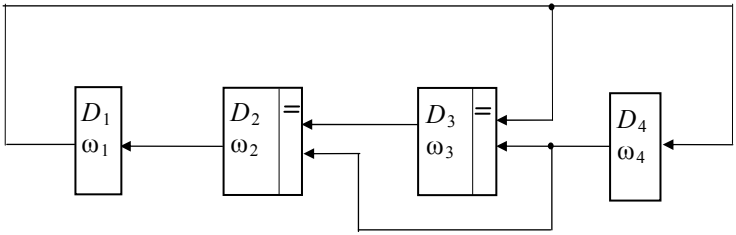


Рис. 10.11. Структура фильтра

Тогда матрицу M можно записать в виде

$$M(x) = \left|\mu_1(x)\, \mu_2(x) \ldots \mu_r(x)\right|.$$

Полиномы $\mu_j(x)$ выберем таким образом, чтобы

$$\left.\begin{aligned} &\mu_1(x) = g_1 + g_2 x + \ldots + g_r x^{r-1} = x^{-1}, \\ &\mu_2(x) = x\mu_1(x) = 1, \\ &\mu_3(x) = x^2 \mu_1(x) = x, \\ &\qquad \ldots \\ &\mu_r(x) = x^{r-1}\mu_1(x) = x^{r-2} \end{aligned}\right\} \bmod\ g(x),$$

где $g(x) = 1 + g_1 x + g_2 x^2 + \ldots + g_r x^r -$ порождающий многочлен, для которого $g_1 x + g_2 x^2 + \ldots + g_r x^r = 1 \bmod g(x)$.

Тогда схема, описываемая матрицей связей

$$M(x) = \left| x^{-1} \;\; 1 \;\; x \;\; ... \;\; x^{r-2} \right|$$

или матрицей коэффициентов

$$M = \begin{vmatrix} g_1 & 1 & 0 & ... & 0 \\ g_2 & 0 & 1 & ... & 0 \\ & & ... & & \\ g_{r-1} & 0 & 0 & ... & 1 \\ g_r & 0 & 0 & ... & 0 \end{vmatrix},$$

позволяет выполнять деление на полином $g(x)$ [13].

Действительно, если исходному состоянию фильтра $\Omega^{(0)} = (\omega_1, \omega_2, ..., \omega_r)$ поставить в соответствие полином

$$\Omega^{(0)}(x) = \omega_1 + \omega_2 x + ... + \omega_r x^{r-1},$$

то состояние фильтра в следующем такте (на вход схемы информация не поступает) равно

$$\Omega^{(1)}(x) = \Omega^{(0)} M_t(x) = \omega_1 x^{-1} + \omega_2 + \omega_3 x + ... + \omega_r x^{r-2} =$$

$$= (\omega_1 + \omega_2 x + ... + \omega_r x^{r-1}) x^{-1} = \Omega^{(0)}(x) x^{-1} \bmod g(x),$$

где $M_t(x)$ – транспонированная матрица связей.

Входную цепь фильтра построим таким образом, чтобы

$$\Omega^{(i)}(x) = [\Omega^{(i-1)}(x) + a_{i-1}] x^{-1} \bmod g(x),$$

т.е. последующее состояние вычисляется прибавлением входного символа к содержимому фильтра и однократным сдвигом его содержимого.

Если начальное состояние $\Omega^{(0)}(x) = 0$ и на вход фильтра поступает последовательность символов $a_0, a_1, ..., a_{n-1}$, а эта последовательность соответствует коэффициентам произвольного полинома степени не более $n-1$

$$A_n(x) = a_0 + a_1 x + ... + a_{n-1} x^{n-1},$$

то фильтр будет последовательно принимать состояния

$$\left.\begin{aligned}
&\Omega^{(1)}(x) = a_0 x^{-1}, \\
&\Omega^{(2)}(x) = (a_0 x^{-1} + a_1)x^{-1} = a_0 x^{-2} + a_1 x^{-1}, \\
&\qquad \ldots \\
&\Omega^{(n)}(x) = a_0 x^{-n} + a_1 x^{-(n-1)} + \ldots + a_{n-1} x^{-1}
\end{aligned}\right\} \bmod g(x).$$

Следовательно, $\Omega^{(n)}(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \bmod g(x)$, так как $x^n = 1 \bmod g(x)$.

После поступления последнего символа a_{n-1} содержимое фильтра будет равно остатку от деления полинома $A_n(x)$ на $g(x)$.

Структура рассмотренного фильтра, описываемого матрицей связи $M(x) = \left| x^{-1} \;\; 1 \;\; x \;\; \ldots \;\; x^{r-2} \right|$, приведена на рис. 10.12.

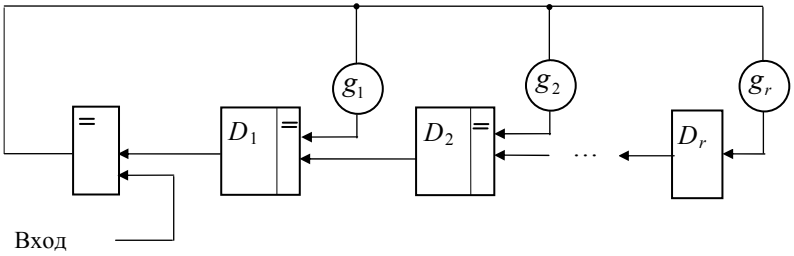


Рис. 10.12. Структура фильтра с входом

При соответствующей организации входной цепи построенная схема может быть использована для вычисления контрольных разрядов разделимого циклического кода (рис. 10.13).

Кодер работает следующим образом. Сначала ключ находится в 1-м положении, и k информационных разрядов $a_0, a_1, \ldots, a_{k-1}$ поступают на вход фильтра, который вначале находится в нулевом состоянии, и непосредственно в линию связи. Затем ключ переключается во 2-е положение. В этом случае обратная связь в фильтре размыкается и $r = n - k$ контрольных разрядов последо-

вательно выдвигаются в линию. Таким образом, к концу передачи кодового слова длиной n фильтр оказывается в нулевом состоянии.

Формально работа кодера описывается следующим образом. После поступления k информационных символов содержимое фильтра равно

$$\Omega^{(k)}(x) \equiv a_0 x^{-k} + a_1 x^{-k+1} + ... + a_{k-1} x^{-1} =$$

$$= x^{-k}(a_0 + a_1 x + ... + a_{k-1} x^{k-1}) = x^{-k} A_k(x) = x^{n-k} A_k(x) \bmod g(x),$$

которое совпадает с контрольным вектором.

На рис. 10.14 представлена блок-схема декодирующего устройства.

Коэффициенты $b_0, b_1, ..., b_{n-1}$ полинома $B_n(x)$ последовательно поступают на вход буферного ЗУ и одновременно на вход многотактного фильтра, с помощью которого $B_n(x)$ делится на $g(x)$.

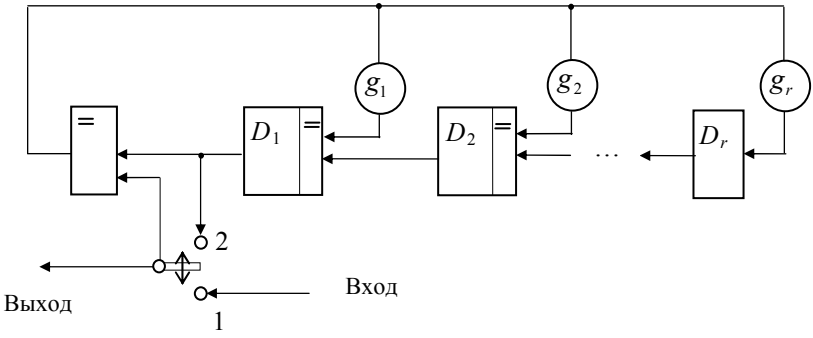


Рис. 10.13. Кодирующее устройство

Если ошибка отсутствует, то после поступления последнего символа b_{n-1} содержимое фильтра будет равно нулю.

Получение ненулевого результата свидетельствует о наличии обнаруживаемой ошибки, полином которой имеет вид $E(x) = x^i e(x)$.

Если продолжить работу декодирующего устройства (синхронно производить сдвиги БЗУ и блока деления в автономном режи-

ме), то последовательные состояния фильтра будут равны

$$\left.\begin{array}{l}\Omega^{(0)}(x)=x^{i}e(x),\\ \Omega^{(1)}(x)=x^{i-1}e(x),\\ \ldots\\ \Omega^{(i)}(x)=x^{0}e(x)=e(x)\end{array}\right\}\bmod g(x),$$

т.е. вычет вида ошибки будет воспроизведен точно через i тактов. В простейшем случае, когда возникла одиночная ошибка $E(x)=x^{i}$, т.е. $e(x)=1$, то на i-м такте $\Omega^{(i)}(x)=1$ или $\Omega^{(i)}=\underset{\omega_r\ldots\omega_2\omega_1}{00\ldots01}$. При появлении такой комбинации на выходе ***логической схемы*** должен появиться сигнал, который инвертирует ошибочный символ.

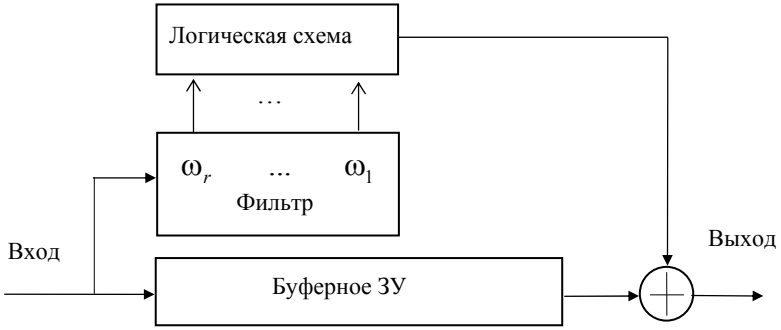


Рис.10.14. Структура декодирующего устройства

10.8. ПОСТРОЕНИЕ МНОГОКАНАЛЬНЫХ КОДИРУЮЩИХ И ДЕКОДИРУЮЩИХ УСТРОЙСТВ ЦИКЛИЧЕСКОГО КОДА ХЭММИНГА

Рассмотренные одноканальные кодеры применяются в системах последовательного действия, где передача символов осуществляется последовательно во времени. В ряде устройств современных управляющих систем используется параллельно-последовательный принцип передачи информации. При параллельно-последовательной передаче информационное слово разбивается на части длиной ν разрядов, эти ν-разрядные части передаются

последовательно. В этом случае необходимо уметь строить линейные фильтры (кодеры), содержащие ν входов и ν выходов.

Последовательные состояния и значения выхода рассмотренных одноканальных линейных фильтров описываются матричными уравнениями

$$\Omega^{(i)} = \Omega^{(i-1)} \cdot M_t + a_{i-1} \cdot F,$$
$$y_i = \Omega^{(i-1)} \cdot B + a_{i-1} \cdot \Psi,$$

где $\Omega^{(i)} = (\omega_1^{(i)},..., \omega_r^{(i)})$ – вектор состояния фильтра в момент времени i; M_t – транспонированная матрица связей; $a_{i-1}(y_i)$ – значение входного (выходного) символа в момент времени $i-1$ (i); F – $(1\times r)$ матрица связей между входом и элементами задержки фильтра; B – $(r\times 1)$ матрица связей между элементами задержки и выходом фильтра; Ψ – (1×1) матрица связей между входом и выходом фильтра.

Если известны M_t, F, B и Ψ, то соответствующие уравнения для ν-канального аналога фильтра, содержащего ν входов и ν выходов, имеют следующий вид (известный результат из ***математического аппарата линейных последовательностных машин***):

$$\Omega^{(i)} = \Omega^{(i-1)} \cdot M_t^{\nu} + A^{(i-1)} \cdot F^*,$$
$$Y^{(i)} = \Omega^{(i-1)} \cdot B^* + A^{(i-1)} \cdot \Psi^* \quad ,$$

где $A^{(i-1)} = (a_1^{(i-1)},..., a_\nu^{(i-1)})$ – входной вектор в момент времени $i-1$; $Y^{(i)} = (y_1^{(i)},..., y_\nu^{(i)})$ – выходной вектор в момент времени i;

$F^* = \begin{vmatrix} F \cdot M_t^{\nu-1} \\ F \cdot M_t^{\nu-2} \\ \dots \\ F \end{vmatrix}$ – $(\nu \times r)$ матрица связей между входами фильтра и элементами задержки; $B^* = \left| B \quad M_t \cdot B \quad \dots \quad M_t^{\nu-1} \cdot B \right|$ – $(r \times \nu)$ матрица связей между элементами задержки и ν выходами фильт-

ра; $\Psi^* = \begin{vmatrix} \Psi & F\cdot B & F\cdot M_t\cdot B & \dots & F\cdot M_t^{\nu-2}\cdot B \\ 0 & \Psi & F\cdot B & \dots & F\cdot M_t^{\nu-3}\cdot B \\ & & & \dots & \\ 0 & 0 & 0 & \dots & \Psi \end{vmatrix}$ – ($\nu \times \nu$) матрица связей между входами фильтра и его выходами.

В качестве примера рассмотрим построение двухканального кодера ($\nu = 2$) для разделимого циклического кода (7,4), порождаемого многочленом $g(x) = 1 + x^2 + x^3$.

Матричное описание одноканального кодера задается матрицами

$$M = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}, \quad M_t = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix},$$

$$F = \begin{vmatrix} g_1 & g_2 & \dots & g_r \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \end{vmatrix}, \quad B = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}, \quad \Psi = \begin{vmatrix} 1 \end{vmatrix}.$$

Структура одноканального кодера представлена на рис. 10.15.

При известных схемотехнических решениях матрицы $B(B^*)$ и $\Psi(\Psi^*)$ в принципе можно и не строить. В большей степени они приводятся лишь для того, чтобы сохранить общую полноту картины построения ν-канального аналога одноканальной линейной последовательностной машины.

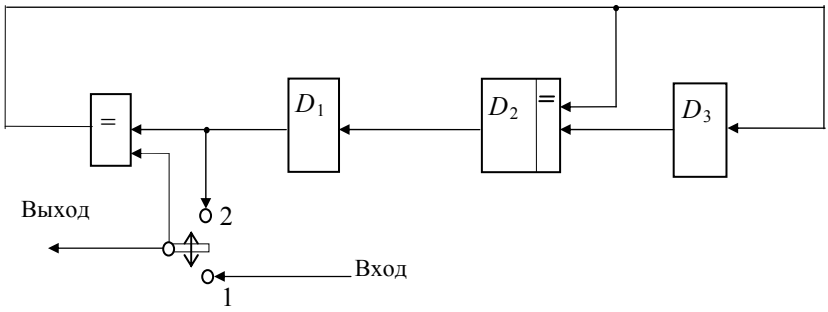


Рис. 10.15. Структура одноканального кодера

Матричное описание двухканального аналога имеет вид

$$M_t^2 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix}, \quad F^* = \begin{vmatrix} F \cdot M_t \\ F \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix},$$

$$M^2 = \begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix}, \quad F \cdot M_t = \begin{vmatrix} 0 & 1 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \end{vmatrix}.$$

Структура двухканального кодера представлена на рис. 10.16.

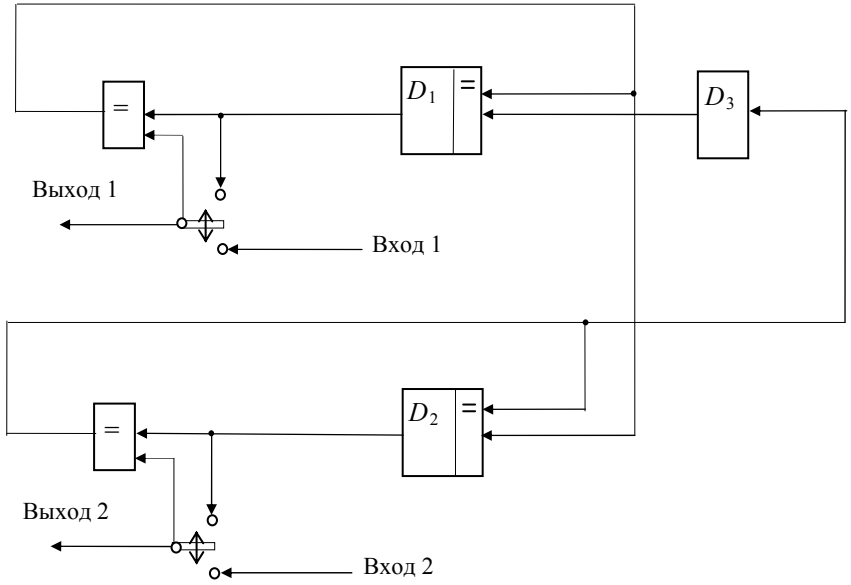


Рис. 10.16. Структура двухканального кодера

В рассмотренном примере количество информационных разрядов k кратно ν. В общем случае это условие не выполняется и после выполнения $z = \left\lceil \frac{k}{\nu} \right\rceil$ тактов (квадратные скобки означают наименьшее большее целое число), содержимое ν-канального

кодера будет равно $R^*(x) = R(x) \cdot x^{-(z \cdot \nu - k)} \bmod g(x)$, где $R(x)$ соответствует контрольному коду. Поэтому, если $(z \cdot \nu - k) \neq 0$, требуется коррекция содержимого фильтра. Однако коррекции можно избежать, если полином $A_k(x)$, соответствующий информационным разрядам, умножить на $x^{(z \cdot \nu - k)}$. В этом случае $R^*(x) = R(x) \cdot x^{-(z \cdot \nu - k)} \cdot x^{(z \cdot \nu - k)} = R(x) \bmod g(x)$. Известно, что умножение равносильно операции сдвига. Реализация сдвига заключается в том, что перед младшим информационным разрядом a_0 приписывается $(z \cdot \nu - k)$ фиктивных нулей. При этом разряд a_0 поступает на $(z \cdot \nu - k) + 1$ вход (канал) кодера, разряд a_1 – на $(z \cdot \nu - k) + 2$ канал и т.д.

Блок-схема ν канального декодирующего устройства представлена на рис. 10.17.

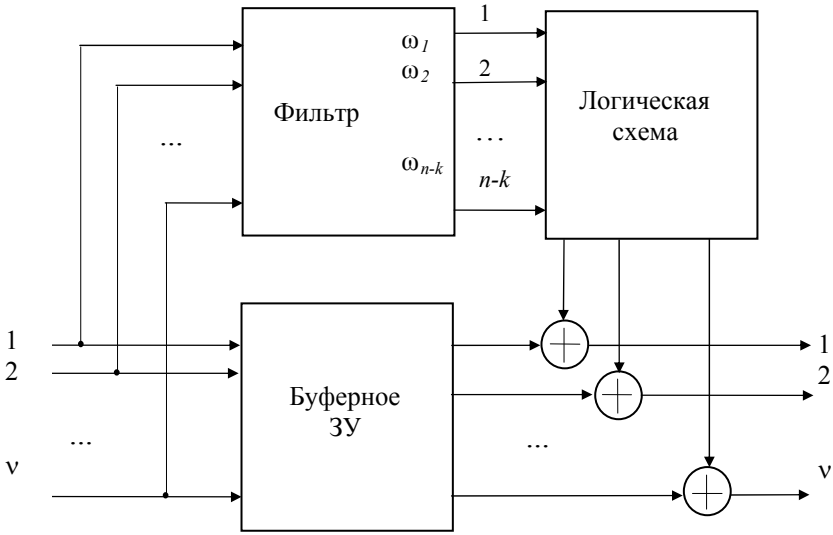


Рис. 10.17. Структура декодирующего устройства

Принцип построения логической схемы заключается в следующем. Для приема всего слова длиной n разрядов требуется

$z' = \left[\frac{n}{\nu}\right]$ тактов. Поэтому при наличии ошибки вида x^i состояние фильтра после приема кодового слова будет равно

$$\Omega^0(x) = x^{i-(z'\cdot\nu-n)} \bmod g(x) ,$$

а последующие состояния при автономной работе будут равны:

$$\Omega^1(x) = x^{i-(z'\cdot\nu-n)-\nu} \bmod g(x),$$
$$\Omega^2(x) = x^{i-(z'\cdot\nu-n)-2\cdot\nu} \bmod g(x),$$
$$\Omega^3(x) = x^{i-(z'\cdot\nu-n)-3\cdot\nu} \bmod g(x),$$
$$\ldots .$$

Пусть $i = j\cdot\nu+\beta \quad (\beta<\nu)$, т.е. ошибка должна быть исправлена в такте j. Тогда на j-м такте автономной работы состояние фильтра равно

$$\Omega^j(x) = x^{j\cdot\nu+\beta-(z'\cdot\nu-n)-j\cdot\nu} = x^{\beta-(z'\cdot\nu-n)} \bmod g(x) ,$$

т.е. логическая схема должна распознавать следующие вычеты:

$$\left.\begin{array}{ll} \text{при } \beta=0 & x^{-(z'\cdot\nu-n)}, \\ \text{при } \beta=1 & x^{1-(z'\cdot\nu-n)}, \\ \quad\ldots & \\ \text{при } \beta=\nu-1 & x^{\nu-1-(z'\cdot\nu-n)} \end{array}\right\} \bmod g(x).$$

Таким образом, логическая схема в ν-канальном декодере содержит ν выходов.

Например, пусть $n=7$, $\nu=2$, $g(x)=1+x^2+x^3$, тогда $z'=4$ и логическая схема будет представлять собой две схемы совпадения на следующие коды:

при $\beta=0$ $\quad x^{-1} \equiv x^{7-1} \equiv x^6 \equiv x^2+x$, т.е. 110;
при $\beta=1$ $\quad x^{1-1} \equiv x^0 \equiv 1$, т.е. 001.

ЗАДАЧИ

10.1. Используется групповой код Хэмминга (7,4), исправляющий одиночные ошибки. На приемном конце принята комбинация 1011011. Какой код был передан?

10.2. Используется разделимый циклический код (7,4) с $g(x) = x^3 + x + 1$. На приемном конце принята комбинация 1001:011. Какой код был передан?

10.3. Используется неразделимый циклический код (7,4) с $g(x) = x^3 + x^2 + 1$. На приемном конце принята комбинация 0000011. Какой код был передан?

10.4. По каналу передаются три команды. Построить групповой код, исправляющий двукратные ошибки включительно.

10.5. Построить разделимый циклический код Хэмминга (7,4) с порождающим многочленом $g(x) = x^3 + x^2 + 1$.

10.6. Построить неразделимый циклический код Хэмминга (7,4) с порождающим многочленом $g(x) = x^3 + x + 1$.

10.7. По каналу передается одна команда. Построить циклический код, исправляющий двукратные ошибки включительно.

10.8. Построить кодирующее и декодирующее устройства для группового кода Хэмминга (3,1) .

10.9. Построить одноканальное кодирующее и декодирующее устройства для циклического кода Хэмминга (3,1) с $g(x) = x^2 + x + 1$.

10.10. Построить циклический код Боуза–Чоудхури–Хоквингема (БЧХ), имеющий длину $N = 21$ и служащий для исправления ошибок кратности $T = 2$ включительно. Продемонстрировать процесс исправления ошибки.

10.11. Построить циклический код БЧХ, имеющий длину $N = 31$ и служащий для исправления ошибок кратности $T = 15$ включительно.

10.12. Построить циклический код Файра, позволяющий обнаружить пакет ошибок длиной $b \le 5$ и исправить одиночную вспышку (пакет) ошибок длиной $l \le 5$.

10.13. Определить избыточность кода Файра (279,265) и сопоставить ее с избыточностью циклического кода БЧХ с близким

значением $n = 255$.

10.14. Найти параметры циклических кодов Файра, полагая $l = m$ и $6 \le l \le 10$.

10.15. Построить четырехканальное кодирующее устройство $(\nu = 4)$ и декодирующее устройство для разделимого циклического кода Хэмминга (31,26), порождаемого многочленом $g(x) = x^5 + x^3 + 1$.

11. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО РЕШЕНИЮ ТИПОВЫХ ЗАДАЧ И ОТВЕТЫ

1.1. Для доказательства достаточно воспользоваться известными свойствами:

$$\cos(k\omega_0 t) = \frac{e^{jk\omega_0 t} + e^{-jk\omega_0 t}}{2},$$

$$\sin(k\omega_0 t) = \frac{e^{jk\omega_0 t} - e^{-jk\omega_0 t}}{2j}.$$

Тогда

$$\frac{1}{2}\left[\left|A_k\right| e^{j(k\omega_0 t + \Theta_k)} + \left|A_{-k}\right| e^{-j(k\omega_0 t + \Theta_{-k})}\right] = \left|A_k\right| \cos(k\omega_0 t + \Theta_k) \ ,$$

с учетом того, что $\left|A_k\right| = \left|A_{-k}\right|$ и $\Theta_{-k} = -\Theta_k$.

1.2. Тригонометрическая форма ряда Фурье имеет вид

$$x(t) = \frac{A_0}{2} + \sum_{k=1}^{\infty} \left|A_k\right| \cos(\ k\omega_0 t + \Theta_k\ ) =$$

$$= \frac{A_0}{2} + \sum_{k=1}^{\infty} (a_k \cos\ k\omega_0 t - b_k \sin\ k\omega_0 t),$$

где $\frac{A_0}{2}$ – постоянная составляющая функции $x(t)$; $\left|A_k\right|, k\omega_0, \Theta_k$ – амплитуда, частота и начальная фаза *k*-й гармонической составляющей; $\left|A_k\right| \cos(\ k\omega_0 t + \Theta_k\ )$ – *k*-я гармоническая составляющая; $\omega_0 = \frac{2\pi}{T}$ – частота основной гармоники (*T* – период колебаний).

Ряд Фурье с учетом свойств периодической функции $x(t)$ приобретает еще более простой вид:

$$x(t) = \sum_{k=1}^{\infty} b_k \sin k\omega_0 t$$ – нечетная функция.

Коэффициент b_k вычисляется в соответствии с известным вы-

ражением

$$b_k = |A_k| \sin \Theta_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin(k\omega_0 t) dt \ .$$

Итак, поскольку функция $x(t)$ – нечетная, разложение будем искать в виде ($a_k = 0$)

$$b_k = \frac{4h}{T} \int_0^{T/2} \sin(k\omega_0 t) dt = -\frac{4h}{T} \frac{1}{k\omega_0} \cos k\omega_0 t \Big|_0^{T/2} =$$

$$= -\frac{2h}{k\pi}(\cos k\pi - 1) = \begin{cases} \frac{4h}{k\pi}, k = 1,3,5,...; \\ 0, k = 2,4,6,... \ . \end{cases}$$

Спектр фаз приобретает вид

$$\Theta_k = -\operatorname{arctg}\left(\frac{b_k}{a_k}\right) = -\operatorname{arctg}\left(\frac{b_k}{0}\right) = -\frac{\pi}{2} \ .$$

Спектральные характеристики меандра приведены на рис. 11.1. Окончательно тригонометрическая форма обобщенного ряда Фурье будет иметь вид

$$x(t) = \frac{4h}{\pi}(\sin \omega_0 t + \frac{1}{3}\sin 3\omega_0 t + \frac{1}{5}\sin 5\omega_0 t + ...) .$$

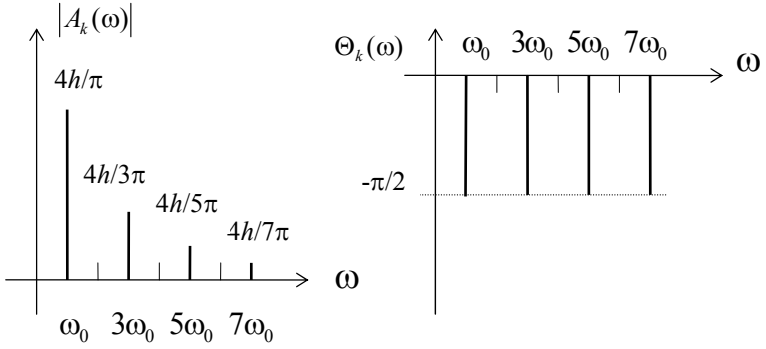


Рис. 11.1. Спектр амплитуд и спектр фаз меандра

1.3. В общем случае разложение в ряд Фурье периодической последовательности прямоугольных импульсов имеет вид

$$x(t)=\frac{\tau h}{T}\left[1+2\sum_{k=1}^{\infty}\frac{\sin\dfrac{k\omega_0\tau}{2}}{\dfrac{k\omega_0\tau}{2}}\cos(k\omega_0 t)\right].$$

Следовательно, спектральные характеристики можно представить следующим образом

$$\frac{A_0}{2}=\frac{h}{2},\quad |A_k(\omega)|=\frac{2h}{k\pi},\quad \Theta_k(\omega)=-\frac{\pi}{2}\left[1+(-1)^{\frac{k+1}{2}}\right],\quad k=1,3,5,\ldots\ .$$

1.4. Обобщенный ряд Фурье имеет вид

$$x(t)=\frac{h}{2}-\frac{4h}{\pi^2}\left[\cos(\omega_0 t)+\frac{1}{3^2}\cos(3\omega_0 t)+\frac{1}{5^2}\cos(5\omega_0 t)+\ldots\right].$$

1.5. Обобщенный ряд Фурье имеет вид

$$x(t)=\frac{2h}{\pi}\left[\sin(\omega_0 t)-\frac{1}{2}\sin(2\omega_0 t)+\frac{1}{3}\sin(3\omega_0 t)-\ldots\right].$$

1.6. Спектральные характеристики представлены на рис. 11.2.

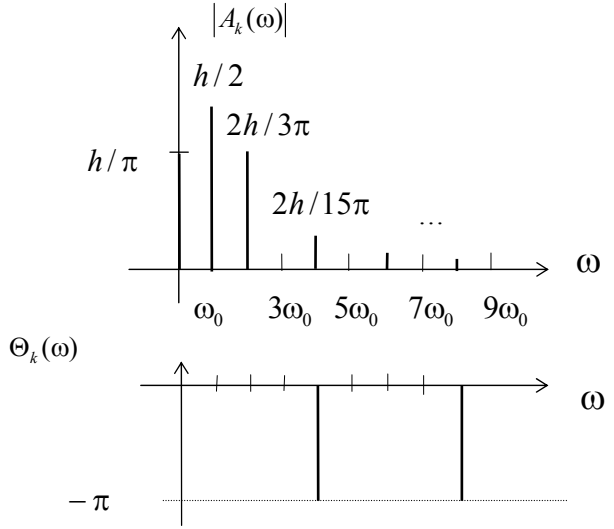


Рис. 11.2. Спектр амплитуд и спектр фаз

1.7. Обобщенный ряд Фурье имеет вид

$$x(t) = 2\frac{h}{\pi} + \frac{4h}{3\pi}\cos(2\omega_0 t) - \frac{4h}{15\pi}\cos(4\omega_0 t) + \frac{4h}{35\pi}\cos(6\omega_0 t) - \ldots .$$

1.8. Полная средняя мощность заданной периодической последовательности прямоугольных импульсов равна

$$P_\infty = \frac{A_0^2}{4} + \frac{1}{2}\sum_{k=1}^{\infty}\left|A_k^2\right| = h^2\left[\frac{1}{4} + \frac{2}{\pi^2}\left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \ldots\right)\right] =$$

$$= h^2\left[\frac{1}{4} + \frac{2}{\pi^2}\frac{\pi^2}{8}\right] = 0{,}5h^2 \ .$$

Результат получен с учетом того обстоятельства, что

$$\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2} = \sum_{n=1}^{\infty}\frac{1}{n^2} - \sum_{n=1}^{\infty}\frac{1}{(2n)^2} = \frac{\pi^2}{6} - \frac{\pi^2}{24} = \frac{\pi^2}{8} \ .$$

Полную среднюю мощность, кроме того, можно было получить более простым способом:

$$P_\infty = \frac{1}{T}\int_{-\tau/2}^{\tau/2} h^2 dt = 0{,}5h^2 \ .$$

В пределах практической ширины спектра размещаются постоянная составляющая и первая, третья, пятая гармоники. Следовательно, средняя мощность сигнала, приходящаяся на практическую ширину спектра, имеет вид

$$P_5 = \frac{h^2}{4} + \frac{1}{2}\left(\frac{2h}{\pi}\right)^2 + \frac{1}{2}\left(\frac{2h}{3\pi}\right)^2 + \frac{1}{2}\left(\frac{2h}{5\pi}\right)^2 \approx 0{,}48h^2 \ .$$

Таким образом, в пределах практической ширины спектра сосредоточено

$$\frac{P_5}{P_\infty} = \frac{0{,}48h^2}{0{,}5h^2} = 0{,}96 \,,$$

т.е. 96 % всей средней мощности сигнала.

1.9. Число гармоник, подлежащих учету, может быть получено из соотношения

$$\frac{\left|A_k\right|}{\left|A_1\right|} = 0{,}1 \ .$$

1.10. На рис. 11.3 приведены первые три гармонические составляющие меандра, а на рис. 11.4 – их графическая сумма.

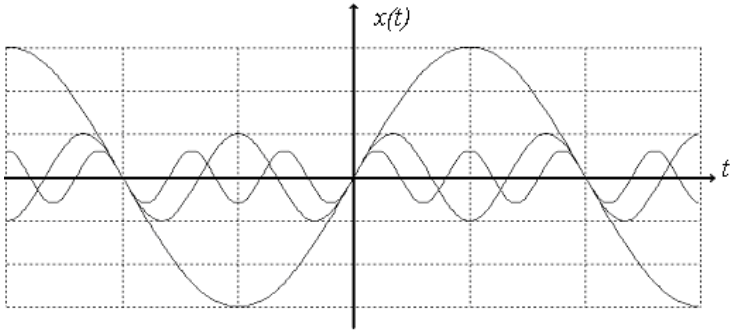


Рис. 11.3. Гармонические составляющие меандра

1.11. Наиболее простой способ решения этой задачи состоит в следующем. Получить спектральные характеристики для сигнала, представленного на рис. 1.4. Сигнал в соответствии с условием задачи отличается от периодического колебания пилообразной формы наличием постоянной составляющей h и сдвигом в сторону запаздывания на величину $T/2$. Следовательно, достаточно изменить известный спектр в соответствии с известными свойствами преобразования Фурье. Особенность будет заключаться только в том, что непрерывная частота, фигурирующая в преобразованиях Фурье, должна быть заменена на дискретную $\omega \to k\omega_0$.

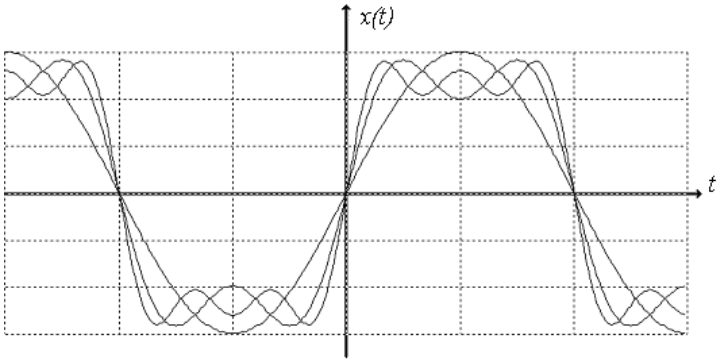


Рис. 11.4. Сумма гармонических составляющих меандра

2.1. Спектральная плотность такого сигнала находится следующим образом:

$$S(j\omega)=\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} h\cos(\omega_0 t)e^{-j\omega t}dt=\frac{h}{2}\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} e^{-j(\omega-\omega_0)t}dt+\frac{h}{2}\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} e^{-j(\omega+\omega_0)t}dt=$$

$$=\frac{h}{2}\left[\frac{e^{-j\frac{(\omega-\omega_0)\tau}{2}}-e^{j\frac{(\omega-\omega_0)\tau}{2}}}{-j(\omega-\omega_0)}+\frac{e^{-j\frac{(\omega+\omega_0)\tau}{2}}-e^{j\frac{(\omega+\omega_0)\tau}{2}}}{-j(\omega+\omega_0)}\right]=$$

$$=\frac{h}{\omega-\omega_0}\sin\frac{(\omega-\omega_0)\tau}{2}+\frac{h}{\omega+\omega_0}\sin\frac{(\omega+\omega_0)\tau}{2}=$$

$$=\frac{\tau h}{2}\frac{\sin\frac{(\omega-\omega_0)\tau}{2}}{\frac{(\omega-\omega_0)\tau}{2}}+\frac{\tau h}{2}\frac{\sin\frac{(\omega+\omega_0)\tau}{2}}{\frac{(\omega+\omega_0)\tau}{2}}=$$

$$=\frac{1}{2}\left\{S_1\left[j(\omega-\omega_0)\right]+S_1\left[j(\omega+\omega_0)\right]\right\}\ .$$

Сравнив полученное выражение с выражением

$$S_1(j\omega)=h\tau\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}$$

спектральной плотности для одиночного прямоугольного импульса такой же длительности и амплитуды h, но без высокочастотного заполнения, видим, что по отношению к спектру прямоугольного импульса спектр импульса высокочастотных колебаний смещен на величину несущей ω_0 и расширен в два раза за счет появления зеркального отображения спектра (рис. 11.5).

Следует отметить, что аналогичный результат может быть получен путем простого использования свойства преобразования Фурье – сдвиг спектра колебания по частоте (2.3.4), в соответствии с которым расщепление спектра $S_1(j\omega)$ на две части, смещенные

соответственно на $+\omega_0$ и $-\omega_0$, эквивалентно умножению функции $x_1(t)$ на гармоническое колебание $\cos\omega_0 t$.

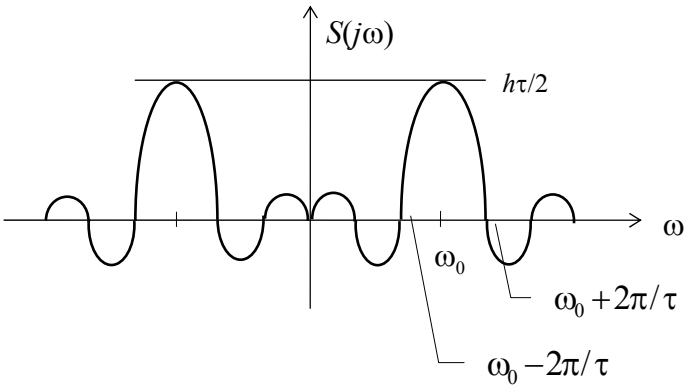


Рис. 11.5. Комплексная форма спектральной плотности одиночного импульса высокочастотных колебаний

Графики модуля спектральной плотности импульса высокочастотных колебаний $S(\omega) = h\tau\left|\dfrac{\sin\dfrac{(\omega-\omega_0)\tau}{2}}{\dfrac{(\omega-\omega_0)\tau}{2}}\right|$ и аргумента спектральной плотности $\Theta(\omega) = -E\left(\dfrac{\omega-\omega_0}{2\pi/\tau}\right)\pi$ приведены на рис. 11.6.

2.2. $S(j\omega) = \dfrac{h}{\sqrt{\beta^2+\omega^2}}e^{-j(\operatorname{arctg}\frac{\omega}{\beta}+\omega t_0)}$.

2.3. Единичная функция $1(t)$ не удовлетворяет условию абсолютной интегрируемости $\int\limits_{-\infty}^{\infty}|x(t)|dt < \infty$ и, следовательно, к ней нельзя применить преобразование Фурье. Однако ее можно рассматривать как образованную из импульса экспоненциальной

формы при неограниченном уменьшении его коэффициента затухания $\beta \to 0$.

Окончательно результат имеет вид $S(\omega) = \dfrac{h}{\omega}$, $\Theta(\omega) = -\dfrac{\pi}{2}$.

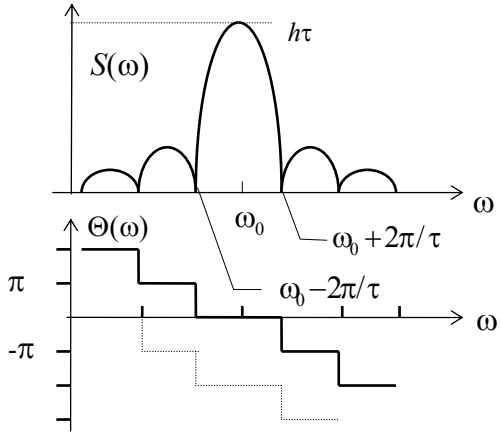


Рис. 11.6. Модуль и аргумент спектральной плотности одиночного импульса высокочастотных колебаний

2.4. Реакция фильтра представлена на рис. 11.7.

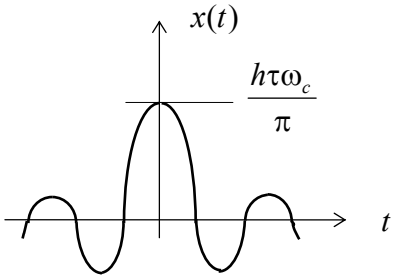


Рис. 11.7. Реакция фильтра нижних частот на дельта-импульс

2.5. Доказательство состоит в нахождении спектральной плотности дельта-функции $S(j\omega) = e^{-j\omega t_0}$ и в последующем применении обратного преобразования Фурье.

2.6. Спектральная плотность выходного сигнала определяется следующим образом:

$$S^*(j\omega) = S(j\omega)K(j\omega),$$

где $K(j\omega)$ – передаточная функция интегрирующей цепочки. Выходной сигнал находится в соответствии с обратным преобразованием Фурье

$$x^*(t) = \begin{cases} \dfrac{1}{2RC} e^{\frac{t-T_0}{RC}}, & t \le T_0; \\ \dfrac{1}{2RC} e^{-\frac{t-T_0}{RC}}, & t > T_0, \end{cases}$$

вид которого представлен на рис. 11.8.

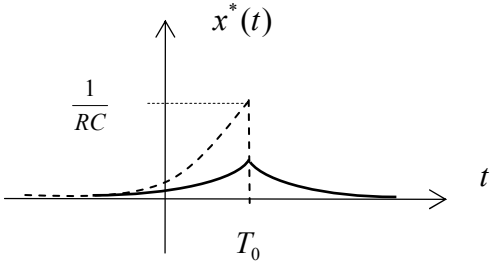


Рис. 11.8. Выходной сигнал

Данное решение находится с использованием известного табличного интеграла $\int_0^\infty \dfrac{\cos px}{q^2 + x^2} dx = \dfrac{\pi}{2q} e^{-|pq|}$.

2.7. Спектральную плотность находим в соответствии со свойством преобразования Фурье о взаимозаменяемости переменных ω и t для четного сигнала (рис. 11.9).

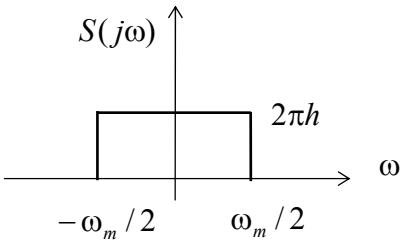


Рис. 11.9. Спектральная плотность импульса

3.1. $K_{XX}^{\text{дп}} = \dfrac{A^2}{2}\cos\omega_0\tau$.

3.2. $K_{XX}^{\text{дп}} = \dfrac{1}{4}\left|A_0\right|^2 + \dfrac{1}{2}\sum_{k=1}^{\infty}\left|A_k\right|^2\cos k\omega_0\tau$. Как видно из полученного выражения, корреляционная функция периодического сигнала с периодом T представляет собой периодическую функцию аргумента τ с тем же периодом. Амплитуда гармоники корреляционной функции равна половине квадрата амплитуды соответствующей гармоники $x(t)$.

3.3. Задача решается с использованием преобразования Хинчина–Винера

$$G_{XX}(\omega) = \int_{-\infty}^{\infty} K_{XX}(\tau)e^{-j\omega\tau}d\tau;$$

$$K_{XX}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_{XX}(\omega)e^{j\omega\tau}d\omega.$$

Спектральная плотность средней мощности представляет собой усредненную энергетическую картину распределения мощности сигнала по частотному спектру (рис. 11.10)

$$G_{XX}(\omega) = \frac{2A\alpha}{\alpha^2 + \omega^2} \quad .$$

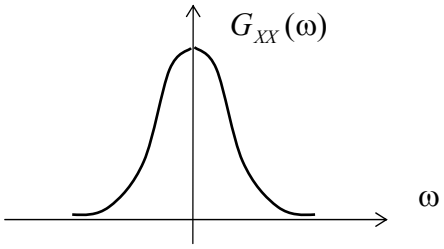


Рис. 11.10. Спектральная плотность мощности

3.4. $G_{XX}(\omega) = G_0$. Спектр равномерен на всех частотах – сигнал типа ”белый шум”, который характеризуется тем, что значения

$x(t)$ в любые два (даже сколь угодно близкие) момента времени некоррелированы.

3.5. $K_{XX}(\tau) = e^{-3|\tau|} + e^{-2|\tau|}$. Задача решается достаточно просто, если $G_{XX}(\omega)$ представить в виде

$$G_{XX}(\omega) = \frac{6}{\omega^2 + 9} + \frac{4}{\omega^2 + 4} \ .$$

3.6. Дисперсия стационарного случайного процесса с нулевым математическим ожиданием может быть вычислена по формуле

$$D[X(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{XX}(\omega) d\omega = 2 \ .$$

3.7. Как известно из теории вероятности, среднеквадратическое значение определяется выражением

$$\sqrt{m[X^2]} = \sqrt{K_{XX}(0)} = \sqrt{\frac{G_0}{2\pi}} \ .$$

3.8. Для заданного стационарного случайного процесса математическое ожидание и дисперсия соответственно равны

$$m[X] = 0 \ , \quad D[X] = \frac{\omega_c G_0}{\pi} \ .$$

3.9. Спектральная плотность мощности выходного сигнала определяется следующим образом:

$$G_{XX}(\omega) = G_0 \left| K(j\omega) \right|^2 = \frac{G_0 \alpha^2}{1 + (\omega\beta)^2} \ ,$$

где $K(j\omega)$ – передаточная функция электрической цепи. Корреляционная функция вычисляется обычным образом:

$$K_{XX}(\tau) = \frac{G_0 \alpha^2}{2\beta} e^{-\frac{|\tau|}{\beta}} \ .$$

4.1. Интегральная кривая распределения энергии сигнала в спектре частот $\lambda(\omega_c) = \frac{Э(\omega_c)}{Э} = \frac{\frac{1}{\pi}\int_0^{\omega_c} [S(\omega)]^2 d\omega}{\frac{1}{\pi}\int_0^{\infty} [S(\omega)]^2 d\omega}$ для экспоненци-

ального импульса $\lambda(\omega_c) = \frac{2}{\pi}\text{arctg}\,\frac{\omega_c}{\beta}$ представлена на рис. 11.11. Задавшись, например, уровнем 0,9, можно приближенно удовлетворить условие теоремы Котельникова об ограниченности спектра сигнала и найти частоту, за пределами которой $S(j\omega) \approx 0$, $\omega_c^{0,9} = \beta\ \text{tg}(0{,}45\pi)$.

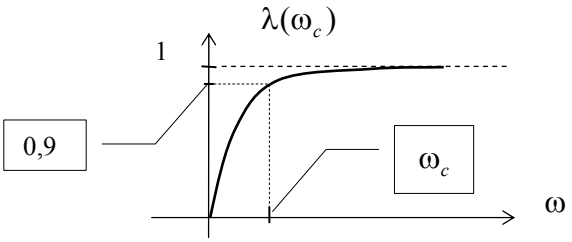


Рис. 11.11. Интегральная кривая распределения энергии

4.2. Интервал дискретизации находится как отношение

$$\Delta t = \frac{\pi}{\omega_c^{0,9}} = \frac{\pi}{\beta\ \text{tg}(0{,}45\pi)}\ .$$

4.3. Известно, что ряд Котельникова представляет собой разложение реализации случайного процесса координатными детерминированными функциями времени $A(t_k) = \text{sinc}\,[\omega_c(t - k\Delta t)]$ с весовыми коэффициентами $x(k\Delta t)$, равными мгновенным значениям сигнала в точках $k\Delta t$.

На рис. 11.12 изображен ряд Котельникова совместно с детерминированными координатными функциями времени для экспоненциального импульса $x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t)\,\frac{\sin[\omega_c(t - k\Delta t)]}{\omega_c(t - k\Delta t)}$.

4.4. Период спектра амплитуд (рис. 4.6) дискретного периодического сигнала в масштабе частот равен $N\Omega = \frac{2\pi}{T}$, а в пределах периода составляющие спектра симметричны относительно середины этого интервала, причем $|A_k| = |A_{N-k}|$. Поэтому вся полезная информация имеется уже в интервале частот, равном $\frac{\pi}{T}$.

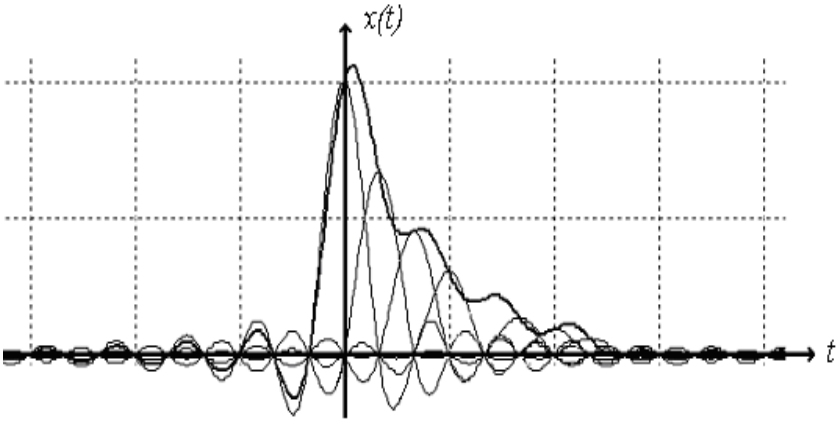


Рис. 11.12. Ряд Котельникова

5.1. Отношение мощностей полезного сигнала и помехи увеличивается в n раз в результате n - кратного отсчета

$$\left(\frac{P_x}{P_\xi}\right)_{\text{вых}} = n\left(\frac{P_x}{P_\xi}\right)_{\text{вх}} .$$

В соответствии с условием задачи

$$\left(\frac{P_x}{P_\xi}\right)_{\text{вых}} = \left(\frac{a_x}{a_\xi}\right)^2_{\text{вых}} = 16 ,$$

$$\left(\frac{P_x}{P_\xi}\right)_{\text{вх}} = \frac{a_x^2}{\sigma_\xi^2} = 0{,}16 .$$

Следовательно, продолжительность обработки сигнала в приемнике должна быть равна $T_{\text{пр}} = nT = \dfrac{0{,}1 \cdot 16}{0{,}16} = 10\,\text{с}.$

5.2. Так как по условию задачи помеха аддитивна и выборка X представляет одномерную величину, то функции правдоподобия $L(a_2)$ и $L(a_1)$ определяются законом распределения помехи:

$$L(a_2) = f(X / a_2) = \frac{1}{\sigma_\xi \sqrt{2\pi}} e^{-\frac{(X-a)^2}{2\sigma_\xi^2}},$$

$$L(a_1) = f(X / a_1) = \frac{1}{\sigma_\xi \sqrt{2\pi}} e^{-\frac{X^2}{2\sigma_\xi^2}}.$$

Отношение правдоподобия при этом

$$\lambda = e^{\frac{a^2}{\sigma_\xi^2}\left[\frac{X}{a} - \frac{1}{2}\right]}.$$

Графики функций правдоподобия $L(a_2)$, $L(a_1)$ и λ приведены на рис. 11.13. При использовании критерия максимума правдоподобия пороговое значение $\lambda_0 = 1$. Тогда условие для порогового значения входного сигнала будет иметь вид $\lambda = 1$ или $\frac{X_{\text{П}}^1}{a} - \frac{1}{2} = 0$.

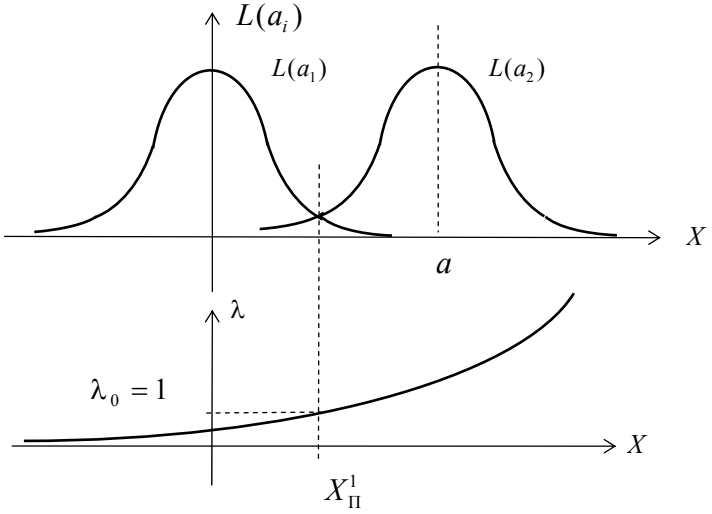


Рис. 11.13. Графики функций правдоподобия $L(a_i)$ и отношения функций правдоподобия λ

Таким образом, приемное устройство – это устройство сравнения, сопоставляющее входной сигнал с пороговым уровнем $X_{\text{П}}^1 = \frac{a}{2}$. Если амплитуда входного сигнала больше уровня $X_{\text{П}}^1$, в принятом сигнале содержится полезный сигнал. В противном случае входной сигнал определяется одной помехой.

5.3. Пороговый уровень измерения входного сигнала находится в соответствии с выражением

$$X_{\text{П}}^2 = \frac{\sigma_\xi^{\,2}}{a} \ln \frac{P(a_1)}{P(a_2)} + \frac{a}{2} + m_\xi \ .$$

5.4. Пороговый уровень измерения входного сигнала будет иметь вид

$$X_{\text{П}}^3 = \frac{\sigma_\xi^{\,2}}{a} \ln \frac{r_{21} P(a_1)}{r_{12} P(a_2)} + \frac{a}{2} + m_\xi \ .$$

5.5. Известно, что пороговый уровень измерения входного сигнала можно определить из соотношения

$$\alpha = \int_{\lambda_0}^{\infty} f(\lambda / a_1) d\lambda = \varepsilon_0 \ .$$

Если перейти от отношения функций правдоподобия λ к одномерной переменной X, то можно получить

$$\alpha = \int_{X_{\text{П}}}^{\infty} f(X / a_1) dX = \varepsilon_0 \ .$$

Подставляя выражение для плотности $f(X / a_1)$, получим новое уравнение

$$\alpha = \int_{X_{\text{П}}}^{\infty} \frac{1}{\sigma_\xi \sqrt{2\pi}} e^{-\frac{X^2}{2\sigma_\xi^{\,2}}} dX = \varepsilon_0 \ .$$

Наконец после замены переменных $z = \frac{X}{\sigma_\xi}$ получаем окончательное уравнение для нахождения $X_{\text{П}}$:

$$\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{X_П}{\sigma_\xi}}^{\infty} e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\infty} e^{-\frac{z^2}{2}}dz - \frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\frac{X_П}{\sigma_\xi}} e^{-\frac{z^2}{2}}dz =$$

$$= \frac{1}{2} - F\left(\frac{X_П}{\sigma_\xi}\right) = \varepsilon_0 ,$$

где $F\left(\frac{X_П}{\sigma_\xi}\right) = \frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\frac{X_П}{\sigma_\xi}} e^{-\frac{z^2}{2}}dz$ – интеграл вероятности.

По условию задачи $F\left(\frac{X_П}{\sigma_\xi}\right) = \frac{1}{2} - \varepsilon_0 = 0{,}45$.

По таблицам интеграла вероятности находим $\frac{X_П}{\sigma_\xi} = 1{,}65$, откуда искомый пороговый уровень $X_П = 1{,}65\sigma_\xi = 2{,}06$.

5.6. Вектор отсчетов помехи $\xi = (\xi_1, \xi_2)$ определяется двумерной плотностью распределения вероятностей $f(\xi) = f(\xi_1\xi_2)$. Полагая помеху стационарной и отсчеты некоррелированными, можно двумерный нормальный закон распределения помехи представить в виде

$$f(\xi) = f(\xi_1)f(\xi_2) = \left(\frac{1}{\sigma_\xi\sqrt{2\pi}}\right)^2 e^{-\frac{\sum\limits_{k=1}^{2}(\xi_k - m_\xi)^2}{2\sigma_\xi{}^2}} .$$

Тогда для двумерного вектора отсчетов $X = (x_1, x_2)$ при разных гипотезах закон распределения примет вид

$$f(X/a_2) = f(x_1/a_2)f(x_2/a_2) = \left(\frac{1}{\sigma_\xi\sqrt{2\pi}}\right)^2 e^{-\frac{\sum\limits_{k=1}^{2}[x_k - (a+m_\xi)]^2}{2\sigma_\xi{}^2}} ,$$

$$f(X/a_1)=f(x_1/a_1)f(x_2/a_1)=\left(\frac{1}{\sigma_\xi\sqrt{2\pi}}\right)^2 e^{-\frac{\sum_{k=1}^{2}(x_k-m_\xi)^2}{2\sigma_\xi{}^2}}.$$

Наконец, отношение функций правдоподобия $\lambda=\dfrac{f(X/a_2)}{f(X/a_1)}$ и $\lambda_0=\dfrac{r_{21}P(a_1)}{r_{12}P(a_2)}$ позволяют окончательно определить наличие или отсутствие полезного сигнала. Если $\lambda>\lambda_0$, то a_2, иначе a_1.

5.7. Условные плотности распределения вероятностей выборки $Y=\{y_1(t),y_2(t)\}$ при разных гипотезах в предположении, что отсчеты помех $\xi_1(t)$ и $\xi_2(t)$ статистически независимы, принимают вид

$$f(Y/a_2)=f(y_1/a_2)f(y_2/a_2)=\left(\frac{1}{\sigma_\xi\sqrt{2\pi}}\right)^2 e^{-\frac{\sum_{k=1}^{2}[y_k-(a+m_\xi)]^2}{2\sigma_\xi{}^2}},$$

$$f(Y/a_1)=f(y_1/a_1)f(y_2/a_1)=\left(\frac{1}{\sigma_\xi\sqrt{2\pi}}\right)^2 e^{-\frac{\sum_{k=1}^{2}(y_k-m_\xi)^2}{2\sigma_\xi{}^2}}.$$

5.8. Апостериорные вероятности гипотез вычисляются по формуле Байеса, которая в данном контексте задачи выглядит следующим образом

$$P(a_i/Y)=\frac{P(a_i)f(Y/a_i)}{\sum_{j=1}^{2}P(a_j)f(Y/a_j)},\ i=\overline{1,2}\ .$$

С учетом условных плотностей распределения вероятностей выборки $Y=\{y_1(t),y_2(t)\}$ при разных гипотезах, полученных в задаче 5.7, искомые вероятности принимают окончательный вид $P(a_1/Y)=0{,}268$, $P(a_2/Y)=0{,}732$.

6.1. Собственная информация некоторого сообщения x_k находится в соответствии с выражением $J(x_k) = -\log_2 p(x_k)$.

Следовательно:

а) $J(x_k) = -\log_2 \frac{5}{15} \cdot \frac{4}{14} \cdot \frac{10}{13} = 3{,}771$ бит;

б) $J(x_k) = -\log_2 \frac{\binom{2}{5} \cdot \binom{1}{10}}{\binom{3}{15}} = 2{,}186$ бит.

6.2. $J(x_k) = -\log_2 (1-p)^2 = 9{,}288$ бит.

6.3. Вероятность $p(1)$ нахождения системы в состоянии 1 находится из уравнения $p(1) = p(1) \cdot \frac{1}{2} + p(2) \cdot \frac{3}{4}$ при условии, что $p(1) + p(2) = 1$.

Тогда $J(x_k) = -\log_2 p(1) = -\log_2 0{,}6 = 0{,}737$ бит.

6.4. $J(x_k) = -\log_2 0{,}4 = 1{,}322$ бит.

6.5. $J(x_k) = -\log_2 \frac{6}{36} = 2{,}585$ бит.

6.6. Энтропия представляет собой математическое ожидание собственной информации сообщений заданного множества X:

$$H(X) = -\sum_X p(x_k) \log_2 p(x_k).$$

Учитывая тот факт, что $\sum_X p(x_k) = 1$, зависимость будет иметь вид кривой, представленной на рис. 11.14.

6.7. Приведем некоторые количественные меры информации, которые позволят найти ответы на все пункты задачи:

$$H(Y/X) = -\sum_{XY} p(x_k y_i) \log_2 p(y_i / x_k),$$

$$H(X;Y) = \sum_{XY} p(x_k y_i) J(x_k; y_i) = H(X) - H(X/Y) =$$

$$= H(Y) - H(Y/X) = H(X) + H(Y) - H(XY),$$

$$H(XY) = -\sum_{XY} p(x_k y_i) \log_2 p(x_k y_i) = H(X) + H(Y/X),$$

$$J(X; y_i) = \sum_{X} p(x_k / y_i) J(x_k; y_i) \ ,$$

$$J(x_k; Y) = \sum_{Y} p(y_i / x_k) J(x_k; y_i) \ ,$$

$$H(Y/x_k) = -\sum_{Y} p(y_i / x_k) \log_2 p(y_i / x_k).$$

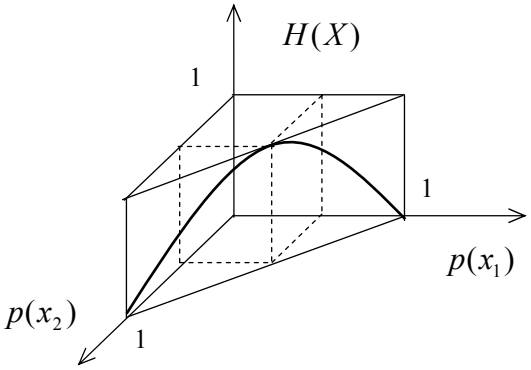


Рис. 11.14. Энтропия $H(X) = F[p(x_1), p(x_2)]$

6.8. Распределение вероятностей взаимной информации приведено на рис. 11.15, функция распределения на рис. 11.16, а взаимная энтропия равна $H(X;Y) = m\left[J(x_k; y_i)\right] = 0{,}86$ бит.

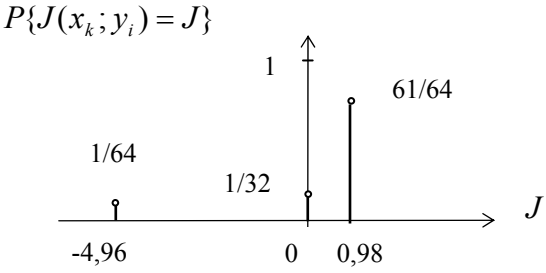


Рис. 11.15. Распределение вероятностей

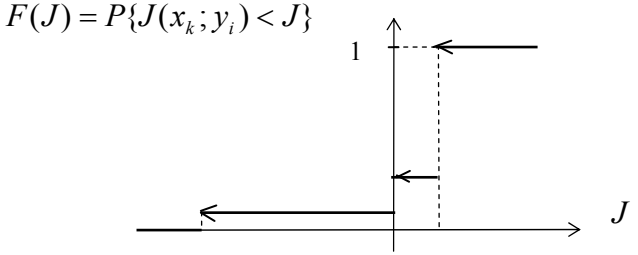


Рис. 11.16. Функция распределения вероятностей

6.9. $H(U) = 1{,}75$ бит, $\bar{n} = 1{,}75$, $p_0 = \dfrac{\overline{n_0}}{\bar{n}} = \dfrac{\frac{1}{2}N + \frac{1}{4}N + \frac{1}{8}N}{\frac{7}{4}N}$.

6.10. Взаимная информация определяет количество информации, которое событие y_i содержит в себе о событии x_k:

$$J(x_k; y_i) = \log_2 \frac{p(x_k / y_i)}{p(x_k)} = \log_2 \frac{p(y_i / x_k)}{p(y_i)} =$$

$$= \log_2 \frac{p(x_k y_i)}{p(x_k) p(y_i)} \quad .$$

Исходные данные представим в формализованном виде

Средство ПВО x_k	$p(x_k)$	$p(y / x_k)$
x_1	1/4	0
x_2	1/2	1/2
x_3	1/4	1

При необходимости недостающие условные вероятности могут быть найдены при использовании формулы Байеса

$$p(x_k / y) = \frac{p(x_k) p(y / x_k)}{p(y)} = \frac{p(x_k) p(y / x_k)}{\sum\limits_X p(x_k) p(y / x_k)} \quad .$$

В результате имеем:

$$J(x_1;y)=\log_2\frac{0}{1/2}=-\infty \quad \text{бит},$$

$$J(x_2;y)=\log_2\frac{1/2}{1/2}=0 \quad \text{бит},$$

$$J(x_3;y)=\log_2\frac{1}{1/2}=1 \quad \text{бит},$$

$$J(x_2;z)=\log_2\frac{p(z/x_2)}{p(z)}=\log_2\frac{\left[p(y/x_2)\right]^3}{\sum\limits_{k=1}^{3}p(x_k)\left[p(y/x_k)\right]^3}=-1{,}322 \quad \text{бит}.$$

6.11. $H(U^{\nu})=\nu\cdot H(U)=81{,}131$ бит.

7.1. Метод конструирования кодовых слов Шеннона–Фано реализует два принципа:

– в каждой позиции кодового слова символы алфавита должны использоваться с равными вероятностями;

– вероятности появления символов в каждой позиции не должны зависеть от расположения всех предыдущих символов.

В табл. 11.1 приводится схема построения множества кодовых слов, а на рис. 11.17 – соответствующее кодовое дерево.

Таблица 11.1

<table>
<tr><th rowspan="2">Сообщения x_k</th><th rowspan="2">Вероятности $p(x_k)$</th><th colspan="4">Символы разрядов кодовых слов</th><th rowspan="2">$\text{Код}_{\text{шф}}$</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th></tr>
<tr><td>x_1</td><td>0,3</td><td rowspan="2">0</td><td>0</td><td rowspan="2">-</td><td rowspan="5">-</td><td>00</td></tr>
<tr><td>x_2</td><td>0,2</td><td>1</td><td>01</td></tr>
<tr><td>x_3</td><td>0,2</td><td rowspan="5">1</td><td rowspan="2">0</td><td>0</td><td>100</td></tr>
<tr><td>x_4</td><td>0,1</td><td>1</td><td>101</td></tr>
<tr><td>x_5</td><td>0,1</td><td rowspan="3">1</td><td>0</td><td>110</td></tr>
<tr><td>x_6</td><td>0,05</td><td rowspan="2">1</td><td>0</td><td>1110</td></tr>
<tr><td>x_7</td><td>0,05</td><td>1</td><td>1111</td></tr>
</table>

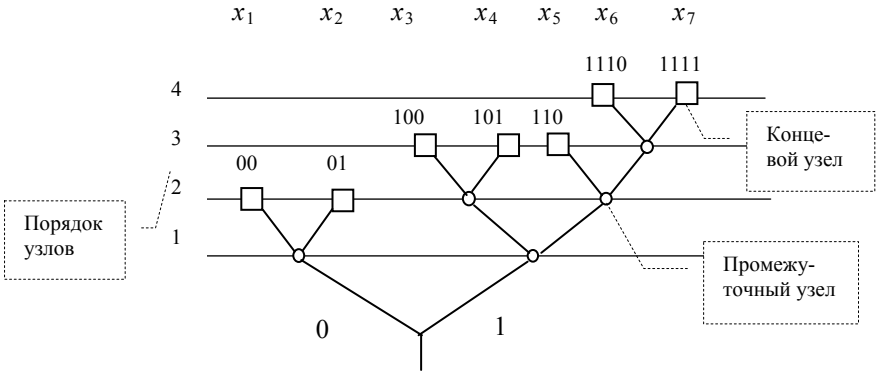


Рис. 11.17. Кодовое дерево кода Шеннона - Фано

Эффективность кодирования составила

$$\text{Э}_{\text{ШФ}} = \frac{H(U)}{\overline{n}\log_2 D} = \frac{2{,}48}{2{,}6} = 0{,}955 \ .$$

7.2. Процедура кодирования приведена на рис. 11.18. Оптимальный троичный код приведен на рис. 11.19 в виде неполного кодового дерева. Код характеризуется следующими параметрами:

$$H(X) = 2{,}43 \ , \quad \overline{n} = \sum_{k=1}^{6} n_k p(x_k) = 1{,}65 \ ,$$

$$\text{Э} = \frac{H(X)}{\overline{n}\log_2 D} = 0{,}93 \ .$$

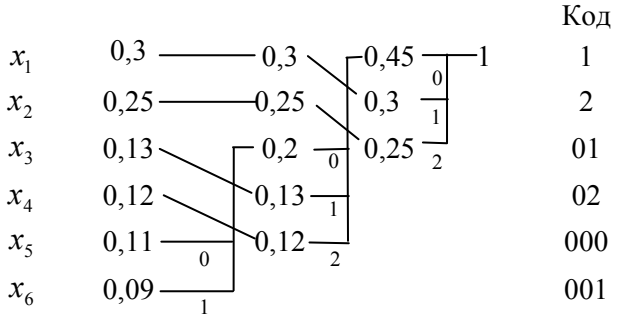


Рис. 11.18. Процедура оптимального кодирования для $D = 3$

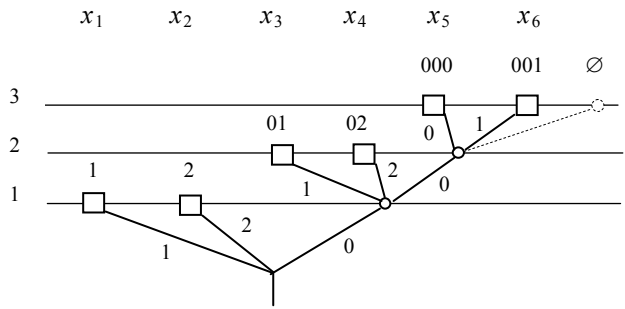


Рис. 11.19. Неполное кодовое дерево

При использовании произвольного алфавита кодировки ($D \neq 2$) процедура Хаффмана предполагает на первом шаге объединение m_0 сообщений с наименьшими вероятностями и D на последующих шагах.

Число m_0 находится из условия

$$\begin{cases} 2 \le m_0 \le D; \\ \left[\dfrac{M - m_0}{D-1} \right] = a, \end{cases}$$

где a – целое положительное число.

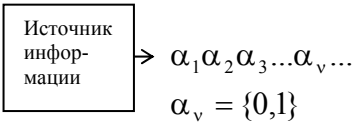
7.3. При применении процедуры Шеннона – Фано целесообразно первоначальное множество расположить в порядке убывания вероятностей и именно в такой последовательности пытаться проводить разбиения на три примерно равновероятных группы. Как показывает практика построения кодов, это обеспечивает наибольшую эффективность кодирования, так как в каждой группе будут фигурировать сообщения с равновеликими вероятностями.

Что касается процедуры Хаффмана, то в ней все шаги строго регламентированы и никаких неоднозначностей кодирования нет. Произвольное присвоение символов алфавита кодировки различным ветвям дерева не нарушает общую структуру оптимального кода.

8.1. Для применения процедуры Хаффмана необходимо построить множество, которое является декартовым произведением мно-

жеств U^3, и найти соответствующее распределение вероятностей (рис. 11.20)

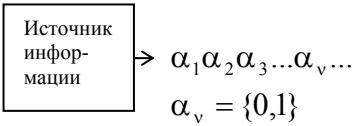
$$p(\alpha^3) = p(\alpha_1\alpha_2\alpha_3) = p(\alpha_1)p(\alpha_2)p(\alpha_3)\ .$$



$i = \overline{1,M}$	α^3	$p(\alpha^3)$	Код
1	000	1/27	1011
2	001	2/27	1000
3	010	2/27	1001
4	011	4/27	11
5	100	2/27	1010
6	101	4/27	010
7	110	4/27	011
8	111	8/27	00

Рис. 11.20. Схема решения задачи для $D = 2$

8.2. Элементы решения задачи приведены на рис. 11.21.



$i = \overline{1,M}$	α^3	$p(\alpha^3)$	Код
1	000	1/64	201
2	001	3/64	21
3	010	3/64	22
4	011	9/64	10
5	100	3/64	200
6	101	9/64	11
7	110	9/64	12
8	111	27/64	0

Рис. 11.21. Схема решения задачи для $D = 3$

8.3. Процедура асимптотически эффективного кодирования состоит из следующих шагов:

– находим наибольшее M число последовательностей $\alpha^\nu \in T$, удовлетворяющее условиям

$$\begin{cases} M < 2^{\nu[H(U)+\varepsilon]}; \\ J(\alpha^{\nu}) \le \nu[H(U)+\varepsilon]; \end{cases}$$

– из условия $D^{n_\nu} \ge M+1$ определяем длину кодовых слов как наименьшее большее целое

$$n_\nu \ge \left\lceil \frac{\log_2(M+1)}{\log_2 D} \right\rceil .$$

Идея кодирования состоит в том, что M последовательностей из ν букв представляются различными кодовыми словами одинаковой длины n_ν, а все остальные последовательности – одним и тем же кодовым словом. Возникающая при этом вероятность неоднозначного кодирования равна вероятности появления сообщений из множества $\alpha^{\nu} \notin T$.

Результаты кодирования приведены в табл. 11.2.

Таблица 11.2

	α^2	$p(\alpha^2)$	$J(\alpha^2) = -\log_2 p(\alpha^2)$	Код
1	u_1u_1	1/4	2	00
2	u_1u_2	1/6	2,58	01
3	u_1u_3	1/12	3,58	11
4	u_2u_1	1/6	2,58	10
5	u_2u_2	1/9	3,17	11
6	u_2u_3	1/18	4,17	11
7	u_3u_1	1/12	3,58	11
8	u_3u_2	1/18	4,17	11
9	u_3u_3	1/36	5,17	11

8.4. Расчет условных энтропий проводится по следующим формулам:

$$H(A) = -\sum_{\alpha_1 \in A} p(\alpha_1) \log_2 p(\alpha_1) = 2,$$

$$H(A/A^1) = -\sum_{\alpha_1 \in A} \sum_{\alpha_2 \in A} p(\alpha_1) p(\alpha_2/\alpha_1) \log_2 p(\alpha_2/\alpha_1) = 1{,}75,$$

$$H(A/A^2) = -\sum_{\alpha_1 \in A} \sum_{\alpha_2 \in A} \sum_{\alpha_3 \in A} p(\alpha_1) p(\alpha_2/\alpha_1) p(\alpha_3/\alpha_1\alpha_2) \log_2 p(\alpha_3/\alpha_1\alpha_2) =$$

$$= -\sum_{\alpha_1 \in A} \sum_{\alpha_2 \in A} \sum_{\alpha_3 \in A} p(\alpha_1) p(\alpha_2/\alpha_1) p(\alpha_3/\alpha_2) \log_2 p(\alpha_3/\alpha_2) = 1{,}75,$$

...

$$H(A/A^\infty) = \lim_{\nu \to \infty} H(A/A^\nu) = 1{,}75.$$

8.5. Результаты кодирования приведены в табл. 11.3. Эффективность кодирования при этом составила

$$Э = \frac{H(A/A^\infty)}{(\overline{n_\nu}/\nu)\log_2 D} = \frac{1{,}75}{3{,}75/2} = 0{,}98 \quad .$$

Таблица 11.3

a_1a_1	1/8	000	a_1a_3	1/32	11000
a_2a_4		001	a_1a_4		11001
a_3a_3		010	a_2a_2		11010
a_4a_2		011	a_2a_3		11011
a_1a_2	1/16	1000	a_3a_1		11100
a_2a_1		1001	a_3a_2		11101
a_3a_4		1010	a_4a_1		11110
a_4a_3		1011	a_4a_4		11111

8.6. В соответствии с основной теоремой кодирования $2 \le \overline{n} < 2{,}1$.

8.7. Метод универсального кодирования (УК – метод Бабкина) применяется в случае неизвестных статистических характеристик источника сообщений и состоит в следующем. Пусть в накоплен-

ном n - блоке имеется k единиц и (n-k) нулей. Пронумеруем последовательно нули и единицы, встречающиеся в n - блоке. Пусть $i_1, i_2, \ldots, i_k$ – номера позиций единиц.

Положим $b_k = \binom{i_1 - 1}{1} + \binom{i_2 - 1}{2} + \ldots + \binom{i_k - 1}{k}$, где $\binom{\nu}{\mu}$ – биномиальный коэффициент, причем полагаем $\binom{\nu}{\mu} = 0$ при $\nu < \mu$.

В этом случае справедливы следующие утверждения:

1. $0 \le b_k \le \binom{n}{k} - 1$.

2. Соответствие между n - блоками и парами чисел (k, b_k) - взаимно однозначно.

Тогда кодовое слово, соответствующее n - блоку, определяется как упорядоченная пара двоичных наборов (k, b_k). При этом длина кодового слова $l_{n,k} =]\log_2(n+1)[+ \left]\log_2 \binom{n}{k}\right[$, где $]x[$ – наименьшее большое целое.

Для порожденной источником последовательности имеем

	i_1	i_2		i_3	i_4		i_5
1	2	3	4	5	6	7	8
0	1	1	0	1	1	0	1

$i_1 = 2$,
$i_2 = 3$,
$i_3 = 5$,
$i_4 = 6$, $\quad n = 8, k = 5$,
$i_5 = 8$, $\quad l_{n,k} = 4 + 6$.

Построим для этого случая таблицу вычисления биномиальных коэффициентов (табл. 9.5).

Таблица 11.4

	n	$\binom{n-1}{1}$	$\binom{n-1}{2}$	$\binom{n-1}{3}$	$\binom{n-1}{4}$	$\binom{n-1}{5}$
i_5	8	7	21	35	35	21
	7	6	15	20	15	6
i_4	6	5	10	10	5	1
i_3	5	4	6	4	1	0
	4	3	3	1	0	0
i_2	3	2	1	0	0	0
i_1	2	1	0	0	0	0
	1	0	0	0	0	0

В таблице слагаемые, соответствующие искомым биномиальным коэффициентам, обведены,

$$b_k = 32, \quad (k, b_k) \to (0101{,}100000).$$

Декодирование производится в обратном порядке; так как k=5, ищем в 5-м столбце наибольшее число, меньшее или равное 32. Вычисляем разность 32-21=11. В 4-м столбце имеем наибольшее число, меньшее или равное 11, и т.д. Таким же образом декодирование производится при любом n.

8.8. Для решения задачи достаточно воспользоваться свойством стационарности дискретного марковского источника. Источник стационарен, так как все распределения вероятностей не зависят от i и справедливо соотношение

$$\sum_{\alpha_{i-1} \in A_{i-1}} p(\alpha_{i-1}) p(\alpha_i / \alpha_{i-1}) = p(\alpha_i) \ .$$

Более того, для стационарного состояния источника можно принять

$$p(\alpha_i = 1) = p(\alpha_{i-1} = 1);$$
$$p(\alpha_i = 0) + p(\alpha_i = 1) = 1;$$
$$J = -\log_2 p(\alpha_i = 1) =$$
$$= -\log_2 0{,}4 = 1{,}322 \ .$$

9.1. Дискретный двоичный канал (ДК) без памяти задается матрицей условных (переходных) вероятностей

$$\left\| p(y_i / x_k) \right\| = \begin{vmatrix} 1-p_1 & p_1 \\ p_2 & 1-p_2 \end{vmatrix}.$$

В теории информации доказывается, что информационная пропускная способность дискретного стационарного канала может быть вычислена по формуле

$$\begin{cases} C = \max\limits_{\{p(x_k)\}} \{H(X;Y)\}, \\ \sum\limits_X p(x_k) = 1, \end{cases}$$

где $H(X;Y)$ – взаимная энтропия.

Вычислим пропускную способность ДК. Пусть $p(x_1) = z$, $p(x_2) = 1 - z$. Тогда, используя формулу полной вероятности $p(y_i) = \sum\limits_X p(x_k) p(y_i / x_k)$, представим вероятности выходного пространства в виде

$$p(y_1) = z(1-p_1) + (1-z)p_2 = \alpha z + p_2,$$
$$p(y_2) = zp_1 + (1-z)(1-p_2) = (1-p_2) - \alpha z,$$

где $\alpha = 1 - p_1 - p_2$.

Подставляя эти значения в

$$H(X;Y) = H(Y) - H(Y/X),$$
$$H(Y) = -\sum_Y p(y_i) \log_2 p(y_i),$$
$$H(Y/X) = -\sum_X \sum_Y p(x_k) p(y_i / x_k) \log_2 p(y_i / x_k),$$

получаем

$$H(Y) = -(\alpha z + p_2)\log_2(\alpha z + p_2) -$$
$$-\left[(1-p_2) - \alpha z\right]\log_2\left[(1-p_2) - \alpha z\right],$$
$$H(Y/X) = -z(1-p_1)\log_2(1-p_1) - zp_1 \log_2 p_1 -$$
$$-(1-z)p_2 \log_2 p_2 - (1-z)(1-p_2)\log_2(1-p_2).$$

Находим экстремум, приравнивая нулю производную $\frac{\partial H(X;Y)}{\partial z} = 0$. После несложных преобразований получаем

$$z_0 = \frac{1 - p_2 D}{\alpha D},$$

где $D = 1 + \exp(-B/\alpha)$,

$$B = p_1 \ln p_1 + (1-p_1)\ln(1-p_1) - p_2 \ln p_2 - (1-p_2)\ln(1-p_2) .$$

Итак, информационная пропускная способность канала вычисляется по формуле $C = H(X;Y)\big|_{z=z_0}$.

Поверхность $C = F[p_1, p_2]$ приведена на рис. 11.22.

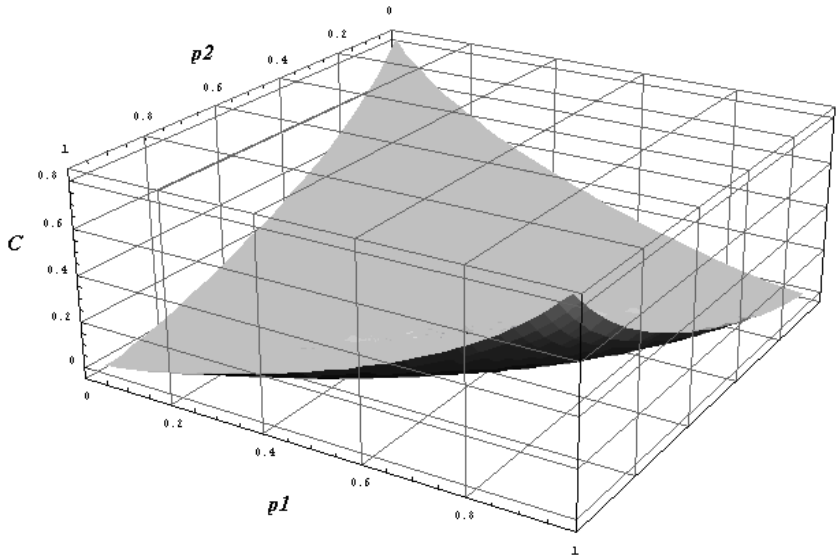


Рис. 11.22. Поверхность C для двоичного канала

9.2. Эквивалентный канал (рис. 11.23) симметричен по входу и выходу. Следовательно,

$$C = 1 + (1-2p+2p^2)\log_2(1-2p+2p^2) + (2p-2p^2)\log_2(2p-2p^2) .$$

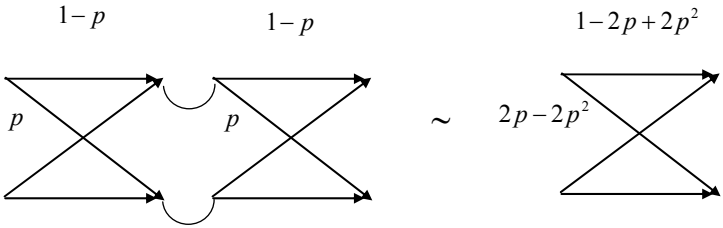


Рис. 11.23. Граф-схема эквивалентного канала

9.3. $C = 0{,}919$ бит.

9.4. Эквивалентный канал – ДСКС. Следовательно,

$$C = (1-p)^2 \log_2 (1-p)^2 + p^2 \log_2 p^2 -$$

$$- (1 - 2p + 2p^2) \log_2 \frac{(1 - 2p + 2p^2)}{2} \quad .$$

9.5. Эквивалентный канал – ДК. Следовательно,

$$C = 1 + (1-p)^2 \log_2 (1-p)^2 + (2p - p^2) \log_2 (2p - p^2) .$$

9.6. $C = 2 + p \log_2 \frac{p}{2} + (1-p) \log_2 \frac{1-p}{2}$.

9.7. $C = 2 + p \log_2 p + (1-p) \log_2 (1-p)$.

Данное решение находится с учетом того обстоятельства, что $\lim\limits_{p \to 0} p \log_2 p = 0$.

9.8. Решив экстремальную задачу с ограничением

$$\begin{cases} C = \max\limits_{\{p(x_k)\}} \{H(X;Y) = H(Y) - H(Y/X)\}, \\ \sum\limits_{k=1}^{3} p(x_k) = 1, \end{cases}$$

получим информационную пропускную способность заданного канала

$$C = \log_2 (1 + 2^{1 + p \log_2 p + (1-p) \log_2 (1-p)}) .$$

10.1. На приемном конце принята комбинация

$$\begin{array}{ccccccc} x_7 & x_6 & x_5 & x_4 & x_3 & x_2 & x_1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \end{array}.$$

Вычислив опознаватель

$$o_1 = x_1 + x_3 + x_5 + x_7 = 1,$$
$$o_2 = x_2 + x_3 + x_6 + x_7 = 0,$$
$$o_3 = x_4 + x_5 + x_6 + x_7 = 1,$$

исправляем 5-й разряд. Следовательно, был передан код 1001011.

Образующая матрица группового кода Хэмминга (7,4) имеет вид

$$M_{n,k} = \left| I_k^T : B_{n-k,k} \right| = \begin{array}{c} 7\ \ 6\ \ 5\ \ 3 \quad\ 4\ \ 2\ \ 1 \\ \hline \left| \begin{array}{ccccccc} 0 & 0 & 0 & 1 & : & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & : & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & : & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & : & 1 & 1 & 1 \end{array} \right| \end{array},$$

который позволяет достаточно просто убедиться в правильности принятого решения.

10.2. На приемном конце принята комбинация $\begin{array}{ccccccc} 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}$.

Производим деление многочленов

$$\begin{array}{l|l} x^6 + x^3 + x + 1 & x^3 + x + 1 \\ \hline x^6 + x^4 + x^3 & x^3 + x \\ \hline x^4 + x + 1 & \\ x^4 + x^2 + x & \\ \hline x^2 + 1 & \end{array}.$$

По таблице ошибок (табл. 10.11) находим, что искажение произошло в разряде 7. Исправляем этот разряд. Следовательно, был передан код 0001011.

Образующая матрица циклического кода Хэмминга (7,4) имеет вид

$$M_{n,k} = \left| I_k^T : B_{n-k,k} \right| = \begin{array}{c} \begin{array}{ccccccccc} 7 & 6 & 5 & 3 & & 4 & 2 & 1 \end{array} \\ \left| \begin{array}{ccccccccc} 0 & 0 & 0 & 1 & : & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & : & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & : & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & : & 1 & 0 & 1 \end{array} \right| \end{array},$$

который позволяет также убедиться в правильности принятого решения.

10.3. n-разрядные комбинации неразделимого циклического кода получаются умножением многочленов, соответствующих k-разрядным информационным комбинациям, на порождающий многочлен $g(x)$.

Искажение произошло в разряде 6. Следовательно, передавался код 0111, после кодирования которого на выходе кодера была комбинация 0100011.

10.4. Параметры группового кода определяются в соответствии с неравенствами $2^k - 1 \geq 3$, $2^{n-k} - 1 \geq n + \dfrac{n(n-1)}{2}$, т.е. $k \geq 2$, $n \geq 7$.

В этом случае составление таблицы опознавателей одиночных ошибок должно производиться с учетом того, что опознаватели двойных ошибок будут образовываться суммированием по модулю два опознавателей одиночных ошибок, составляющих эту двойную ошибку.

Удовлетворить этому требованию можно только при параметрах группового кода (8,2), который в свою очередь характеризуется образующей матрицей

$$M_{n,k} = \left| I_k^T : B_{n-k,k} \right| = \begin{array}{c} \begin{array}{ccccccccc} 8 & 5 & & 7 & 6 & 4 & 3 & 2 & 1 \end{array} \\ \left| \begin{array}{ccccccccc} 0 & 1 & : & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & : & 1 & 1 & 0 & 0 & 1 & 1 \end{array} \right| \end{array}.$$

10.5. Циклический код строится путем приписывания к каждой неизбыточной комбинации информационных символов остатка от деления этой комбинации на $g(x)$. При этом под информацион-

ные символы принято отводить k старших разрядов, а под проверочные – $(n-k)$ младших. Процесс построения кода приведен в табл. 11.5.

Таблица 11.5

Код	$a_i(x)x^3$	$r_i(x)=\left]\dfrac{a_i(x)x^3}{g(x)}\right[$	$\widetilde{a}_i(x)=a_i(x)x^3+r_i(x)$	
0001	x^3	x^2+1	x^3+x^2+1	0001:101
0010	x^4	x^2+x+1	x^4+x^2+x+1	0010:111
0100	x^5	$x+1$	x^5+x+1	0100:011
1000	x^6	x^2+x	x^6+x^2+x	1000:110

10.6. Известно, что неразделимый циклический код получается умножением многочленов, соответствующих k-разрядным информационным комбинациям, на порождающий многочлен $g(x)$. Процесс построения кода приведен в табл. 11.6.

Таблица 11.6

Информационный код	$a_i(x)$	$\widetilde{a}_i(x)=a_i(x)\cdot g(x)$	Циклический код
0001	1	x^3+x+1	0001011
0010	x	x^4+x^2+x	0010110
0100	x^2	$x^5+x^3+x^2$	0101100
1000	x^3	$x^6+x^4+x^3$	1011000

10.7. Параметры циклического кода определяются в соответствии с неравенствами: $2^k-1\ge 1$, $2^{n-k}-1\ge n+n(n-1)/2$, т.е. $k\ge 1$, $n\ge 5$.

Выбираем порождающий многочлен $\deg g(x)=n-k=4$, который являлся бы делителем двучлена $x^5+1\equiv 0 \bmod g(x)$. В разложении $x^5+1=(x+1)(x^4+x^3+x^2+x+1)$ такой многочлен единст-

венный – $g(x) = x^4 + x^3 + x^2 + x + 1$. Тогда образующая матрица принимает вид $M_{5,1} = |1:1111|$.

Таблица одиночных и двойных ошибок приведена в табл. 11.7.

Таблица 11.7

i	$\xi_i(x)$	$\left]\frac{\xi_i(x)}{g(x)}\right[$	Код - опознаватель
1	1	1	0001
2	x	x	0010
3	x^2	x^2	0100
4	x^3	x^3	1000
5	x^4	$x^3 + x^2 + x + 1$	1111
6	$x+1$	$x+1$	0011
7	x^2+1	x^2+1	0101
8	x^3+1	x^3+1	1001
9	x^4+1	$x^3 + x^2 + x$	1110
10	$x^2 + x$	$x^2 + x$	0110
11	$x^3 + x$	$x^3 + x$	1010
12	$x^4 + x$	$x^3 + x^2 + 1$	1101
13	$x^3 + x^2$	$x^3 + x^2$	1100
14	$x^4 + x^2$	$x^3 + x + 1$	1011
15	$x^4 + x^3$	$x^2 + x + 1$	0111

10.8. Аппаратная реализация группового кода Хэмминга (3,1) с образующей матрицей $M_{3,1} = |1:11|$ приведена на рис. 11.24.

10.9. Аппаратная реализация циклического кода Хэмминга (3,1) с образующей матрицей $M_{3,1} = |1:11|$ приведена на рис. 11.25.

10.10. Коды БЧХ строятся по заданным *N* и *T*. Значение *k* (число информационных разрядов) неизвестно, пока не найден порождающий многочлен кода *G*(*x*). В конкретном случае

$$G(x) = \text{НОК}[F_1(x), F_3(x), ..., F_i(x), ..., F_{D-2}(x)] \quad ,$$

где $F_i(x)$ – минимальные многочлены, а D=2T+1 – минимальное расстояние Хэмминга кода БЧХ. Длина кода БЧХ, как правило, определяется из выражения $N = 2^m - 1$, где m - любое целое число. Иногда возникает ситуация, когда вычисленное по формуле $m = \log_2(N+1)$ получается не целым. Тогда необходимо пользоваться формулой $m = \log_2(NC_0 + 1)$, подбирая минимальное C_0 такое, чтобы результат оказался целочисленным. Число проверочных разрядов кода БЧХ $r \le mT$, отсюда число информационных разрядов $k \ge \dfrac{2^m - 1}{C_0} - mT$.

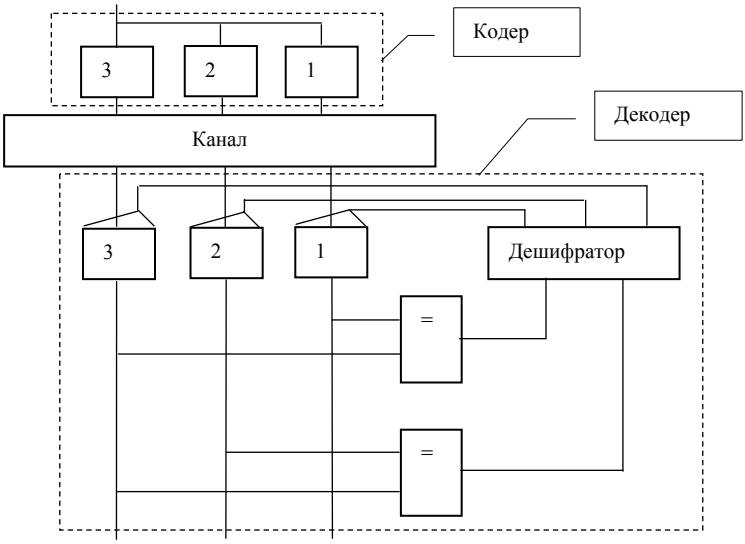


Рис. 11.24. Кодер и декодер группового кода (3,1)

Если $r > \dfrac{2^m - 1}{C_0}$, то код БЧХ не определен, так как число проверочных разрядов не может быть больше длины кода.

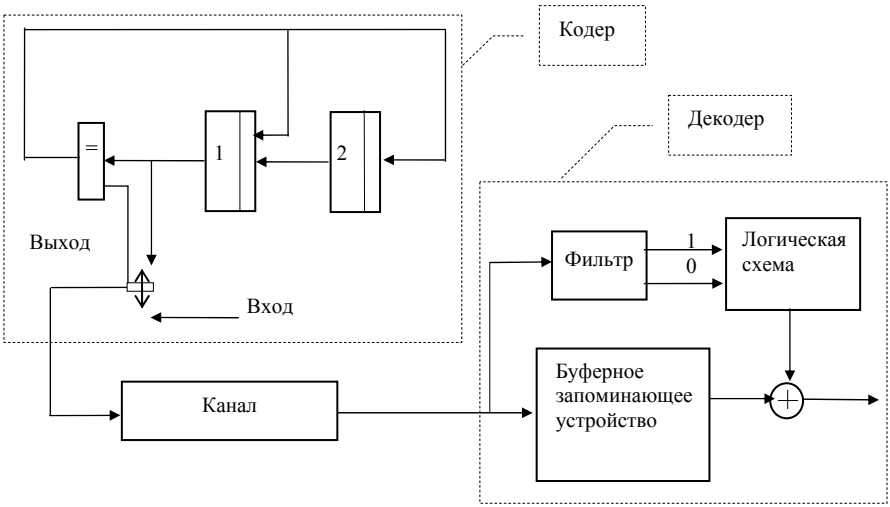


Рис. 11.25. Кодер и декодер циклического кода (3,1)

Порождающий многочлен кода длины N=21, служащий для исправления двухкратных ошибок T=2, получается следующим образом. Формула $m = \log_2(N+1)$ к положительным результатам не приводит, а формула $m = \log_2(NC_0 + 1)$ при $C_0 = 3$ дает m=6. Далее обращаемся к таблице минимальных многочленов (табл. 10.6). Выбираем минимальные многочлены $F_1(x)$ и $F_3(x)$, так как D-2=3. Умножаем индексы выбранных многочленов на C_0 и окончательно получаем

$$G(x) = \text{НОК}[F_3(x), F_9(x)] =$$
$$\text{НОК}[x^6 + x^4 + x^2 + x + 1, x^3 + x^2 + 1] =$$
$$= x^9 + x^8 + x^7 + x^5 + x^4 + x + 1 \quad .$$

Следовательно, $r = \deg G(x) = 9$ и $k = 12$. Строки образующей матрицы кода (21,12) находятся как остаток от деления произведения $x^j G(x)$ на двучлен $x^N + 1$ (рис. 11.26).

$$M_{21,12} = \left| \begin{array}{l} 000000000001{:}110110011 \\ 000000000010{:}011010101 \\ 000000000100{:}110101010 \\ 000000001000{:}011100111 \\ 000000010000{:}111001110 \\ 000000100000{:}000101111 \\ 000001000000{:}001011110 \\ 000010000000{:}010111100 \\ 000100000000{:}101111000 \\ 001000000000{:}101000011 \\ 010000000000{:}100110101 \\ 100000000000{:}111011001 \end{array} \right|$$

Рис. 11.26. Образующая матрица кода БЧХ (21,12)

Традиционный подход исправления ошибок достаточно трудоемкий, так как получить $\sum_{i=1}^{T} \binom{N}{i}$ опознавателей не представляется целесообразным. Воспользуемся следующим алгоритмом декодирования.

Допустим, что передавалась кодовая комбинация $A(x)$, построенная на основе образующей матрицы кода БЧХ (N, k) с порождающим многочленом $G(x)$, способным исправлять ошибки кратности T включительно.

В результате воздействия помех в канале связи в приемнике получена комбинация $A^*(x) = A(x) + E(x)$, где $E(x)$ – многочлен ошибки. Делим $A^*(x)$ на $G(x)$ и вычисляем вес W полученного остатка. Если $W > T$, производим циклический сдвиг предыдущего делимого влево на один разряд и опять выполняем деление на многочлен $G(x)$.

Эта процедура повторяется до получения остатка с весом не больше T. Если $W \le T$, складываем последнее делимое с полученным остатком и производим циклический сдвиг полученной комбинации вправо на столько разрядов, на сколько осуществлялся сдвиг влево. Получаем истинную переданную комбинацию.

Процесс исправления ошибок для циклического кода БЧХ (21,12) приведен на рис. 11.27.

10.11. Образующая матрица кода БЧХ (31,1) имеет вид

$$M_{31,1} = \left| 1{:}1111111111\ 1111111111\ 1111111111 \right| .$$

10.12. Циклический код Файра позволяет обнаружить пакет ошибок длиной b и исправить одиночную вспышку (пакет) ошибок длиной l. Порождающий многочлен кода определяется из выражения $g(x) = p(x)(x^c + 1)$, где $p(x)$ – неприводимый многочлен степени m. Длина кода n равна наименьшему общему кратному e и c: $n = \text{НОК}\{e, c\}$, где e – показатель для $p(x)$, т.е. наименьшее число, при котором

$$x^e + 1 \equiv 0 \bmod p(x) .$$

Передаваемое сообщение:
A(X)= 11000000000011101100.
Кратность ошибки: T= 2.
Вектор ошибки:
E(X)= 11000000000000000000.
В результате действия помех в канале связи на приемном конце была принята комбинация:
A"(X)=00000000000011101100.
Проводим деление принятой комбинации на G(X) .

Цикл N 1

Остаток: 11101100 с весом W = 5.
Так как W>T , то надо произвести циклический сдвиг A"(X) влево на 1 разряд и снова произвести деление на G(X) .

Цикл N 2

Остаток: 111011000 с весом W = 5.
Так как W>T , то надо произвести циклический сдвиг A"(X) влево на 1 разряд и снова произвести деление на G(X) .

Цикл N 3

Остаток: 11 с весом W = 2.
Теперь, когда W<=T , можно произвести сложение по модулю 2 циклически сдвинутой комбинации A"(X) и полученного остатка.
Комбинация A"(X) , циклически сдвинутая влево на 2 разряда:
00000000000001110110000.
Сумма циклически сдвинутой комбинации A"(X) и полученного остатка:
00000000000001110110011.
Наконец, проведем обратный циклический сдвиг этой суммы вправо на 2 разряда.
Получили переданное сообщение:
11000000000011101100.

Рис. 11.27. Процесс исправления ошибок

В частности, если $p(x)$ является примитивным многочленом, то $e = 2^m - 1$. Степень двучлена должна удовлетворять условию $c \geq b + l - 1$, причем c не должно делиться нацело на $2^m - 1$.

Число проверочных разрядов равно $r = m + c$. Код позволяет обнаружить пакет ошибок длиной $b \le l$ и исправить вспышку ошибок длиной $l \le m$.

В соответствии с условием задачи степень $\deg p(x) = m \ge l$. Степень двучлена $c \ge b + l - 1 = 9$.

Выберем в качестве неприводимого многочлена $p(x) = x^5 + x^2 + 1$. Порождающий многочлен имеет вид $g(x) = (x^5 + x + 1)(x^9 + 1)$. Длина кодовой комбинации $n = (2^5 - 1)9 = 279$, а число проверочных символов $m = 9 + 5 = 14$. Следовательно, код Файра имеет параметры (279,265).

10.13. Избыточность кода, характеризуемая отношением числа проверочных символов к длине комбинации, в данном случае составит 14/279=0,05.

Любой циклический код, рассчитанный на исправление T-кратных ошибок, может быть использован для исправления пакетов ошибок длины не более T. Однако избыточность кода будет существенно выше по сравнению с циклическим кодом Файра, имеющим ту же корректирующую способность. Так, если для исправления пакетов ошибок длины 5 воспользоваться БЧХ кодом с близким значением $n = 255$, то при $T = 5$ число проверочных символов составит 40, а избыточность – 40/255=0,16.

Это подтверждает вполне очевидный вывод: исправить пакет ошибок проще, чем то же количество ошибок, рассредоточенных по всей комбинации.

10.14. Параметры циклических кодов Файра с $l = m$ приведены в табл. 11.8.

Таблица 11.8

i	(n,k)	$l = m$
1	(693,676)	6
2	(1651,1631)	7
3	(3825,3802)	8
4	(8687,8661)	9
5	(19437,19408)	10

10.15. Результаты получены с использованием компьютерного учебника ОТИК 4.15 (рис. 11.28).

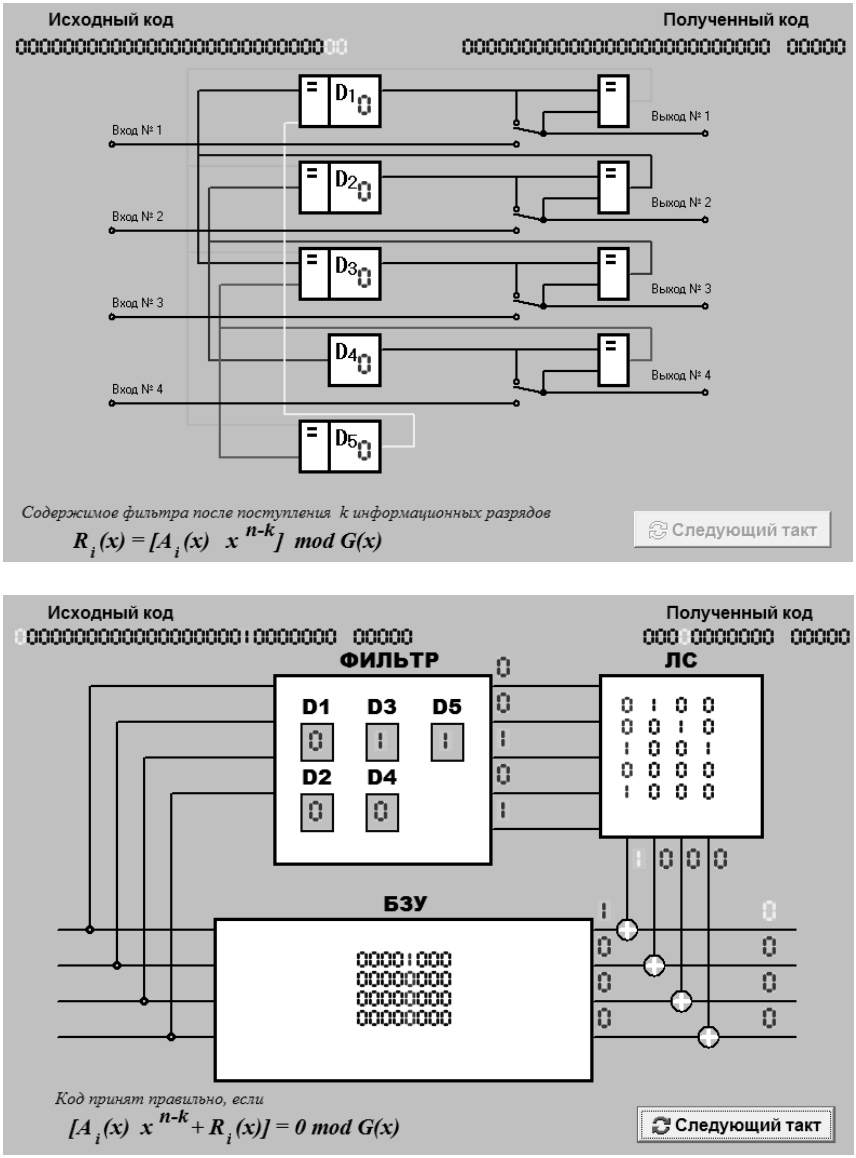


Рис. 11.28. Результаты синтеза кодера и декодера

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

(жирным шрифтом выделена основная литература)

1. Котоусов А.С. Теория информации. Учебное пособие для вузов. – М.: Радио и связь, 2003.

2. **Липкин И.А. Статистическая радиотехника. Теория информации и кодирования. – М.: Вузовская книга, 2002.**

3. **Гоноровский И.С. Радиотехнические цепи и сигналы. – М.: Советское радио, 1986.**

4. **Темников Ф.Е. и др. Теоретические основы информационной техники. – М.: Энергия, 1979.**

5. Кузьмин И.В., Кедрус В.А. Основы теории информации и кодирования. – Киев: Вища школа, 1986.

6. Федосеев Ю.Н. Элементы теории передачи и кодирования дискретной информации в АСУ. – М.: МИФИ, 1981.

7. Блейхут Р. Теория и практика кодов, контролирующих ошибки: Пер. с англ. – М.: Мир, 1986.

8. Абдуллаев Д.А., Арипов Н.Н. Передача дискретных сообщений в задачах и упражнениях. Учебное пособие для вузов. – М.: Радио и связь, 1985.

9. **Березкин Е.Ф. Основы теории информации и кодирования. Лабораторный практикум: Учебно-методическое пособие. – 2-е изд., перераб. и доп. – М.: МИФИ, 2009.**

10. Фано Р. Передача информации. Статистическая теория связи. – М.: Мир, 1965.

11. Дмитриев В.И. Прикладная теория информации. Учебник для вузов по специальности "Автоматизированные системы обработки информации и управления". – М.: Высшая школа, 1989.

12. Хэмминг Р.В. Теория кодирования и теория информации. – М.: Радио и связь, 1989.

13. Хетагуров Я.А., Руднев Ю.П. Повышение надежности цифровых устройств методами избыточного кодирования. – М.: Энергия, 1974.

14. Древс Ю.Г., Реутов В.Ф. Ортогональные преобразования сигналов и их применение в технике. – М.: МИФИ, 1988.

15. **Березкин Е.Ф. Сборник задач по курсу "Основы теории информации и кодирования". – М.: МИФИ, 2002.**

16. Панин В.В. Основы теории информации. Ч.1. – М.: МИФИ, 2001.

17. Панин В.В. Основы теории информации. Ч.2. Введение в теорию кодирования. – М.: МИФИ, ФГУП ИСС, 2004.

18. Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов: Учебное пособие для высших учебных заведений. – М.: Радио и связь, 1990.

19. Панин В.В. Основы теории информации: учебное пособие для вузов. – 2-е изд., испр. и доп. – М.: БИНОМ. Лаборатория знаний, 2007.

20. Литвинская О.С., Чернышев Н.И. Основы теории передачи информации: учебное пособие. – М.: КНОРУС, 2010.

Приложение 1. Таблица значений интеграла вероятностей

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$$

x	$F(x)$	x	$F(x)$	x	$F(x)$	x	$F(x)$
0,00	0,0000	0,25	0,0987	0,50	0,1915	0,75	0,2734
0,01	0,0040	0,26	0,1026	0,51	0,1950	0,76	0,2764
0,02	0,0080	0,27	0,1064	0,52	0,1985	0,77	0,2794
0,03	0,0120	0,28	0,1103	0,53	0,2019	0,78	0,2823
0,04	0,0160	0,29	0,1141	0,54	0,2054	0,79	0,2852
0,05	0,0199	0,30	0,1179	0,55	0,2088	0,80	0,2881
0,06	0,0239	0,31	0,1217	0,56	0,2123	0,81	0,2910
0,07	0,0279	0,32	0,1255	0,57	0,2157	0,82	0,2939
0,08	0,0319	0,33	0,1293	0,58	0,2190	0,83	0,2967
0,09	0,0359	0,34	0,1331	0,59	0,2224	0,84	0,2995
0,10	0,0398	0,35	0,1368	0,60	0,2257	0,85	0,3023
0,11	0,0438	0,36	0,1406	0,61	0,2291	0,86	0,3051
0,12	0,0478	0,37	0,1443	0,62	0,2324	0,87	0,3078
0,13	0,0517	0,38	0,1480	0,63	0,2357	0,88	0,3106
0,14	0,0557	0,39	0,1517	0,64	0,2389	0,89	0,3133
0,15	0,0596	0,40	0,1554	0,65	0,2422	0,90	0,3159
0,16	0,0636	0,41	0,1591	0,66	0,2457	0,91	0,3186
0,17	0,0675	0,42	0,1628	0,67	0,2486	0,92	0,3212
0,18	0,0714	0,43	0,1664	0,68	0,2517	0,93	0,3238
0,19	0,0753	0,44	0,1700	0,69	0,2549	0,94	0,3264
0,20	0,0793	0,45	0,1736	0,70	0,2580	0,95	0,3289
0,21	0,0832	0,46	0,1772	0,71	0,2611	0,96	0,3315
0,22	0,0871	0,47	0,1808	0,72	0,2642	0,97	0,3340
0,23	0,0910	0,48	0,1844	0,73	0,2673	0,98	0,3365
0,24	0,0948	0,49	0,1879	0,74	0,2703	0,99	0,3389

x	$F(x)$	x	$F(x)$	x	$F(x)$	x	$F(x)$
1,00	0,3485	1,30	0,4032	1,60	0,4452	1,90	0,4713
1,01	0,3438	1,31	0,4049	1,61	0,4463	1,91	0,4719
1,02	0,3461	1,32	0,4066	1,62	0,4474	1,92	0,4726
1,03	0,3485	1,33	0,4082	1,63	0,4484	1,93	0,4732
1,04	0,3508	1,34	0,4099	1,64	0,4495	1,94	0,4738
1,05	0,3531	1,35	0,4115	1,65	0,4505	1,95	0,4744
1,06	0,3554	1,36	0,4131	1,66	0,4515	1,96	0,4750
1,07	0,3577	1,37	0,4147	1,67	0,4525	1,97	0,4756
1,08	0,3599	1,38	0,4162	1,68	0,4535	1,98	0,4761
1,09	0,3621	1,39	0,4177	1,69	0,4545	1,99	0,4767
1,10	0,3643	1,40	0,4192	1,70	0,4554	2,00	0,4772
1,11	0,3665	1,41	0,4207	1,71	0,4564	2,10	0,4821
1,12	0,3686	1,42	0,4222	1,72	0,4573	2,20	0,4861
1,13	0,3708	1,43	0,4236	1,73	0,4582	2,30	0,4893
1,14	0,3729	1,44	0,4251	1,74	0,4591	2,40	0,4918
1,15	0,3749	1,45	0,4265	1,75	0,4599	2,50	0,4938
1,16	0,3770	1,46	0,4279	1,76	0,4608	2,60	0,4953
1,17	0,3790	1,47	0,4292	1,77	0,4616	2,70	0,4965
1,18	0,3810	1,48	0,4306	1,78	0,4625	2,80	0,4974
1,19	0,3830	1,49	0,4319	1,79	0,4633	2,90	0,4981
1,20	0,3849	1,50	0,4332	1,80	0,4641	3,00	0,4986
1,21	0,3869	1,51	0,4345	1,81	0,4649	3,20	0,4993
1,22	0,3888	1,52	0,4357	1,82	0,4656	3,40	0,4996
1,23	0,3907	1,53	0,4370	1,83	0,4664	3,60	0,4998
1,24	0,3925	1,54	0,4382	1,84	0,4671	3,80	0,49993
1,25	0,3944	1,55	0,4394	1,85	0,4678	4,00	0,499968
1,26	0,3962	1,56	0,4406	1,86	0,4686	4,50	0,499997
1,27	0,3980	1,57	0,4418	1,87	0,4693	5,00	0,49999997
1,28	0,3997	1,58	0,4429	1,88	0,4699		
1,29	0,4015	1,59	0,4441	1,89	0,4706		

Приложение 2. Таблица значений двоичного логарифма $-\log_2 p$

p	$-\log_2 p$	p	$-\log_2 p$	p	$-\log_2 p$	p	$-\log_2 p$
0,01	6,6439	0,26	1,9434	0,51	0,9714	0,76	0,3960
0,02	5,6439	0,27	1,8890	0,52	0,9434	0,77	0,3770
0,03	5,0589	0,28	1,8365	0,53	0,9159	0,78	0,3584
0,04	4,6439	0,29	1,7859	0,54	0,8890	0,79	0,3401
0,05	4,3220	0,30	1,7370	0,55	0,8625	0,80	0,3220
0,06	4,0589	0,31	1,6897	0,56	0,8365	0,81	0,3040
0,07	3,8365	0,32	1,6439	0,57	0,8110	0,82	0,2863
0,08	3,6439	0,33	1,5995	0,58	0,7859	0,83	0,2688
0,09	3,4739	0,34	1,5564	0,59	0,7612	0,84	0,2515
0,10	3,3219	0,35	1,5146	0,60	0,7370	0,85	0,2345
0,11	3,1844	0,36	1,4739	0,61	0,7131	0,86	0,2176
0,12	3,0589	0,37	1,4344	0,62	0,6897	0,87	0,2009
0,13	2,9434	0,38	1,3959	0,63	0,6666	0,88	0,1844
0,14	2,8365	0,39	1,3584	0,64	0,6439	0,89	0,1681
0,15	2,7370	0,40	1,3219	0,65	0,6215	0,90	0,1520
0,16	2,6439	0,41	1,2863	0,66	0,5995	0,91	0,1361
0,17	2,5564	0,42	1,2515	0,67	0,5778	0,92	0,1203
0,18	2,4739	0,43	1,2176	0,68	0,5564	0,93	0,1047
0,19	2,3959	0,44	1,1844	0,69	0,5353	0,94	0,0893
0,20	2,3219	0,45	1,1520	0,70	0,5146	0,95	0,0740
0,21	2,2515	0,46	1,1203	0,71	0,4941	0,96	0,0589
0,22	2,1844	0,47	1,0893	0,72	0,4739	0,97	0,0439
0,23	2,1203	0,48	1,0589	0,73	0,4540	0,98	0,0291
0,24	2,0589	0,49	1,0291	0,74	0,4344	0,99	0,0145
0,25	2,0000	0,50	1,0000	0,75	0,4150	1,00	0,0000

Приложение 3. Список примитивных неприводимых многочленов с минимальным числом ненулевых коэффициентов

Степень m	Многочлены $g_i(x)$, принадлежащие максимальному показателю $e = 2^m - 1$, т.е. $x^e + 1 \equiv 0 \bmod g_i(x)$
1	$x+1$
2	x^2+x+1
3	$x^3+x+1,\ x^3+x^2+1$
4	$x^4+x+1,\ x^4+x^3+1$
5	$x^5+x^2+1,\ x^5+x^3+1$
6	$x^6+x+1,\ x^6+x^5+1$
7	$x^7+x+1,\ x^7+x^3+1,\ x^7+x^4+1$
8	$x^8+x^4+x^3+x^2+1,\ x^8+x^6+x^5+x^3+1$
9	$x^9+x^4+1,\ x^9+x^5+1$
10	$x^{10}+x^3+1,\ x^{10}+x^7+1$
11	$x^{11}+x^2+1,\ x^{11}+x^9+1$
12	$x^{12}+x^6+x^4+x+1,\ x^{12}+x^{11}+x^8+x^6+1$
13	$x^{13}+x^4+x^3+x+1,\ x^{13}+x^{12}+x^{10}+x^9+1$
14	$x^{14}+x^{10}+x^6+x+1,\ x^{14}+x^{13}+x^6+x^4+1$
15	$x^{15}+x+1,\ x^{15}+x^{14}+1$
16	$x^{16}+x^{12}+x^3+x+1,\ x^{16}+x^{15}+x^{13}+x^4+1$
17	$x^{17}+x^3+1,\ x^{17}+x^{14}+1$
18	$x^{18}+x^7+1,\ x^{18}+x^{11}+1$
19	$x^{19}+x^5+x^2+x+1,\ x^{19}+x^{18}+x^{17}+x^{14}+1$
20	$x^{20}+x^3+1,\ x^{20}+x^{17}+1$

Степень m	Многочлены $g_i(x)$, принадлежащие максимальному показателю $e = 2^m - 1$, т.е. $x^e + 1 \equiv 0 \bmod g_i(x)$
21	$x^{21} + x^2 + 1$, $x^{21} + x^{19} + 1$
22	$x^{22} + x + 1$, $x^{22} + x^{21} + 1$
23	$x^{23} + x^5 + 1$, $x^{23} + x^{18} + 1$
24	$x^{24} + x^7 + x^2 + x + 1$, $x^{24} + x^{23} + x^{22} + x^{17} + 1$
25	$x^{25} + x^3 + 1$, $x^{25} + x^{22} + 1$
26	$x^{26} + x^6 + x^2 + x + 1$, $x^{26} + x^{25} + x^{24} + x^{20} + 1$
27	$x^{27} + x^5 + x^2 + x + 1$, $x^{27} + x^{26} + x^{25} + x^{22} + 1$
28	$x^{28} + x^3 + 1$, $x^{28} + x^{25} + 1$
29	$x^{29} + x^2 + 1$, $x^{29} + x^{27} + 1$
30	$x^{30} + x^{23} + x^2 + x + 1$, $x^{30} + x^{29} + x^{28} + x^7 + 1$
31	$x^{31} + x^3 + 1$, $x^{31} + x^{28} + 1$
32	$x^{32} + x^{22} + x^2 + x + 1$, $x^{32} + x^{31} + x^{30} + x^{10} + 1$
33	$x^{33} + x^{13} + 1$, $x^{33} + x^{20} + 1$

Приложение 4. Разложение двучлена $x^n + 1$ на неприводимые сомножители над полем $GF(2)$

Двучлен $x^n + 1$	Неприводимые сомножители
$x + 1$	$x + 1$
$x^2 + 1$	$(x + 1)^2$
$x^3 + 1$	$(x + 1)(x^2 + x + 1)$
$x^4 + 1$	$(x + 1)^4$
$x^5 + 1$	$(x + 1)(x^4 + x^3 + x^2 + x + 1)$

Двучлен x^n+1	Неприводимые сомножители
x^6+1	$(x+1)^2(x^2+x+1)^2$
x^7+1	$(x+1)(x^3+x+1)(x^3+x^2+1)$
x^8+1	$(x+1)^8$
x^9+1	$(x+1)(x^2+x+1)(x^6+x^3+1)$
$x^{10}+1$	$(x+1)^2(x^4+x^3+x^2+x+1)^2$
$x^{11}+1$	$(x+1)(x^{10}+x^9+x^8+x^7+x^6+x^5+$ $+x^4+x^3+x^2+x+1)$
$x^{12}+1$	$(x+1)^4(x^2+x+1)^4$
$x^{13}+1$	$(x+1)(x^{12}+x^{11}+x^{10}+x^9+x^8+x^7+x^6+x^5+$ $+x^4+x^3+x^2+x+1)$
$x^{14}+1$	$(x+1)^2(x^3+x+1)^2(x^3+x^2+1)^2$
$x^{15}+1$	$(x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)\times$ $\times(x^4+x^3+x^2+x+1)$
$x^{16}+1$	$(x+1)^{16}$
$x^{17}+1$	$(x+1)(x^8+x^7+x^6+x^4+x^2+x+1)\times$ $\times(x^8+x^5+x^4+x^3+1)$
$x^{18}+1$	$(x+1)^2(x^2+x+1)^2(x^6+x^3+1)^2$
$x^{19}+1$	$(x+1)(x^{18}+x^{17}+x^{16}+x^{15}+x^{14}+x^{13}+x^{12}+$ $+x^{11}+x^{10}+x^9+x^8+x^7+x^6+x^5+x^4+$ $+x^3+x^2+x+1)$
$x^{20}+1$	$(x+1)^4(x^4+x^3+x^2+x+1)^4$
$x^{21}+1$	$(x+1)(x^2+x+1)(x^3+x+1)(x^3+x^2+1)\times$ $\times(x^6+x^5+x^4+x^2+1)(x^6+x^4+x^2+x+1)$

Двучлен x^n+1	Неприводимые сомножители
$x^{22}+1$	$(x+1)^2(x^{10}+x^9+x^8+x^7+x^6+x^5+$ $+x^4+x^3+x^2+x+1)^2$
$x^{23}+1$	$(x+1)(x^{11}+x^9+x^7+x^6+x^5+x+1)\times$ $\times(x^{11}+x^{10}+x^6+x^5+x^4+x^2+1)$
$x^{24}+1$	$(x+1)^8(x^2+x+1)^8$
$x^{25}+1$	$(x+1)(x^4+x^3+x^2+x+1)\times$ $\times(x^{20}+x^{15}+x^{10}+x^5+1)$
$x^{26}+1$	$(x+1)^2(x^{12}+x^{11}+x^{10}+x^9+x^8+x^7+x^6+x^5+$ $+x^4+x^3+x^2+x+1)^2$
$x^{27}+1$	$(x+1)(x^2+x+1)(x^6+x^3+1)(x^{18}+x^9+1)$
$x^{28}+1$	$(x+1)^4(x^3+x+1)^4(x^3+x^2+1)^4$
$x^{29}+1$	$(x+1)(x^{28}+x^{27}+x^{26}+x^{25}+x^{24}+x^{23}+x^{22}+$ $+x^{21}+x^{20}+x^{19}+x^{18}+x^{17}+x^{16}+x^{15}+x^{14}+$ $+x^{13}+x^{12}+x^{11}+x^{10}+x^9+x^8+x^7+x^6+x^5+$ $+x^4+x^3+x^2+x+1)$
$x^{30}+1$	$(x+1)^2(x^2+x+1)^2(x^4+x+1)^2(x^4+x^3+1)^2\times$ $\times(x^4+x^3+x^2+x+1)^2$
$x^{31}+1$	$(x+1)(x^5+x^2+1)(x^5+x^3+1)\times$ $\times(x^5+x^3+x^2+x+1)(x^5+x^4+x^2+x+1)\times$ $\times(x^5+x^4+x^3+x+1)(x^5+x^4+x^3+x^2+1)$

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Евгений Феофанович Березкин

ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ И КОДИРОВАНИЯ

Учебное пособие

Редактор Е.Е. Шумакова
Оригинал-макет изготовлен Е.Ф. Березкиным

Подписано в печать 08.12.2009. Формат 60x84 1/16.
Печ.л. 19,5. Уч.-изд.л. 19,5. Тираж 100 экз.
Изд. № 1/1/4. Заказ № 3

Национальный исследовательский ядерный университет «МИФИ».
115409, Москва, Каширское шоссе, д.31.

ООО «Полиграфический комплекс «Курчатовский».
144000, Московская область, г. Электросталь, ул. Красная, д. 42